

DEPARTAMENT DE METAFISICA I TEORIA DEL
CONEIXEMENT

CARÁCTER, CIRCUNSTANCIAS Y ACCIÓN. EL PAPEL DE
LA SUERTE EN LA DETERMINACIÓN DE LA
RESPONSABILIDAD MORAL.

SERGI ROSELL TRAVER

UNIVERSITAT DE VALÈNCIA
Servei de Publicacions
2009

Aquesta Tesi Doctoral va ser presentada a València el dia 17 de juliol de 2009 davant un tribunal format per:

- Dr. Carlos Thiebaut Luis-André
- Dr. Josep Corbí Fernández de Ybarra
- Dr. José Luis Prades Celma
- Dra. María Álvarez Alonso
- Dr. Vicente Sanfèlix Vidarte

Va ser dirigida per:
Dr. Carlos Moya Espí

©Copyright: Servei de Publicacions
Sergi Rosell Traver

Dipòsit legal: V-4178-2010

I.S.B.N.: 978-84-370-7701-7

Edita: Universitat de València
Servei de Publicacions
C/ Arts Gràfiques, 13 baix
46010 València
Spain
Telèfon:(0034)963864115

Universitat de València
Facultat de Filosofia i Ciències de l'Educació
Departament de Metafísica i Teoria del Coneixement



CARÁCTER, CIRCUNSTANCIAS Y ACCIÓN
EL PAPEL DE LA SUERTE EN LA DETERMINACIÓN DE LA
RESPONSABILIDAD MORAL

Tesis Doctoral

Presentada al título de Doctor en Filosofía,
con la mención de “Doctor Europeus”

Doctorando: Sergi Rosell Traver
Director: Carlos Moya Espí
Programa de Doctorado: Razón, Lenguaje e Historia
(627 165F)

2009

**Departament de Metafísica i Teoria del Coneixement
Universitat de València**

Tesis Doctoral:

**CARÁCTER, CIRCUNSTANCIAS Y ACCIÓN
EL PAPEL DE LA SUERTE EN LA DETERMINACIÓN DE LA
RESPONSABILIDAD MORAL**

**Presentada al título de Doctor en Filosofía,
con la mención de “Doctor Europeus”**

**Doctorando: Sergi Rosell Traver
Director: Carlos Moya Espí
Programa de Doctorado: Razón, Lenguaje e Historia (627 165F)**

**Departament de Metafísica i Teoria del Coneixement
Universitat de València**

PhD Dissertation:

**CHARACTER, CIRCUMSTANCES, AND ACTION
THE ROLE OF LUCK IN DETERMINING MORAL RESPONSIBILITY**

**In Candidacy for the Degree of Doctor of Philosophy,
With the Title of “Doctor Europeus”**

**Candidate: Sergi Rosell Traver
Supervisor: Carlos Moya Espí
PhD Graduate Programme: Reason, Language, and History (627 165F)**

2009

© Sergi Rosell Traver
Universitat de València
València, 2009

En memoria de mi querido hermano, Javi

Contenidos

Agradecimientos	9
Introduction	15
Précis (Resumen)	I

PARTE I

Capítulo 1. Las atribuciones de responsabilidad	23
Capítulo 2. La suerte moral. <i>Williams y Nagel sobre la suerte moral</i>	39
Apéndice: <i>La suerte moral y el debate sobre el libre albedrío y la responsabilidad moral</i>	81

PARTE II

Capítulo 3. El caso contra la suerte moral. <i>Articulación y respuesta</i>	93
Capítulo 4. La suerte constitutiva. <i>Sobre necesidad y contingencia en nuestros orígenes e identidad</i>	145
Capítulo 5. La suerte formativa. <i>Responsabilidad por el carácter</i>	179
Capítulo 6. Suerte circunstancial. <i>Atrapados por las circunstancias</i>	239
Capítulo 7. Suerte resultante. <i>Justificación retrospectiva y emociones morales</i>	299

PARTE III

Capítulo 8. El rechazo del control y sus consecuencias	359
Capítulo 9. Teoría, práctica y choque de intuiciones	413
Conclusions (Conclusiones)	451
Bibliografía	463
Índice detallado	487

Agradecimientos

El proyecto de escribir *esta* tesis empezó allá por la primavera de 2004, si bien la idea de escribir *una* tesis fue algo anterior. Ya en 2002, durante mi estancia en Birmingham, traté de hallar, sin éxito, un tema de investigación, al que dedicar unos cuantos años de mi vida. Por lo menos, descubrí algo —que no es poco—: mi pasión por la *manera* analítica de hacer filosofía y mi interés por los temas *prácticos*. Posteriormente, de vuelta en Valencia, asistí al curso de doctorado impartido por Josep Corbí y Julián Marrades, que combinaba de un modo brillante cuestiones conceptuales —típicamente abstractas—, entre las que se hallaba sin duda el tema de la suerte moral, con datos empíricos —hechos sociales e históricos— y ejemplos literarios, que situaban a las primeras en un contexto real y concreto, dotándolas así de una especial vitalidad. No hay duda de que este modo de abordar los problemas filosóficos, que también he encontrado en California, ha influido en mi propia práctica.

A principios del verano de 2004, tras serme concedida la Beca V Segles, establecí un contacto más estrecho con Carlos Moya, que había aceptado ser el director de esta tesis. Carlos me propuso diversos temas, relacionados con los intereses más bien imprecisos que le había expresado, en los que me garantizaba su competencia. Realmente, no era sólo *competencia*, como me he podido dar cuenta durante estos años de trabajo en común, sino un minucioso conocimiento y un hábil dominio de muchos de los diversos temas que integran lo que me gusta llamar *filosofía de la agencia*; un área, que es en realidad una confluencia de distintas

disciplinas filosóficas, que ha despertado un inusitado entusiasmo en mí. Fue entonces cuando pude empezar a vislumbrar las conexiones de la cuestión de la suerte moral con otros problemas pertenecientes a la filosofía de la acción, el problema del libre albedrío y la responsabilidad moral, la filosofía moral, la filosofía de la mente, la epistemología, la metafísica, etc.; y cuando, finalmente, me decidí por el tema al que he acabado dedicando esta investigación doctoral.¹

Por supuesto, es mucha la gente que, filosófica o extrafilosóficamente y en mayor o menor medida, ha contribuido a hacer posible esta tesis.

En primer lugar, me gustaría referirme a los miembros del grupo Phrónesis de Valencia, con los que he podido compartir tantos seminarios, y de los que sin duda he aprendido mucho, como son Josep Corbí, Álex Fontcuberta, Manolo Garcés, Miracle Garrido, Tobies Grimaltos, Javier Hernández Iglesias, Marta Moreno, Carlos Moya, Eduardo Ortiz, José Luis Pérez, Lino San Juan y Jordi Valor. Es indudable que mi pensamiento filosófico se ha ido formando, en gran medida, en armonía y contraste con el de otros miembros del grupo. También me ha sido de gran ayuda el contacto con Fernando Broncano, Manuel García-Carpintero, Genoveva Martí, Josep Lluís Prades, Jennifer Saul, Gianfranco Soldati y Christine Tappolet, miembros todos ellos de Nomos Network. Quiero, asimismo, citar los nombres de Óscar Benito, Noemí Calabuig, Salvador Cuenca, Alberto J. Gómez, Javier Hernández Iglesias, Mi-

¹ La financiación principal para la realización de esta tesis ha procedido de la Beca de Investigación Predoctoral “V Segles”, otorgada por la Universitat de València. Asimismo, la participación en los proyectos de investigación “Realismo, deliberación y verdad” (BFF2000-1073-C04-03) y “Creencia, responsabilidad y acción” (HUM2006-04907), financiados por el Ministerio de Educación y Ciencia del Gobierno de España, me ha aportado fondos complementarios. No puedo más que expresar mi sincera gratitud para con ambos organismos.

riam Hoyo, Joan B. Llinares, Julián Marrades, Bernat Martí, Joan David Mateu, Vicente Raga, Begoña Ramón, Vicente Sanfélix y Vanessa Vidal, amigos y miembros (o ex miembros, como yo mismo) del Departament de Metafísica i Teoria del Coneixement, así como a Enric Casaban, Vicente Claramonte y Valeriano Iranzo, también amigos, pero del Departament de Lògica i Filosofia de la Ciència. Guardo, asimismo, un recuerdo muy especial del profesor Josep Lluís Blasco y del último curso de doctorado que pudo impartir, y al que tuve el honor de asistir.

Agradezco especialmente a Josep Corbí, Salva Cuenca, Tobies Grimaltos y Eduardo Ortiz que leyeron y me comentasen diferentes capítulos de esta tesis. Ni que decir tiene que mi deuda con Carlos Moya, que leyó todo el manuscrito y me comentó detenidamente cada uno de sus párrafos, es enorme. Nuestro desacuerdo fundamental, pero cordial, respecto a la posición correcta en el tema de esta tesis, más que un obstáculo, ha sido extraordinariamente estimulante y enriquecedor.

Otras personas que me he ido encontrado en seminarios, congresos y departamentos de filosofía de otras universidades y que, en modos y grados muy diversos, me han ayudado (algunos sin que seguramente lo sospechen), son María José Alcaraz, Paloma Atencia, David Brink, Óscar Cabaco, Mario De Caro, José Chaves, George Couvalis, Esa Díaz-León, Ángel García, Manuel Hernández Iglesias, Christopher Hookway, Max Kölbel, Diego Lawler, Frank Lihoreau, Josep Macià, Roberto Mordacci, Dana Nelkin, Carlo Penco, Thomas Pink, Erick Ramírez, Luis Robledo, Sònia Roca, Esther Romero, David Segarra, Jesús Vega y Gary Watson. A Dana Nelkin le agradezco especialmente que me aceptara como *visiting graduate* en el departamento de filosofía de la Universidad de California, San Diego, así como que leyera y comentara diversos materiales que forman parte de esta tesis. A Roberto Mordacci le agradezco su comentario a

mi presentación de materiales de esta tesis en el IV Latin Meeting on Analytic Philosophy (Génova, septiembre de 2007); también a Esa Díaz-Leon, su comentario para el X Taller d'Investigació en Filosofia (Barcelona, enero de 2008); y a Ricardo Caracciolo, el suyo para el II Encuentro Hispano-Argentino de Filosofía Analítica (Buenos Aires, septiembre de 2008).

Una versión previa del capítulo 2 fue publicada como “Nagel y Williams acerca de la suerte moral”, en *Revista de Filosofía* 31 (2006), 1: 143-165. Presenté una versión simplificada de las dos primeras secciones del capítulo 3 en el V Congreso de la SEFA y en el IV Latin Meeting for Analytic Philosophy; que apareció en actas como “Is the Case Against Moral Luck Successful?”, C. Penco *et al.* (eds.) *Proceedings of the IV Latin Meeting in Analytic Philosophy*. Génova: 2007, pp. 33-44. Materiales que integran el capítulo 4 fueron presentados en un congreso en Braga (Portugal) y en otro en Berlín, y publicados en las actas de éste último como “Constitutive Luck and Conceivable Identity”, en H. Bohse and S. Walter (eds.) *Selected Contributions to GAP.6, Sixth International Conference of the Society for Analytical Philosophy, Berlin, 11–14 September 2006*. Paderborn: mentis Verlag, 2007, pp. 1031-1040. En el XVII Congrés Valencià de Filosofia, celebrado en Valencia en la primavera de 2008, y en el II Encuentro Hispano-Argentino de Filosofía Analítica, que tuvo lugar el otoño del mismo año en Buenos Aires, leí una versión de las dos últimas secciones del capítulo 5 con el título de “Carácter, conocimiento y responsabilidad”. Las secciones del capítulo 6 en las que me ocupé del escepticismo sobre los rasgos de carácter han sido presentadas en diferentes congresos celebrados en Granada, Ginebra, Milán y Barcelona. Un antecedente, bastante superado, de parte del capítulo 7 es la po-

12

nencia “Un postimpresionista pels Mars del Sud: Williams i la retrospectivitat de la justificació”, presentada en el XVI Congrés Valencià de Filosofia (2006). Las diferentes reflexiones acerca de la fragilidad del juicio moral, la historia del piloto Claude Eatherly y las sutilezas de la psicología moral del mal, que he expuesto en varios congresos, también han inspirado algunos párrafos de los capítulos 7 y 8. Finalmente, la última sección del capítulo 8 es básicamente lo contenido en “¿Es la idea de responsabilidad moral culturalmente variable?”, artículo publicado en A. Alonso *et al.* (eds.), *Surcar la cultura*. Valencia: Pre-Textos, 2006.

Durante la realización de este trabajo, he recibido un apoyo muy especial de mi familia: de mis padres, Toni y Tere, y de mi hermano, Javi. Sin el cariño constante de los tres, junto con la afectiva y paciente comprensión de mi novia Eva, las cosas hubieran sido, por decir poco, mucho más difíciles.

* * *

Desgraciadamente, en el último momento, me veo obligado a añadir algo más. La finalización de esta tesis estará siempre trágicamente unida a un suceso profundamente doloroso que golpeó nuestras vidas: el 16 de noviembre de 2008, Javi perdió la vida en un accidente de tráfico.

Aristóteles dijo que la felicidad (*eudaimonia*) trasciende la propia vida. En concreto, afirmó que “parece que para el hombre muerto existen un mal y un bien —lo mismo que existen para el que vive, pero no se da cuenta” (EN I, 10). Algunos de nuestros deseos y expectativas, especialmente aquellos que tienen que ver con la buena o mala fortuna de nuestros seres queridos, pueden ser satisfechos tras nuestra muerte. De este

modo, si el filósofo está en lo cierto, la felicidad de una persona podría aumentar o disminuir póstumamente. El mismo Aristóteles reconoció el carácter un tanto paradójico de esta posición. Yo mismo no me siento mucho más persuadido por ella que él —ciertamente, la idea de placer *sentido*, o de satisfacción *experimentada*, tiene un peso mucho mayor en nuestra noción moderna de felicidad. Pero ante la posibilidad de que resulte ser cierta, desearía que la finalización de este trabajo, en el que he invertido mucho tiempo, esfuerzo e ilusiones, contribuyera a aumentar la felicidad de mi querido hermano Javi. Dedico esta tesis a su memoria.

Valencia, mayo de 2009

Introduction

The main topic of this dissertation is the so-called phenomenon of “moral luck”. This was a new phrase, introduced in the late seventies of the twentieth century by Bernard Williams and Thomas Nagel, for an old issue that goes back to the Greeks. Indeed, it may be the case that the very idea of moral luck was less difficult to accept for the Greek conception of agency and ethics than it is for our modern conception.

In general, the notion of “moral luck” may sound wrong. If there is something which seems to us especially misleading regarding our general purpose of being fair in treating people, or our wish of living in a just world, that appears to be luck. Undeniably luck intervenes in our ordinary lives in very different ways, to the extent that it is not easy to imagine a world with no luck. But, in spite of that, we feel obligated to find out ways of neutralizing it. What is more, we tend to believe that if something does not depend on the agent, it is unfair to hold her morally responsible for it.

To be sure, we commonly assume that admired and admirable intellectual, aesthetical, or athletic talents, gifts, or even achievements, hinge partially on accident and luck. We all acknowledge that they depend on biology, upbringing, education, social environment, opportunities, and so on; but tend to deny that those factors can make a moral difference. We wonder what is so special about morality—why moral status should be beyond the influence of luck?

The issue of moral luck is usually motivated as a puzzle resulting of a tension between the belief that we cannot blame or praise differently two or more people due to things beyond their control, and what seems to be our ordinary practice of judging people differently regarding what actually happens. Here you have a prototypical case showing that sort of tension. Two people, after drinking in a pub, decide to come back home by driving their respective cars. In the way, one of them loses control of her car, comes off the road, hits a pedestrian and runs him over. The other driver equally loses control of her car, comes off the road, but neither hits a pedestrian nor runs anyone over because there was nobody there. It appears from this picture that depending on something that is beyond the control of either agent, just one of them will be responsible for a death and will putatively deserve more blame; whereas the other, even though she is equally at fault or makes the same mistake, will be judged with more leniency and will not be responsible for killing anybody. In the upshot, it seems that the latter is morally luckier than the former. However, this outcome collides with what appears as a *prima facie* principle ruling attributions of moral responsibility:

(CP) An agent *A* is to be morally responsible for *x* only if she has control over *x*.

In particular, it seems that it is a consequence of our idea of fairness that we ought not to judge persons differently for *doing the same thing* or on the basis of factors beyond their control. A corollary of the Control Principle would be like this:

(Col) People ought not to be morally assessed differently if the only differences between them are due to factors beyond their control.

In the face of this clash of intuitions, a feeling of perplexity, as a first reaction, arises.

Yet, the issue is broader than it seems. Moral luck is a wide-ranging phenomenon that extends beyond our assessment of the consequences of certain actions—what has been called *consequential* or *resultant* moral luck. It may also affect our assessment regarding an agent having to face some relevant circumstances, or having received some influences, and not others, or possessing a certain constitution. These other kinds of moral luck will be described briefly; but we can already present what is the main discrepancy about.

The debate is, firstly, about the very existence of such a phenomenon. Is it moral luck a real phenomenon or a mere appearance? This question is more difficult to answer than it seems at first sight, especially in the light of the different kinds of moral luck proposed. And, secondly, if there is a real conflict between our practices of moral judgment, or particular intuitions, and what we think are the principles ruling them, or general intuitions, then something appears to be wrong; and, then, either our practices or our principles are likely in need of revision.

Therefore, an option is to try to deny the very existence of moral luck as a real phenomenon, as numerous philosophers have done. This side argues for a convenient reinterpretation of what appeared to be a genuine conflict. However, another strategy against moral luck will consist in accepting the existence of such a phenomenon in our practices, that we judge people for things beyond their control, but denouncing the correctness of those practices. On the other side, philosophers argue for the

reality of moral luck. And that can be done either by accepting the impossibility of getting rid of any of the conflicting intuitions, so reaffirming the initial perplexity—as Nagel does, to the extent of defending the existence of a real paradox—or by favoring our common practices and denying a true role for the Control Principle.

Roughly, that is the picture. In this dissertation I will argue for the reality of moral luck while denying that such a phenomenon entails a real paradox.

The Dissertation has three parts. Part I is devoted to explain what the phenomenon of moral luck, if real, would be about. In Part II a reconstruction and rebuttal of the main arguments against moral luck is offered. And, finally, Part III answers significant questions that arise after the recognition of the fact that luck can make differences in judging people morally. I will claim that luck necessarily interferes in our moral agency, so that it is impossible to isolate attributions of moral responsibility from luck; in addition, there exist important normative considerations that favor preferring a world with moral luck to one without it. No doubt others have also offered arguments of these two sorts—descriptive and normative—in favor of moral luck; nevertheless I will offer here new arguments and new defenses of others' arguments. It follows a précis of the contents of the dissertation, in which the particularities of my own arguments are stressed.

Before moving on that, a brief remark on methodology is not irrelevant. In this dissertation I make use of two kinds of extra-philosophical sources. On the one hand, I draw on some empirical literature, mainly on those findings in social psychology that have recently been employed by some philosophers to undermine our notion of charac-

ter traits, but also upon some surveys of folk intuitions brought to light the last years by experimental philosophers. Both sets of empirical data will be relevant for the discussions in which they will be used. In general, the utilization of empirical data follows from the conviction—which I share—that philosophical claims must meet the pertinent empirical constraints.

On the other hand, biographical case studies, taken from history or literature, are also adduced to illustrate particular points. These cases add some flesh to abstract claims. In particular, resorting to literary passages is usually profitable given that novelists often anticipate philosophical ideas, especially when philosophers look for descriptions more than explanations.

Précis¹

Part I

As advanced, the first part of this dissertation aims at characterizing the phenomenon of moral luck. In Chapter 1, I begin by offering a rough depiction of what can be called *the system of moral responsibility*, i.e. the practices, reactive attitudes, beliefs, concepts, ruling principles, and so on, associated to our attributions of moral responsibility. A quite neutral characterization, faithful to general folk intuitions, is intended, in order to frame our subsequent discussion.

Ordinarily, we hold ourselves and others responsible for what we and they do or, more exactly, for part of what we and they do. Being responsible is being accountable for one's own acts. I propose the following basic scheme for responsibility attributions:

(AR) *A* is responsible for *O*, in virtue of *C*.

Where *A* is the agent who bears the responsibility attribution, also called *locus* or *perpetrator*; *O* is the object of responsibility, that thing for what responsibility is attributed, also known as the *content* of responsibility; and *C* are the conditions for attributing responsibility to *A*. In order to be considered a moral agent, in the relevant sense of this term that justifies

¹ This is a quite detailed outline of each chapter's content, especially intended for *Doctor Europeus'* international evaluators. Then, regular evaluators and general (Spanish-language) readers may skip this précis and proceed to go deep into the main text.

attributions of moral responsibility, one has to possess certain psychological capacities, i.e. the capacities for understanding and reasoning, as well as capacities for forming desires, intentions and plans, for deliberating, making decisions, and acting according to reasons. Those are necessary conditions for agency. In general, it can be distinguished between a *reflective power*, i.e. those cognitive and volitive capacities that play a role in assessing and selecting actions; and a *self-control power*, i.e. the capacities that motivate and direct one's behavior in virtue of the relevant cognition and volition.

So, to be rightly considered as a responsible agent, it seems that it is necessary that one possesses these powers or capacities. Moreover, the existence of a causal link (in a broad sense) between the agent and the object of responsibility is generally required for responsibility attributions. It seems that causal responsibility is also a necessary condition for moral responsibility. Indeed, moral responsibility is a subclass of what can be called personal or normative responsibility, in which we can include, as well as moral responsibility, legal responsibility, epistemic responsibility, and so on. Every kind of responsibility is defined by its particular criteria for the relevant responsibility attribution—moral, legal, epistemic or whatever.

On the other hand, attributions of moral responsibility involve a particular sort of response. Being morally responsible for an action appears to involve that a particular reaction is deserved—although it may not be justified in a particular context. We tend to think that people who behave rightly and show admirable traits deserve praise and reward, whereas people who act wrongly deserve blame and punishment. In other words, our practice of attributing responsibility to persons is linked to holding them to deserve praise or blame, reward or punishment, gratitude

or resentment; responses that constitute a complex network of concepts round the idea of moral responsibility.

Something about C conditions must also be said. These conditions are referred to particular attributions of moral responsibility, and not (as the previous ones) to general capacities agents must satisfy. Namely, we need to establish further requirements that warrant that the object of attribution is properly attributed to the agent. Those requirements can be characterized, in a way that comes from Aristotle, as two general conditions. One is the epistemic condition:

(EC) An agent liable to attribution of moral responsibility has to have an appropriate degree of relevant knowledge of the circumstances of her behavior.

The other, the control condition:

(CC) An agent liable to attribution of moral responsibility is to be in control of her behavior in an appropriate degree and in the relevant aspects.

That is, the agent does not have to ignore the relevant information about the circumstances, nor lack the appropriate control over her behavior. By merging both conditions, we reach the above-mentioned Control Principle:

(CP) An agent *A* can be morally responsible for *x* only if she has control over *x*.

(In CP, “control” includes both the epistemic (EC) and the freedom aspects (CC) of control.) Then, although CP seems, at first glance, quite

vague, we see that there is a long story about it, which we should *read* in it. Anyway, the main point here is that this principle is generally acknowledged as a necessary condition for attribution of moral responsibility. However, the moral luck phenomenon comes to question this assumption. It is worth highlighting that we do not need a strong notion of control in order to raise our issue; it works with any notion of control.

I finish Chapter 1 by proposing some distinctions in the description of our ordinary idea of the responsibility system that will acquire true meaning in the last part of this dissertation.

Chapter 2 is a detailed presentation of the moral luck issue. I present, examine, and compare Thomas Nagel's and Bernard Williams' approaches to it.² Particularly, I focus on Nagel's account, that contains the most straightforward setting out of the moral luck phenomenon and a sketch of the two main sorts of response to it—roughly, either to opt for practices or to opt for the principle—which have articulated the debate. Although, of course, that does not mean that I am going to ignore either the particularities of Williams' account, or its significance.

It is said that a case of moral luck occurs when we morally judge an agent for something (an action, a trait, etc.) that in a significant aspect depends on factors beyond her control. More generally, what is at issue is whether luck can make *a moral difference*. We saw above a typical example of moral luck, concretely of resultant moral luck, i.e. the drunk drivers' case. But there are (putatively) more sorts of cases, and kinds of moral luck. Consider this other example. John intends to kill someone, but when he is in the position to do it, he misses his shot. His counterpart Sean also goes ahead and performs all previously necessary actions to kill

² Nagel (1979) and Williams (1981).

someone, and finally he does not miss his shot. If our evaluation of each agent varies, if we judge with a higher degree of severity the successful murderer than the unsuccessful one, and the only difference is factors beyond their control, we will have another case of resultant moral luck. *Resultant moral luck* is moral luck in the way that actions or projects of an agent result.

Consider now a putative case of situational luck. Rudolf and Adolf are two German citizens, living in the Nazi era, with Nazi sympathies. Rudolf, because of business, has to move out Germany before Hitler seizes power; whereas Adolf stays in Germany for all the Nazi period. This being so, only Adolf has the opportunity of making his Nazi sympathies effective and becoming, say, an SS member and, eventually, head of a concentration camp. We can stipulate that if the émigré had stayed in Germany he would have acted in the same horrible way. But, do we mean that we have to assess the expatriate businessman Rudolf as harshly as Nazi head Adolf? If we answer “no”, luck will make a moral difference. *Situational moral luck* is the luck of being in one or other place, at one or other time, that can affect the way we are morally judged.

The two remaining Nagelian kinds of moral luck are constitutive and causal moral luck. By constitutive moral luck it is meant luck involved in the fact that people possess different inclinations, capacities and temperament, given that their possession is, at least, not completely in the agent’s power, and the possession of some and not others can make a difference in the way one responds to moral requirements. Finally, causal moral luck is luck in how one is determined by antecedent circumstances. This kind of moral luck has been assimilated to the traditional debate about free will and moral responsibility.

However useful it is, this classification can be modified in order to get a more functional one. In the main text, I compare it mainly with proposals by Williams, Michael Zimmerman and Susan Hurley, and conclude that the most efficacious classification should distinguish between these four main kinds of moral luck: resultant, circumstantial, formative, and constitutive. The two first kinds are common to Nagel's classification. Formative luck refers to upbringing and education, i.e. character formation; all sorts of previous experiences that (with her constitutive traits) made one into the person that one currently is. Constitutive luck is more closely linked to originary traits, those constitutive traits that one received. Moreover, formative and constitutive luck can be brought together under the label of antecedent luck. But also formative and circumstantial luck can be grouped together due to the fact that they are indeed two kinds of luck in circumstances (antecedent and present). In showing the vague differences between kinds, I try to challenge the (more or less implicit) idea that they are independent and clearly distinct ways in which luck can influence our judgments about people.

It should be added that there are some important differences in the way Nagel and Williams present the issue. Nagel stresses the idea that the phenomenon of moral luck involves an insurmountable paradox. The phenomenon is not the result of an arbitrary imposition of stringent standards, due to philosophical or theoretical misunderstandings, but of a consistent application of our ordinary standards of moral judgment. This phenomenon leads to skepticism about moral responsibility and is common to other forms of skepticism. From a different outlook, Williams is mainly concerned with our conception of morality and how this is fundamentally mistaken. Particularly, he raises the issue as an objection to the Kantian conception of morality that, according to him, is primarily our modern

conception. For Kant, moral value is the supreme value and is immune to luck; this is what warrants “its role of solace to the world’s injustice”. In addition, for this view, the notions of morality, rationality, justification and supreme value are intrinsically linked. To thread this conception, Williams sets out to show that moral value cannot be the supreme value and immune to luck, at once. In order to yield this conclusion he presents a case—Gauguin’s choice—in which either moral value is overcome by other kind of value or it is subject to luck. I will further spell out this case in chapter 7, where I will argue for William’s view (with amendments). In general I will try to point out the unsatisfactory character of Nagel’s final response (that we face an insuperable paradox), and also that William’s account can be clarified and defended in a particularly interesting way (to my purposes).

In the final section of this chapter, I present the main options—above mentioned—that philosophers can follow in confronting the issue. And a new difficulty, due to the (putative) existence of different kinds of moral luck, is added. It is worth noticing that an argument intended to reject or defend moral luck may not work equally well (or badly) for each kind of moral luck. So, in order to be successful a case for or against moral luck should be applicable to all kinds of this phenomenon—resultant, circumstantial, formative and constitutive moral luck—, or, alternatively, consist in a combination of strategies. Or even a further option is to advocate a combination of the rejection of some kind(s) of moral luck with the recognition of other(s).

The Appendix to Part I is an attempt to cast some light on the relationship between moral luck and the debate about free will and moral responsibility. Although both topics share important overlapping areas, neither of

them should be seen as a sub-area of the other. Particularly, I reject the idea that moral luck is a subtopic within the free will and moral responsibility literature. Moral luck is indeed a new threat to our concept of moral responsibility, partially independent in many aspects of traditional discussions on free will and determinism. Both topics are not coincident, at least completely, especially regarding the way they are primarily studied. In my opinion, the moral luck issue is more interestingly studied as a topic in moral psychology than as a metaphysical question. Any way, the moral luck issue will remain undetermined even though the problem of free will and the possibility of moral responsibility were eventually solved.

Being faithful to this conclusion, I will try to primarily frame the discussion in moral-psychological terms and avoid the metaphysical problem of skepticism about moral responsibility—although I will have something to say about it in Chapter 8.

Part II

Given that my aim is to vindicate the reality of moral luck, my main opponents will be the *deniers*, i.e. those philosophers who deny the existence of such a phenomenon. In part II, I address directly the most important group of responses to the moral luck problem, according to which the existence of moral luck is an illusion. In Chapter 3 I begin by reconstructing what I call the Global Case against Moral Luck; after that, I will try to reject it. The remaining chapters of this part (4-7) are devoted to each of the four previously distinguished kinds of moral luck. In each chapter the same methodology is followed: I begin by reconstructing the best arguments against the particular kind of moral luck at stake and then I proceed to put forward a rebuttal of them. My overall argumentative

strategy combines a general rejection of the Global Case with a particular defense of every kind of moral luck.

In the face of a case like that of the drunk drivers, the denier has a clear response: if what they strictly controlled (getting drunk and driving, but not the presence or absence of some pedestrians in their way) was the same, then they deserve the same responsibility judgment. The important thing for moral appraisal is what you strictly control—in this case, the action you performed, and not the particular consequences of your action.

The most common rejection of moral luck includes the so-called *Epistemic Argument* (EA), which claims that luck really does not interfere with someone's moral status, but with our knowledge of her/him, given that we are not omniscient beings and our knowledge is mediated by the available evidence. A person may be lucky or unlucky in the transparency of what she/he deserves, but that does not mean that luck can make a moral difference, i.e., can affect *what she/he really or ultimately deserves*.

Additionally, in order to get a Global Case we have to extend the idea to previous kinds of moral luck. In dealing with cases of resultant luck, the denier locates the locus of responsibility in the agents' action, independently of the actual consequences. Now, it is obvious that this resort is not viable for other kinds of moral luck. Not just consequences are different in cases of, say, circumstantial luck, but agents' actual actions are different as well. Then, the denier has to relocate the locus of responsibility in something like agent's intentions, will, dispositions, etc. The central idea can be restated in this way: a person's moral desert is a function of *what she would have done if she had had the relevant chance*.

This *prima facie* intuitively appealing move naturally extends to antecedent (formative and constitutive) moral luck; but now one has even

to renounce to locate the locus of responsibility in the actual intentions, will or dispositions of the agent. This schema by Michael Zimmerman illustrates well the idea:

Given the Control Principle,
If (i) P made decision d in what he believed to be situation s ,
(ii) P* would have made decision d if he had been in a situation that he believed to be s , and
(iii) P*'s being in a situation that he believed to be s was not in his (restricted) control,
Then: whatever moral credit or discredit accrues to P for making d accrues also to P*.³

Both agents, in virtue of counterfactual considerations, would be equally praiseworthy or blameworthy—granting what was called the Corollary of CP. At the end, we reach a notion of desert *essentially counterfactual*, as *a pure function of agent's control*.

In order to rebut the Global Case against moral luck, I begin by distinguishing two broad positions in this general strategy: a moderate and a radical strategy. On the one hand, the Moderate Strategy (MS) tries to vindicate the Control Principle by means of EA. Moreover, MS defenders are characteristically committed to what I call the *Conservative Claim*: that we can acknowledge the error EA unmasks and also maintain our practices because our ordinary judgments depend on the available evidence. Given our cognitive finiteness we are justified to utter some judgments, responsive to the human accessible evidence, that irremediably differ from real desert. This latter claim is indeed independent of the EA core; it is an additional commitment (EA is an essentially negative

³ Zimmerman (1987), p. 381. See also Zimmerman (2002).

argument, whereas the Conservative Claim is a particular positive development among others), but it is typical of MS positions. Finally, MS defenders commit themselves to counterfactual moderate claims about moral judgment like this: we judge people for the way that it is *plausible for us to think* they would have acted, provided that they had had the chance.⁴ Then, the locus of judgment is the agent's actual character, dispositions or intentions.

In sum, this position aims to vindicate the Control Principle, while trying to avoid a radical revisionism of our ordinary judgments. (Significant variations occur within MS: some versions stress the distinction between legal or pragmatic judgments and moral judgment, and other versions insist on the difference between acts of blame and moral blameworthiness, etc.) However, this strategy is only available for resultant luck and certain kind of circumstantial luck. MS, due to its very moderate essence, cannot take account for antecedent kinds of moral luck, because in such cases neutralizing luck in moral judgment would mean going beyond the agents' actual characters or dispositions. What defines MS as such is its conservative or moderate commitments, especially the commitment to the idea of the *plausible*. By its very nature, MS cannot be extended to antecedent kinds of moral luck, failing then to become a real Global Case against moral luck. Its extension would involve renouncing to the moderate commitments, as defenders of the Radical Strategy do. Anyway, MS can still work as a particular argument against resultant and (some sorts of) circumstantial luck, isolated or as part of either *mixed strategies* (that combine different arguments for different kinds of moral luck) or *hybrid strategies* (that accept some kinds of moral luck and reject others).

⁴ See, esp., Richards (1986), pp. 173-4.

Then, I move on to dealing with the Radical Strategy, my real target in this chapter. Beyond MS, it is possible to pursue “the implications of the denial of the relevance of luck to moral responsibility” to their “logical conclusion”.⁵ RS starts from the idea that the Control Principle is the only guide to determine moral desert. RS defenders have no problem to claiming that the agent’s *actual* intentions, will, character or dispositions cannot be the basis for attributing responsibility, because they depend, at least partially, on factors beyond the agent’s control. Then, RS defenders appeal to the idea of basing (true) desert in those dispositions that the agent could have had given her possible counterfactual histories, i.e. dispositions that she would have had if her factual history had been different, among the whole of her life’s possible histories. Of course, RS gets rid of every commitment (to actual intentions, the “plausible” restriction, etc., characteristic of RS), beyond the control principle itself. RS rejects also something like the Conservative Claim, and becomes a radically revisionist position: our current moral judgments are incorrect, given that they do not satisfy the Control Principle, and should be revised. It is, indeed, an *error theory*: it is not the case that our ordinary moral judgments do not try to refer to the ultimate desert of the judged agent (or only try to do it imperfectly), but they actually try to, but systematically fail. In sum, RS is a more ambitious and comprehensive enterprise than MS, but, as I will try to show, that enterprise is achieved by losing its intuitive character and possibilities of application.

We have seen that, as far as they are external to the agent and not free of contingencies, consequences and overt actions were not good candidates for being the locus of ultimate responsibility/desert. Character,

⁵ Zimmerman (2002), p. 556. See also Zimmerman (1987) and Greco (1995).

intentions or choices could not be better replacements, since even what character or dispositions one has or what choices one makes is partly a matter of luck. Then, since luck appears again on stage the issue is not solved, but simply postponed—and, remember, true desert cannot be a matter of luck in any way, but only a function of one's control.

A move deniers recommend here (as a partial answer) is to make a distinction between a 'factual true desert' and an 'essential true desert'.⁶ The factual true desert would be a function of what one would have freely chosen and done in a diversity of situations, *given the person's actual history*, while the essential moral desert would be a function of what the person would have freely chosen and done in a diversity of situations, *including a diversity of possible histories*. The factual true desert would depend on those dispositions one has, given one's factual history; and the essential true desert would depend on a broader set of dispositions, which *includes the agent's counterfactual possible histories*. It is the latter that would keep luck away in the way required to avoid moral luck. But by making this move, the proponent of an ultimate true desert faces new problems.

I posit here two sorts of objections that affect the applicability of this proposal.

Firstly, by splitting up an agent's actual moral record from her true desert, a big gap emerges between the actual assessments we do and people's true deserts. This is, no doubt, an undesirable consequence. However, the denier might acknowledge that it is really difficult to make a judgment about true desert or essential moral worth, but insist that this does not imply radical scepticism about true desert. Limited judgments about true desert can be reasonable, even though we must be very cau-

⁶ For this strategy, see Greco (1995), p. 94.

tious about making them.⁷ Moreover, in so far as circumstances in which a person indeed chooses and act are a subset of the overall range of circumstances in which that person would have chosen or acted, “a person’s moral record provides a window on that person’s moral worth”.⁸ Even a putatively positive consequence might be drawn: this sort of reasonable scepticism about true desert would undermine our righteousness when blaming others who were less lucky in the situations they had to face.⁹ But this is a move more easily accessible to the moral luck defender, without the necessity of positing such entities as essential true deserts: a consequence of acknowledging moral luck is that our judgements of moral responsibility are more *fragile* than we could previously think. And it is not very hard to see that, once we dissociate true desert from our moral record, the link between them is definitively cut, and to stop at that moderate scepticism or to talk in terms of such a magic window is just the result of a decision, or a mere act of faith.

Secondly, this proposal would imply an implausible radical revisionism, which would actually be impossible to carry out. The RS theorist is that all those people that, in certain circumstances, would have freely acted in the same way as, e.g., the Nazi collaborator acted, are as blameworthy (culpable) as he is. And the same goes for moral praiseworthiness or laudability. The outcome is that we are all blameworthy (and praiseworthy) for countless things, due to countless counterfactuals putatively true about each of us, that we have never imagined. Consequently, Zimmerman prescribes a general and radical revision of our practices, including moral judgments, punishment, etc.¹⁰ But, as it is obvious, this revision

⁷ This view is defended by Richards (1986), Rosebury (1995), and Greco (1995).

⁸ Greco (1995), p. 93. See also Richards (1986).

⁹ See Greco (1995), p. 93-4.

¹⁰ See Zimmerman (1987) and (2002).

is indeed impossible to be put into effect. How can we revise (change) our actual judgments on the base of a multiplicity of essentially counterfactual judgments? (I will come back to the issue of revisionism in Chapter 9.) On the other hand, this idea involves a universalization of positive and negative attributions of moral responsibility that threatens to dissolve the very idea of moral responsibility. Such a universalization, if possible, would eliminate distinctions in personal desert; and this, no doubt, would have morally pernicious consequences.

However, those worries about applicability are not my main objection against the denier. My main contention is that the notion of (essential) true desert turns out to be impossible, not only to know, but also to fix, even in ideal conditions.

By making the move of going backwards and relying on the dispositions the agent would have had given her *counterfactual possible histories*, the proponent of an essential true desert takes progressive steps backwards that ultimately reduce the agent's identity to nothing, to a bare self with no properties. Pursued to its logical conclusions, the denier's view, which rests on the idea that what ultimately matters is *only what exclusively depends on the agent*, becomes meaningless, since it happens that finally nothing *exclusively* depends on the agent. No character or identity remains to be assessed as the agent's deep self (putative locus of her true desert). In other words, there is finally no agent on whom anything might depend. Therefore, the very idea of an *essential true desert*—as actual-enactment independent (indeed radically independent of the actual world)—becomes, at the end of the day, unintelligible.

Certainly, it is quite legitimate to feel that attributions of moral responsibility must be deep, must reflect something “really belonging to the person”. And, then, it is a fair aim to try to separate, to a certain extent,

some more internal and essential traits of an agent from external formative, environmental and circumstantial factors. However, that cannot involve an image of the agent as essentially consisting in a fixed or substantial self that stands behind her various psychological and physical dispositions. But this is exactly the image of the self that the radical strategy against moral luck necessarily presupposes.

So, RS turns to be inapplicable and, more importantly, incoherent. It is worth noticing that if RS were only inapplicable, that would suppose an important obstacle for its defense but would not preclude that it were the ideally correct position. Notwithstanding, by turning out to be also incoherent, RS is finally discredited.

Neither MS nor RS accomplished their intended role of constituting a Global Case against moral luck. Nevertheless, it is still possible to reject the existence of moral luck (in all its kinds) by means of a mixed strategy, consisting of a combination of RS with a particular argument for constitutive moral luck. On the other hand, the very rebuttal of the Global Case does not involve that all kinds of moral luck are, *ipso facto*, vindicated; but only —and this is quite a lot— that to find a single general strategy is not possible.

I finish Chapter 3 by comparing my argument against the Global Case with a known argument about it, due to Michael Moore, that I dub as the *all-or-nothing argument*.¹¹ Although both arguments work as a *reductio* and share important similarities, they are significantly dissimilar. In general, my argument is less ambitious than Moore's, given that it does not preclude that some kinds of moral luck may still be convincingly denied. But this is not a flaw of my argument, so long as Moore's pretension

¹¹ See Moore (1997), Chapter 6.

is not achieved: there could still be particular reasons for not accepting particular kinds of moral luck. I cite here Zagzebski's idea of the cumulative effect of luck and the distinct agential and moral signification of the different kinds of moral luck.¹² On the other hand, my argument primarily attacks the notion of true or unconditioned desert, so long as it is essential for preventing that luck "taints" the moral judgment that someone deserves. My conclusion, unlike Moore's, follows necessarily, given that it does not depend on a conditional. So, I judge it more solid than Moore's, although this argument will have to be complemented by individual responses to the rejection of each particular kind of moral luck—this will be my task for the remaining chapters of Part II.

Chapter 4 is devoted to constitutive luck. As advanced, it is still possible to reject the existence of moral luck (in all its kinds) by means of proposing a specific argument against constitutive moral luck.

Two important arguments have been put forward against constitutive moral luck. The first one denies the coherence of the very notion of constitutive luck. The very constitution of an agent, this argument goes, cannot be luck for anybody, since it is required anybody or something previous on *whom* or *which* luck rests; so, the notion of factors under or beyond one's control cannot be applied to received personality or character. The second, in turn, aims at neutralizing factors of constitutive luck by means of the agent's posterior intervention on her own future constitution or character. The latter has more to do with questions of formative luck, and then will be tackled in the next chapter.

The central idea against the notion of constitutive luck is that, if the existence of someone depends on possessing certain essential proper-

¹² Zagzebski (1996), p. 72.

ties, so that her existence is impossible without them, there cannot be said meaningfully that that person is lucky regarding her possession of those properties. So, as Rescher says, “[o]ne cannot meaningfully be said to be lucky in regard to who one is, but only regarding what happens to one. Identity must precede luck.”¹³

Hurley has pursued this argument further by making a parallel with the problem of harm in the discussions of the impact of environmental policies on the future generations, an issue that Parfit dealt with in terms of his famous Non-Identity Problem. The issue is that if different environmental policies have cumulative effects on everyday life, then different sperm and egg cells will unit and so different people will be conceived. Similarly, it cannot be said that Mozart was lucky in having the innate talent for music he had, because that innate talent for music was part of his original constitution. Saying the opposite is to isolate a person from his essential physical and mental properties, condemning oneself to a *bare-self* view of identity. In sum, Hurley maintains that the notion of constitutive luck violates the requirement according to which one person can only exist given the conditions that permit her existence, and only with the particular set of (*essential*) properties derived from these conditions.¹⁴

In this sort of objection the very notion of luck plays a fundamental role. Rescher (and Hurley follows him at this point) has defined luck as follows. When we characterize some development as lucky for someone, we mean that (i) the affected person receives a benefit or a loss, a good or an evil, and (ii) that the outcome came about “by accident” or “chance”. Something like this definition of the notion of luck is the one

¹³ Rescher (1993), p. 155.

¹⁴ Hurley (2003).

Pritchard puts in modal terms.¹⁵ And working on this account I produce my own characterization:

*Luck**. An event *e* is lucky for a subject *S*, when *e* occurs in the actual world but does not occur in most of the possible worlds nearest to the actual world where the relevant initial conditions for *e* are the same as in the actual world; and *e* has a positive or negative value for *S*.

This view of luck involves (i) that any event that does not come about by chance or accident is not a lucky event (the *chance constraint*), and (ii) that nothing can be a matter of luck (for someone) if there is no possible world in which the individual exists without it (the *metaphysical constraint*). The individual's identity across possible world is fixed by means of preserving the original zygote from which that individual comes from, that includes her essential properties. (i) has the consequence that no event that is not (probabilistically) accidental or chancy for the agent can be a matter of luck (for her). On the other hand, (ii) has the consequence that it could not be said that, for instance, someone was lucky to have the father that had.

After making clear the assumptions on which the Hurley-stile impugnation of the idea of constitutive luck rests, I try to diagnose what is wrong about them. First of all, this account of the notion of luck is far from being uncontroversial. As can be seen at a first glance, it clashes with certain ordinary attributions of luck, certain cases that we ordinarily consider as a matter of luck. In addition, it moves away from the predominant view in the moral luck debate. The vast majority of participants in this debate understand luck merely as *lack of control*. Anyway, my

¹⁵ See Pritchard (2006) and (2005).

contention is that this view of luck unjustifiedly presupposes that (coherent) attributions of luck regarding any trait of person P necessarily involves the fact that P could not really have had the trait in question. I explain this. Although it is right that P's identity, and the physical possibilities of its future variation, are constrained by the satisfaction of the metaphysical constraint, application of the predicate "being (good or bad) luck" does not entail such a constriction. (I stipulate a distinction between what is *conceivable* and what is *imaginable* to make this difference more explicit.) If this is so, we could rightly say that it is bad luck to be a black person in a society with the institution of slavery (of blacks), or to be a woman in a misogynist society, even though those are essential properties of them.

Does this mean that the mechanism of possible worlds, so successful in producing suitable accounts of very different notions, is not valid for the notion of luck? Indeed it has provided useful analysis of this very notion in epistemology. The point is that, given the essentialist constraint that by its very nature this mechanism imposes, it does not get by well with matters of essential properties. In fact, some philosophers have thought that, in the case of luck, this is not a limit for the proposed characterization, but of the idea of luck itself; the possession of some or other essential properties is a matter of *fortune*, not of luck. The idea is that the previous characterization suits the predicate "being lucky", but "being fortunate" exceeds its limits. The initially received inclinations, dispositions or traits of character of someone are her *destiny*, are things that constitute her identity. However, this distinction seems too weak and stipulative to do the work its proponents entrust to it—even though it rests on some wise considerations about the difference of that part of our identity

or character which we begin with and that part of it due to subsequent factors, of a much more contingent nature.

I propose an alternative way of salvaging the modal account of the notion of luck. There is a sense—the one relevant for our topic—according to which we can continue saying that *it was a matter of bad luck for someone to have the fathers that she had*, even accepting the idea that attributions of luck cannot violate the metaphysical constraint. That is, when in our context we say something like this, the question is not that she was unlucky of coming from the zygote she comes from, but of being born and having grown up in a certain family that is the cause of some morally adverse traits he has. Think, for instance, of the (nasty) idea of being Himmler's son. The point is not that his children were unlucky for coming from Mr. and Mrs. Himmler's gametes, but for being born and growing up in an extremely Nazi family, an environment of fanaticism, whose influence is morally pernicious. But this is not necessarily so, because, even keeping fixed the children's zygotic origin, the vital history of Himmler could have been very different. This solution is not helpful regarding one's originary (and genotypic) essential properties, but could be enough to salvaging the modal account regarding constitutive (phenotypic) *moral* luck.

I reach the conclusion that it is undeniable that the notion of constitutive luck—that does not presuppose that agents could have had a different essential constitution—is far from being incoherent. Moreover, one's character, so long as it is definitory of one's moral identity, plays an outstanding role in our moral appraisal of others, even though this does not excludes the relevance of the relationship of an agent with his own character and degree of control on it. I finish this chapter by going deeply into the question of the moldability or possibilities of implementation or

modification of constitutive properties. Here, our enquiry is about formative possibilities and the limits of what is coherently conceivable about the variations of personal traits. These matters advance the topic of the following chapter.

Chapter 5 is devoted to the issue of formative luck. In presenting the previous chapter I said that there were two key arguments against constitutive luck, and only the first was discussed. The second argument rests upon the idea of self-formation of one's character. It is true that we have different sets of traits that we simply receive or inherit, that is, we simply find within us; nevertheless—the argument goes—we do not judge and are not judged for this, but for our *formed character*, in which the agent's intervention is crucial. Consequently, in this chapter I explore the issue of the agent's power in forming and reforming her own moral character and of her life's ups and downs, which leads each person to build herself (in more or less extent). This self-formative power could serve—it is the denier's hope—to achieve the longed-for neutralization of constitutive inequalities. Particularly, I will do this by examining the main philosophical theories about moral responsibility, i.e. those theories that want to provide the necessary and sufficient conditions for an agent being rightly considered an appropriate candidate for moral responsibility attributions. My purpose is not to directly assess those theories or to confront their relative merits. I will neither try to decide whether the proposed conditions are sufficient (or even necessary) for moral responsibility. What I will be concerned with is whether, in the case in which those conditions were correct and fulfilled, they could block the influence of (constitutive/formative) luck in moral judgment. In my view, none of them can do this.

I begin by those theories that focus on the agent's *reflective self-control*. According to these accounts, what primarily distinguishes persons is their capacity for self-evaluation, i.e. the power of reflecting on their own desires and willings, and making them fit (particularly, making fit one's first- and second-order desires or willings). Harry Frankfurt, a key figure in this approach, maintains that the crucial point, regarding moral responsibility, is "whether the dispositions that [the agent] has [...] are features which he identifies and which he thus by his own will incorporates into himself as constitutive of what he is."¹⁶ And that even if, after all, we are inescapably formed by and subject to circumstances beyond our control. "The causes to which we are subject may also change us radically, without thereby bringing it about that we are not morally responsible agents."¹⁷

However, some philosophers have thought that what is fundamental is not a kind of *merely volitional* reflective self-control (of a hierarchy of desires), but a *rational* kind of reflective self-control; that is, a reflective capacity primarily consisting in weighing reasons by which the agent is able to adapt her attitudes to reasons.¹⁸ This position announces a main project against antecedent luck: the project of the *rational overcoming* of moral luck, whose main idea is that possession of a rational capacity neutralizes the different circumstances of everyone's life history regarding her moral tendencies. In the stronger versions of this idea, we are all rational beings, beings that participate in Reason, a fact that lets us to access—if we exercise correctly this capacity—to know what is right in every situation and for every aspect. However, it is more plausible to ap-

¹⁶ Frankfurt (1988), pp. 171-2.

¹⁷ Frankfurt (2002a).

¹⁸ See Watson (1975).

peal to a more moderate *normative capacity*, that would qualify us not only for *self-revision* (as in the reflective model), but also for *self-correction* of our mental states.

As exemplary views of this approach I examine Susan Wolf's well-known theory and a recent proposal by John McDowell. According to Wolf, "an agent cannot have the kind of freedom and control necessary for responsibility unless, when making choices about values and actions, she can understand the significant features of her situation and of the alternatives among which her choice is to be made."¹⁹ To be a morally responsible agent would involve being capable of appreciating the reasons that there are for correcting oneself, doing "the *right* thing for the *right* reasons" or "acting in accord with the True and the Good". Also McDowell has recently insisted that one determines oneself in so far as one thinks or acts the way one does for reasons that, thinking and acting as one does, are responsive to the reasons that there are.²⁰ In both cases, responsiveness to the reasons that there are is referred both to theoretical reasons (reasons for believing), and to practical reasons (reasons for acting).

This view is not blind to the fact that normative competence is socially acquired and social circumstances are very different for different persons. Particular attention deserves those agents with several deficits in their upbringing and education that result in a diminished rational capacity. Wolf excludes *insane* people (remember her JoJo's case) from being morally responsible agents because of their lack of appreciation of the True and the Good. However there remains a big grey area of people with less than an ideal rational capacity, of whom it is too restrictive to say that they are not fully morally responsible agents. The idea was that the acqui-

¹⁹ Wolf (1990), p. 117.

²⁰ McDowell (2007), p. 1.

sition of a normative competence equips the agent with resources to revert previous influence of luck, but to be really effective this competence must be very sophisticated, and it is implausible to demand from people—as a necessary condition for moral responsibility—the possession of a very sophisticated normative competence. My point here can be put as a dilemma: either we exclude too much people from being morally responsible agents, or we cannot preclude an important influence of luck. It seems to me right to conclude that the rational overcoming project turns out to be unsuccessful; influence of luck cannot be prevented this way.

On the other hand, there are a lot of ways of personal development, changes of outlook or even conversions, which are not due to rational considerations, but merely to growing up, falling in love, having shocking experiences, growing old, etc. We cannot claim that these changes turn agents into non-morally responsible people, unless, again, we were disposed to drastically reduce the sphere of morally responsible people. In human formation and development, things such as parental monitoring, influence of family, teachers, friends, peers, etc., are common. Particularly, a good formation requires good parents, good teachers, fortunate social relationships, etc. No doubt, a good formation depends at a great extent on a *good alien control*, i.e. on a beneficial control of people in our setting with the power (in very different degrees) for that. Our moral character and agential identity depend on—are unconceivable without—the web of interpersonal relationships in which we were and are involved.

There is an alternative to what I called the rational overcoming project; that is, the project of *historical overcoming* of moral luck. The denier of antecedent moral luck does not necessarily have to be an opti-

mistic rationalist. A distinct move stresses the importance, in neutralizing constitutive luck, of controlling one's self-formative decisions. This approach focuses on the agent's formative history; in particular, on those initial and especially relevant decisions for her subsequent (current) character's configuration of her own character. Its key idea is that if the agent is ultimately responsible for these decisions, formative of her character, there will be no further question of constitutive moral luck. Her character will be the result of her own (initial) decisions, and not something merely inherited. This is what justifies attributions of moral responsibility.

To characterize this project I appeal to Robert Kane's theory.²¹ (Kane never talks about *moral* luck—as neither do the philosophers previously considered in this chapter—but his theory, so long as it is meant to neutralize luck in one's decision-making, especially in one's self-forming decisions, could be a good tool to eliminate constitutive luck.) Kane describes self-forming decisions as indeterministic processes in which conflicting efforts of one's will bring about a decision, regarding which the agent has *plural voluntary control* over the set of options: she could bring about both options voluntarily, intentionally (knowingly) and rationally (having reason for that). In the case of self-formative decisions it is finally by acting in one way or other that the agent makes up her mind to do A or to do B. And it is by means of decisions and actions of this kind that the agent acquires a newly formed character, which determines a large number of her future actions, for which she will be fully morally responsible so long as they spring from her character and her character was formed by herself. This sort of project emphasizes the idea of being the creator of one's own character, rather than the idea of rightly responding to reasons—to the reasons that there are, or to the True and

²¹ See, esp., Kane (1996).

the Good—, although both emphases are not mutually exclusive, of course.

But this project is also problematic regarding luck-neutralization. First of all, any theory of this kind has to face this dilemma: i) the more basic and initial a self-forming decision is, the less grounded it will be, the less control the agent will have over it; ii) contrarily, in order for decisions to be sufficiently robustly grounded, they will have to rely on a (partially) previously formed character, but then they cannot be the most basic and initial. This dilemma poses a strong obstacle to the aim of showing that agents can be unconditioned authors of their own character, not being their self-forming decisions a function of their previous character, but neither being arbitrary decisions.

On the other hand, a lot of philosophers have objected to Kane that a lack of determination between the agent's mental states and her decisions and actions is of no help in guaranteeing the agent's control over her decisions and actions. I do not pursue this classical objection further. Rather, I stress the idea that decisions or actions especially relevant for character formation constitute thresholds that compromise the agent's future responsibility, even in spite of the temporal distance and the high degree of indetermination between those decisions and actions, and their future consequences and repercussions. Limitations in their being brought forward are reflection of the essential *lack of transparency* of the relationship among self-forming decisions, the character that one is developing and repercussions in future actions and identity. Imagine, for instance, that a person achieves to develop moderation as a trait of character of her. Some years later she is involved in a situation in which the right moral response is an immoderate, radical or drastic, response; but she is not capable, given her character, of that kind of response. In cases like this, the

agent will feel what can be called *character regret*. The result is a non-eliminable presence of luck regarding the link between self-forming decisions and future actions. The relationship among self-forming decisions, character formation and future actions becomes transparent only retrospectively.

On the other hand, there is a different way of dealing with this kind of issue present in the literature on moral responsibility, that primarily focuses in what we called the *epistemic condition* for moral responsibility and in the tracing of its antecedents in the agent's life history. The idea here is that for an agent to be morally responsible for some belief, that belief has to be culpably acquired. We can trace back the presence of that belief in the agent's web of beliefs to an act of *culpable ignorance*, from which the present belief comes from. Some theorists maintain that an agent can be morally responsible only if she believes that she is acting on a morally wrong belief, or she has intentionally developed a bad disposition. Most of those theorists situate the locus of moral responsibility in an act of the will, considered by them the agent's control centre. Against this narrow approach, I argue for the idea that an agent is blameworthy if she simply developed or reinforced in her the habit due to which acquisition of certain attitudes does not appear to be incorrect to her, not being it necessary that this development or reinforcement was intentional or deliberate. My position aims to be anti-volitionist regarding formation of habits or dispositions, and externalist regarding the epistemic condition for attribution of moral responsibility.

In the last section of this chapter, I discuss those who deny that agents might be directly blamed for their characters. The idea is that it is unfair to blame someone for a trait that one simply has and regarding which the agent is passive. The only fair locus of moral judgment is ac-

tion or acts of the will. Of course, I try to reject this position, by arguing for a particular version of what has recently been called *attributivism* (the idea that an agent is morally responsible not just for actions or acts of the will, but also for those of their traits that can be rightly attributed to her), that respects the distinction between the badness of a trait and its blameworthiness. The difference lies, according to me, in whether the trait pertains or not to an agent with rational capacities, who is able to assess and revise it, although she never did it. I finally identify as important sources of this problem the foundationalist presupposition that there is just one basic locus for attributing moral responsibility, as well as the excessive emphasis on the work of the will. Nevertheless, the issue is far from being settled; the following chapter will resume the dialectics character/action by considering a crucial ingredient in this dialectics: the situation.

I conclude that neither the rational overcoming project nor the historical overcoming one achieve the aim of neutralizing constitutive and, in general, antecedent moral luck. A more general conclusion is this: the aim of fully eliminating the uncontrollable factors involved in personal constitution and circumstances of formation and development from the sphere of moral responsibility is just a vain wish, doomed to failure. It is not risky to say that this is simply a fact about how we acquire our condition of morally responsible agents in the real world. Luck is not only an inescapable factor in our constitution and formation, it is a necessary factor. Indeed, it will be by studying the ways luck is present in our constitution and formation, that we will be able to achieve a better and more realistic understanding of our moral character and our moral judgments about it.

To summarize what we have advanced for now in Part II, it can be said that, at this point, if my arguments are correct, the aim of achieving a complete rebuttal of moral luck, in all its kinds—by means of the Global Case alone, or in combination with denying the coherence of the notion of constitutive luck, or even by neutralizing it through either the agent’s rational capacity or her self-creation—, is no more attainable. However, there remains, as an alternative to surrendering oneself to this result, to concede that moral luck cannot be globally rejected, in particular that constitutive and formative luck are inescapable, but to hold nonetheless that circumstantial and resultant luck should be rejected; that is, to defend a *hybrid strategy*. Indeed, this is not an odd position, and its rationale is compelling: “Okay, we are who we are,” one could say, “we all have had diverse fortune regarding our constitution and our life histories have surely been very dissimilar in opportunities, but it is one thing to acknowledge this is and it is another thing to allow that particular circumstances or consequences of our actions make a moral difference.” In fact, most of the participants in the debate have focused on these two kinds of moral luck, while putting the antecedent kinds aside or even completely ignoring them. Yet, I will argue for both circumstantial (next chapter) and resultant luck (Chapter 7).

Chapter 6 is devoted to circumstantial luck. As in the previous chapters, I begin by presenting the attack on circumstantial moral luck and, then, I put forward my response. In order to do so, I will retrieve cases, arguments and distinctions considered in Chapter 3, though now specifically restricted to circumstantial luck. However I will resort no more to antecedent factors in the world and the agent’s life history, but to the differences in subsequent facts that the denier excludes from moral evaluation.

Let us consider this case. Actual Judge and Counterfactual Judge are two judges in a Court. Both are corrupt: both would accept a bribe, if one were offered to them. But the bribe is only offered to Actual Judge, who takes it. Counterfactual Judge, due to her different work schedule, is not the judge of the defendant who offers the bribe and, then, no bribe is offered to her and, eventually, she cannot take it. Nevertheless, for the denier both of them deserve the same moral judgment. However, as Nagel said: “[w]e judge people for what they actually do or fail to do, not just for what they would have done if circumstances had been different.”²²

Before discussing this conclusion, let us look at this other case. George and Georg both intend to kill someone. In order to do that, both of them get a gun, study their respective victim’s route, etc. Eventually, both are in front of their respective victims, but only George shoots and kill her, whereas Georg, due to the presence of some people in the immediacies that could identify him, has to abort his plan. The outcome is that, although both had the same intentions, only George is guilty of a murder; not Georg. It can be said that George is responsible for something, namely to bring about a death, and Georg is not. Hence, it seems that the moral judgment either deserves is different. However, the denier of moral luck argues that both are equally blameworthy, so long as their intentions were the same.

One may understand those intentions, alleged to be the only locus for moral responsibility, as either the product of the agent’s decisions or as a part of her character. In any case, it seems that a stable mental quality is looked for as locus of moral judgment. Hence, I focus on the claim that what we should judge is the agent’s quality of character independently of her or his actual actions, and explore some difficulties for this view. (The

²² Nagel (1979), p. 34.

option of focusing specifically in particular decisions is pursued at length below in connection with resultant luck in next chapter.)

Preliminarily, it is worth pointing out how misleading it is the transparent picture of agency as consisting of some agency-related mental traits, from which, to be sure, actions spring, but which are independent of these actions themselves; the view, that is, that character is fixed independently of overt action. No doubt, a person may be greedy, envious or cowardly, but behave perfectly well thanks to an effort of the will. And we can blame someone for being greedy, envious or cowardly, even if he refrains from behaving out of those character traits. As Nagel says: “Even if one controls the impulses, one still has the vice. An envious person hates the greater success of others. He can be morally condemned as envious even if he congratulates them cordially and does nothing to denigrate or spoil their success.”²³ However, this does not mean that the trait is completely independent of its corresponding action. One’s character is partially defined by one’s behavior. Moreover, it seems intrinsically worse to have a trait and act on it than just having it. There exists a dialectical relationship among character, actions and situations, which is far from being transparent.

Anyway, the claim that what we should judge is the agent’s character presupposes a particular view of character: a specially stable, or robust, view of it. Why? Because the rationale of this claim is the aim of judging a stable mental quality of the agent, something in the agent not subject to variation due to non-controllable circumstances. However, it is not evident that this view of character is empirically appropriate. In order to clarify this point I proceed to examine and evaluate the discussion on the nature of character traits.

²³ Nagel (1979), p. 32-3.

As traditionally understood, character traits seem to be relatively broad dispositions to respond with the trait-relevant behavior to the trait-relevant situation. However, research in social psychology has produced some results that appear to some to undermine our ordinary talk of character traits. In particular, there is a set of experiments, grouped under the heading of *situationism*, regarded as giving support to something like this broad claim: systematic observation of behavior contradicts people's behavioral reliability that we would expect in virtue of our ordinary beliefs and standard theoretical constructions of personality. The empirical findings include, *inter alia*, the discovery that subjects who had just found a dime were 22 times more likely to help someone than those who did not find a dime, that passersby not in a hurry were 6 times more likely to help an unfortunate persons in apparent significant distress than passersby in a hurry, that subjects were less likely to help when ambient noise was at extraordinary high levels, that college students role-playing as "guards" in a simulated prison subjected "prisoner" students to intense verbal and emotional abuse, or that subjects repeatedly shocked a screaming victim at the polite request of an experimenter, etc. In the main text, I describe what can be considered the more relevant experiments in this tradition.

These results have been generalized this way: *situational factors, even those that in principle are apparently irrelevant (not rationally salient), play an outstanding role in determining behavior, a role that is higher than has been traditionally acknowledged.* And the so-called Fundamental Attribution Error—a thesis about people's appraisal of other people's behavior—has been proposed:

(FAE) People have an inflated belief in the importance of personal factors, together with the failure to recognize the importance of situational factors in affecting behavior.

In particular, those results have currently been seen as specifically threatening to character trait attributions. The Situationist Claim about Character Traits can be put as follows:

(SC) Traditional character (or personality) traits do play a less relevant role in prediction and explication of behavior than particular situational factors.

In my opinion, both claims rightly follow from the experimental results, but they remain indeterminate as stated. The point is this: Okay, experimental results show that our behavior is less dependent on our personal factors/traits than we traditionally thought; or the other way round, situational factors are more important in producing behavior than we thought. But, to what extent? How less dependent on personal traits, and how more on situational factors, is our behavior? Some philosophers, especially Gilbert Harman and John Doris—call them eliminativists (about character traits)—have maintained that these empirical findings cast serious doubt upon the very existence of character traits, everyday attributions notwithstanding.²⁴ I oppose this conclusion.

My reply begins wondering why it is supposed that these results exclusively undermine character traits, and not other personal factors. For instance, in Milgram's experiment most of the subjects in the face of the experimenter's demands acted a way contrary to that which they previously thought they would. Therefore, situation prevails over previous personal factors. But personal factors include more things than just the puta-

²⁴ See Harman (1999) and Doris (2002).

tive character traits. Other dispositions that a person is usually said to have are dispositional beliefs and desires, intentions, emotional dispositions and attitudes, and also enduring preferences, goals, plans, values, etc. So, it seems that what would follow from the experimental results is the non-functionality, in producing behavior, of all that set of personal factors. But, in that case, the consequence seems to be that behavior completely depends on the situation, which involves a sort of behavioral nihilism difficult to accept. In fact, the very dichotomical distinction between personal and situational factors does not appear to be very promising. Anyway, my argument works as follows: if one accepts that the experimental results as widely generalize—this being the only way to be justified to deny the existence of character traits—then one would have also to deny the role in behavior of the rest of personal factors as well—which means accepting behavioral nihilism—given that no principled way of applying the experimental results just to character traits and not to other personal factors has been offered.

Actually, eliminativists do not deny either that people possess other kinds of dispositions (which are robust character traits), or even that there are behavioral regularities ruled by those other dispositions a person possesses. What experimental results refute, according to them, is just the existence of *global* character traits—i.e. cross-situationally consistent and temporally stable dispositions of character—but not what they call *local* character traits, whose scope is limited to similar situations and do not presuppose consistency and general stability. Their behavioral reliability would be *highly specific*, highly contingent on context—a contextualized predictability insufficient for grounding attribution of broader traits. Nevertheless, the distinction between global and local traits is more problematical than it seems. My general suspicion is that it might also be an *ad*

hoc distinction, motivated by the necessity of avoiding total confrontation to our everyday intuitions, within the eliminativist project. Anyway, the point I want to stress here is that the distinction appears to be *too weak*, and the identification of local traits is not less problematic than the identification of global traits. Individuating situations and, particularly, identifying “similar situations” is not a simple question, especially taking into account that we aim to identify similar situations, not regarding objective stimuli, but regarding the subject’s construal of them, which is sensitive to the slightest variation. Then, once we take the risk of going beyond particular, individual occurrences, we will necessarily face problems similar to the ones faced by the fixation of traditional character traits. Doris acknowledges that “perhaps [behavior] should be explained by reference to situational continuity alone rather than by local traits.”²⁵ However, he is undoubtedly committed to the idea that “some behavioral tendencies are reliable enough to warrant the postulation of enduring dispositions”, and that those dispositions are fixed by the “determinative features of situations”—especially because he, like Harman, wishes to avoid a *completely counterintuitive* revisionism of our intuitions and practices. In sum, the problem of fixing *determinative features of situations* is an important objection to eliminativism as long as they regard the notion of local traits—defined as “specified quite narrowly”, that “should satisfy our conditional standard”²⁶—as an appropriate way of responding to some of our everyday intuitions. I finish this section by proposing some prospects for an appropriate account of the nature of character traits.

²⁵ Doris (2002), p. 145.

²⁶ Doris (2002), chap. 4.

Although I reject eliminativism about character traits, I accept the importance of the situationist findings for our conception of character.²⁷ This allows me to defend, without giving up the idea that character traits can rightly be the locus for attributions of moral responsibility, that the idea of limiting moral responsibility attributions exclusively to character, on the assumption that is a stable property of the agent, is unjustified given the actual instability of character. The ruin of a too self-sufficient view of the agent and her character is an undeniable consequence of these findings. Moreover, they have some particular repercussions for personal autonomy, moral responsibility and, so, for moral luck that need to be stressed. The experiments make clear that we often—more often than we usually think—ignore the true motives of our actions, which supposes an important deficiency in self-knowledge—think of the experiments of Isen and Levin on moods—, or that we do not manage to turn our motives or intentions into actions, what involves a deficiency in self-control—think of Milgram’s experiments. No doubt, both results have significant repercussion on personal autonomy.

Let us turn back to the discussion about circumstantial moral luck. An alternative way of rejecting circumstantial moral luck, which avoids a commitment with the existence of robust character traits (and it is then compatible with eliminativism), is to anchor responsibility attributions upon decisions and particular intentions or motives. Given my argument for rejecting eliminativism, it can be anticipated that this move does not look very promising; but let us consider it. Think of a case of marital infi-

²⁷ My rejection of eliminativism is intended as a contribution to the debate on the (non-)existence of character traits as well. That justifies the large number of pages dedicated to it in this chapter.

delity. If it turns out that there are no (robust) character traits it is useless to try to reinforce one's character in the way of developing a disposition for resisting critical sexual attraction. In this case, the best way to refrain from extramarital sexual relations would be to avoid the occasion. However, apart from it being an advice belonging to proverbial wisdom, it seems counterintuitive to locate the blame in not having avoided the occasion and not in having been unfaithful. The very fact of going to the dinner is not blamable in itself. It is only so if it ends in infidelity. We begin to see that the effort to avoid circumstantial luck yields some absurdities. (I explore the idea that most of the time we do not just judge intentions, but also actions in connection with resultant luck.)

Let us retrieve the classical case of circumstantial luck. Adolf and Rudolf support Nazism and are euphoric with Hitler's immediate rise to power. However, Adolf has to emigrate to Argentina due to business. Rudolf, staying in Germany, has the opportunity to join SS and becomes a high-rank officer. Adolf continues his support to Hitler, but being out of Germany, he cannot get involved in so active a support as Rudolf. The denier insists that both are equally condemnable. But, do they really deserve the same moral judgment? And, in virtue of what do they deserve being treated equally? It cannot be due either to their actions, or to their decisions or particular intentions. Both agents cannot share exactly the same intentions, given their different histories and environments. They can share some general intentions (or determinable intentions, or plans) but not the specification of these intentions (determinate intentions, plan's fulfillment)—different circumstances shape different intentions. My contention is that this difference in their determinate intentions justifies a difference in moral judgment.

The denier could insist on his view and argue that if the example is powerful this is due to its ambiguity. If the details were refined, we would see that either both counterparts are so similar that there will be no relevant moral difference between them, or they are so dissimilar that they will not be considered counterparts in the sense relevant to the case any more.²⁸ But this seems to involve dissolving the issue, more than solving it, in that it bypasses the intended significance of the cases of circumstantial luck: how different circumstances give room for different opportunities of action, as well as raise different demands on the agent. In particular, certain situations are tests, some of them highly demanding, to which only some of us are subjected.²⁹

It appears that regarding a lot of traits, if not all, we should wait for their behavioral actualization in order to be in a good position to judge their possessor. Even regarding predictions of our own behavior, we often err. Sometimes we fail to act in the way we desired and thought we were going to act. We never are completely transparent, neither to others nor to ourselves. Moreover, our behavior frequently makes us to *reinterpret* our previous intentions or mental states, on the basis of our performing of one or another action when facing certain situation—a reinterpretation that is not fully a discovery, but a constitutive part of the intention or trait. It is not that before certain action one was interpreting wrongly her will and after the performance she got the correct interpretation; it is rather that a lot of traits are simply undetermined until certain behavior fixes them. One's character depends, partially but unavoidably, on one's actions.

²⁸ See Pritchard (2006) for this account.

²⁹ In the main text, I connect these remarks with Hanna Arendt's claim of *the banality of evil*—indeed, this was the starting point of Milgram's research, which he himself took as its demonstration.

Finally, there is a somewhat different way in which circumstances can influence moral responsibility: through confrontation of moral dilemmas. The specificity of this kind of circumstantial moral luck depends on the position subscribed regarding the possibility of real conflicts of values. The difference between genuine moral dilemmas and the previous cases of circumstantial luck is that in the former the agent cannot help acting wrongly; whereas in the latter the agent faces a moral test, that may be really hard to pass but not impossible. The defense of the existence of genuine moral dilemmas is associated to a tragic view of life, i.e. a view according to which people, sometimes, cannot avoid moral guilt. In the next chapter I resume the issue of dilemmatic choices due to conflicts of values.

In sum, cases of circumstantial moral luck, which both disturbing and fascinating, show us important aspects of our moral experience— aspects the denier disregards.

Chapter 7, which finishes Part II, is devoted to resultant (or consequential) luck: the kind of luck involved in the way things turn out from agents' actions and their influence in moral judgment. If my preceding arguments are correct, this is the only kind of moral luck that can still be resisted. Indeed, this is likely the kind of moral luck that most philosophers reject. Again, I will present the main arguments against resultant moral luck and try to undermine them. In my reply I will draw on Williams' argument in favor of moral luck; particularly, on his claim about the retrospective character of justification and his notion of agent-regret. I will argue for the relevance of actually causing harm in assigning moral responsibility.

Remember the murderers' case, now specified as an example of resultant luck. George and Georg have gotten a gun, studied his victim route, etc. Eventually, both shoot on their respective victims, resulting in George killing his victim, but not Georg. Georg failed the shot. Although both had the same intentions, and acted in the same way, only George killed someone and is guilty of a murder. Georg did not commit a murder, just attempted to commit one. It can be said that George is responsible for something—bringing about a death—that Georg is not. However, in Michael Zimmerman's terms, both are equally responsible in *degree*, even though the *scope* of their responsibility is different. George is responsible of more things than Georg, but both are equally responsible or responsible in the same degree. Luck does not make a moral difference, given that the only relevant factor for moral culpability is the *degree* and not the *scope* of moral responsibility.

On this account, what counts is just the agent's will or decision and its consequent action. However, in defense of moral luck, we can vindicate the role played by the different reaction that each result triggers in us, and the fact that such distinct results will have divergent repercussions in the victim and the agent's character and identity. The complete articulation of this response will take up all Chapter 7.

In addition, there are other ways in which resultant luck could make a moral difference. Think about the drunk drivers' case. It is evident that they neither *intended*, nor *decided*, to kill, in any reasonable meaning of these words. Instead, the denier claims that in cases of this sort attribution of moral responsibility depends exclusively on negligent or reckless action, actual outcomes aside. (Let us define negligence as *an unjustified assumption of risk*, an assumption that is unjustified when the action's potential harm is bigger than its potential benefits, given the probabili-

ties.) For the denier, both drunk drivers are equally negligent; both subjected other people to a morally unjustified risk.

But, in order for this account to work, the (moral) risk taken should be defined with accuracy and from the agent's perspective at the time of action, or *ex ante*.³⁰ Against the denier's view it may be adduced an opposite intuition pointing out that an important part of our assignation of desert appears to spring from the very fact of actually causing harm or even death. Pressing on this intuition, I argue that, at least in some cases, *it seems that the very negligence exists only if the outcome is harmful*. The rationale for this claim is this: ordinarily, we take risks, moral risks included, but it would be a too rigorous view of morality and life in general, and indeed impracticable, to make moral blame independent of the actual outcome. The standard of demand cannot ordinarily be so high. We cannot judge both drivers with the same harshness; and neither can we treat them with the same leniency. (I do not intend to deny that, obviously, there can be cases of plain risk, determinable independently of results.)

There exists, finally, a third sort of consequential luck, closely related to the last one: when decisions must be made under uncertainty. Think, for instance, about leaders of a clandestine organization, who must decide when to attack the cruel government of a dictatorial régime that subjugates the country. In this scenario, decisions cannot be based on minimally accurate knowledge of what consequences each action will bring about—since no knowledge of this sort is available. Many uncontrollable factors may eventually influence the outcome. Unavoidably, the rebel force's leaders face up a very hard and risky decision. If insurrection is successful, the country will break free; but if it fails, the régime will

³⁰ See Rosebury (1995), pp. 518–519; and Enoch & Marmor (2007).

reply with a bloody repression that will produce a large cost in lives and suffering.

Both in cases of decision under uncertainty and of negligent behavior, the denier circumscribes attribution of moral responsibility *exclusively* to the very decision of acting, from an *ex ante* perspective—which just takes into consideration the foreseeability or probability of consequences in virtue of the relevant information available to the agent at the time of deciding. On the contrary, I will argue that the retrospective perspective has also a part to play in assigning moral responsibility. In order to advance towards this conclusion, let us look at a case famously proposed by Bernard Williams.

The case is broadly based on Paul Gauguin's life—however, being a thought-experiment, some deviances of his real biography must not be taken into heart. Williams' Gauguin has to decide between these two options: either (i) staying with his family, that he feels responsible for; or (ii) moving to a Pacific island, where he thinks he will be able to develop his gift for painting and become a great painter. Both options are mutually exclusive—carrying out one involve the impossibility of carrying out the other—, and both are regarded as valuable by him. Eventually, Gauguin chooses option (ii), i.e. he emigrates to Tahiti pursuing his passion for art, and abandoning his wife and children. A right understanding of this case is not easy, given its polyphonic character. Part of my aim in this section is to shed more light on it.

So, given this scenario, the question is whether we can say that his decision was or was not justified; that is, was it rational for him, given his interests, to act in the way he did? This is the question Williams wants to answer. In order to clarify the question, I propose to tell the very choice

between (i) and (ii) from the justification for opting for (ii). Let us begin with the conflict between rational and moral justification, and after that we will consider the issue of the retrospective character of justification.

The first thing to notice is that Gauguin's case is not a *moral* dilemma. The conflicting choice involves a moral value (his obligations regarding his family) and a non-moral value (regarding self-perfection or individual flourishing: his artistic vocation, that constitutes his life's project). It is rather a *practical* conflict or dilemma: both are practical values. Now, if it is recognized that the tension, the dilemmatic nature of the choice, remains, even if it is a conflict between two moral values, then we are in the right way to the conclusion that moral value or obligation does not always overcome other sorts of value. However, something more has to be shown.

No doubt, most of you will allege that the right option for Gauguin would be to stay with his family; apart from some initial distortions, nobody should doubt that moral value ought to prevail. Nevertheless, Williams' counterargument here is to say that it is absurd or unreasonable that morality obliges people to renounce to the hard core of their projects, by means of which they precisely give meaning to moral demands. In Gauguin's case, responding to the moral demand—submitting oneself to it—would involve a too high personal sacrifice. It would suppose to give up his life's project, which is what gives meaning to his existence, namely his artistic vocation and cultivation of painting in an environment prone to developing his potentialities. This vocation may constitute a practical necessity for Gauguin, resulting even of a *discovery about himself*.³¹ Morality clashes sometimes with practical necessity, practical *must*—a broader sphere including our life's plans, vocations, categorical desires, and so on.

³¹ See Williams (1981), p. 130.

If, given the excessive sacrifice that would have been for Gauguin staying in Europe with his family and renouncing to develop his artistic gifts, we acknowledge that he was (rationally) justified to opt for pursuing his vocation, then we are accepting that rational justification can overcome moral justification.

Gauguin's case, as presented by Williams, may seem a little bit artificial as long as it appears that all depends on a certain and punctual decision. In time t of his life, Gauguin stakes all: he must decide between leaving and staying, and this is a definitive choice. This contrasts with most cases—as well as with Gauguin's real life—in which discovery is often gradual or merely negative. I propose Rembrandt's story as a more plausible example of gradual decision or conviction.

Yet, I endorse Williams' conclusion; regarding conflict of values, the case is convincing: moral value is not always the supreme value. However, accepting the clash between rational and moral justification and the non-supremacy of moral justifications is independent of endorsing the retrospective nature of justification. Someone could accept that rational justification may on occasions overcome moral justification, but deny any influence of results in being justified—even if the situation will become tragic if the project fails.

It is evident that Gauguin cannot be completely sure regarding his options for becoming a great or good painter, until he carries out his attempt and finally do or do not achieve it. Although, of course, he may (and should) have some information pointing to the plausibility or reasonability of such a prediction: he has already painted promising pictures, teachers and experts have praised his work and encourage him to go on painting, etc. But even granting the existence of reasonable signs pointing to a promising future, what is the ground for saying that Gauguin's choice

was justified? Williams claims that, given this situation, the only thing that will justify or unjustify Gauguin's choice will be the results of his project. Only if he is successful in carrying out his project, he will be able to think that he was justified in acting as he did. Given that it is impossible to foresee the project's success or failure, it has to be conceded that justification will be "essentially retrospective".³² Things being so, and as long as the project's actual success or failure is something beyond the agent's control, rational justification will necessarily be open to the influence of luck. (Although not any kind of luck; there is a distinction between intrinsic and extrinsic luck that is relevant here. Only intrinsic luck is linked to unjustification.)

I draw on Williams' argument, but I will follow my particular route to rebut the denier. I will argue for the claim that the retrospective perspective is an ingredient of moral assessment, as well. But, particularly, I claim that *what actual harmful consequences do is to raise the standard of moral assessment*. This is especially clear in cases of negligent behavior and decisions under uncertainty, but it is also true, although more controversial, of cases of bad intention.

Doubtless, it is *prima facie* reasonable to claim that justification excludes the retrospective perspective: someone is or is not justified in her decision regarding those aspects *in time t in which she decides*. If the agent meets the conditions for good deliberation, i.e. if she acts according to a right reasoning from the available evidence in time of reflection, it will be inappropriate to blame him, no matter how many further pieces of evidence would lead him to a different decision. Either the agent deliber-

³² Williams (1981), p. 24.

ated rightly, being thus justified; or she erred, and so was unjustified.³³ However, this view disregards *the force that how things turn out has in our lives, and particularly in our decisions*. When we face hard choices we are in a precarious position; what we are troubled for, the very aim of decision, is success.³⁴ The attempt of distinguishing between two different aspects of justification—namely, justification at the time of one’s taking a choice (or *ex ante* justification) and retrospective justification or circumscribed-to-consequences justification (*ex post* justification)—clashes with the fundamental importance that, we, as agents made up by our actions and its consequences, give to changes we bring about in the world. Hence, proposing such a *technical* distinction is gratuitous, as long as it is inert. What the claim of the retrospective nature of justification stresses is that one is linked with one’s actions in a way that exceeds one’s relation to one’s own actions as an *ex ante* rational deliberator.

Regarding other aspects of Gauguin’s case, in the main text I reject William’s claim that the feeling of gratitude we feel towards him for having created great paintings that have made better our world, which obviously depends on his project’s success, involves endorsing Gauguin’s choice. I also propose an alternative presentation of Gauguin’s case that, it seems to me, stresses more rightly the point of moral conflict, i.e. the future well-being of his family.

There is a complementary way, already advanced, of arguing for the role the retrospective perspective plays in moral assessment, which is based on the more intense feeling of condemnation we feel against agents who actually brought about harm than to their counterparts. Even granting

³³ For this view, see Rosebury (1995), Latus (2001) and Pritchard (2006).

³⁴ Cf. Moore (1997), p. 232-3.

this fact, the denier can dismiss it as part of a reflective moral judgment: differences in reactive attitudes should not make a difference in moral desert. They can be misleading, as long as we can also feel similar revulsions toward inanimate objects, like the ripper's knife, even though it is plainly absurd to think of such objects as morally responsible.³⁵ Moreover, the very concept of moral responsibility "loses its coherence if this residual unhappiness [the moral feeling] is listed among 'moral' sentiments".³⁶

To argue against this stance, I begin by looking into the nature of some feelings ordinarily involved in moral assessment. Particularly interesting is agent-regret, a kind of regret a person feels regarding her own actions, whose "constitutive thought" is that it would have been much better had she done otherwise.³⁷ The relevant conscience of having done some harm is that that harm is the consequence of one's own actions. This feeling is especially exemplary in showing how first- and third-personal perspectives give access to fundamentally diverse facts. On the one hand, the agent regrets the harm done by her, even if she was not at fault. On the other hand, spectators will ordinarily try to decrease the agent's regret, consoling her, insisting she was not at fault. But the agent herself cannot simply adopt the third-personal perspective and tell herself that she has nothing to regret (about herself), given that she was not at fault. It is one's causing harm what brought about the feeling of regret.

Agent-regret appears to be a moral sentiment, at least in the minimal sense that its associated thought is a negative moral assessment by the agent of her own action, as long as she sees her action as contrary to

³⁵ Richards (1986), p. 179.

³⁶ Rosebury (1995), p. 517. See also Zimmerman (2002), p. 562.

³⁷ Williams (1981), p. 27.

morality. However, the denier can respond to this by endorsing the moral nature of this feeling, while rejecting at the same time its connection to ascription of moral responsibility. Indeed, Susan Wolf argues for the existence of a *nameless virtue* which consists in “taking responsibility for one’s actions and their consequences”.³⁸ This is a virtue of generosity that goes beyond one’s responsibility for the harm done. Our sentiment would not be a reliable guide to moral desert, but to other aspect of our moral experience. Indeed, William’s case can be used to illustrate this. However, the role played by agent-regret in William’s argument, according to the interpretation I offer, is mainly to show the significance of actual harm for moral appraisal. And this significance is easy to understand if we consider morality as particularly concerned with guaranteeing respect for persons.

I finish by considering the moral sentiments of guilt—a feeling about one’s own culpability—and the third-personal sentiments of indignation and resentment. These sentiments contrast with agent-regret, provided that they are necessarily connected with voluntariness or moral fault. All of them can be good guides to moral desert and are especially connected to harm done. I vindicate the role of our attitudes and emotional reactions in finding out what we value and care about.

In addition, it is worth pointing to a tension in harm-directed moral feelings, related to the previously noticed asymmetries between the first-person and the third-person perspectives. These feelings, linking the agent’s action with a harm done, may focus more strongly in one or other element. Indeed, “A regrets having brought about H” is ambiguous between “A regrets H” and “A regrets being the person who caused H”. This ambiguity is due to the existence of two different focuses for regret:

³⁸ Wolf (2001), p. 13.

the harm done (H) or one's own agency (A). It is both one's awareness of the harm done and of one's being the person who caused it that leads us to repair our wrong deeds, in the victim and in our own identity. The ideal is a balanced consideration of both elements.

In sum, the existence of the sort of reactive attitudes discussed highlights both our concern for the harm done and the agent's contribution to it. These affective reactions are a kind of imprecise judgment that, when tracing facts, uncover relevant aspects of the situation's *moral truth*. We will lose something important if we exclude them from our understanding of our moral appraisal of other people and ourselves. Putting them aside would involve fatally distorting moral reality. (In the next chapter I will put emphasis on the causal history linking the agent with the harm done, a link that is partially normative.)

Part III

The last part of this dissertation focuses on the main repercussions of acknowledging the existence of moral luck. I will argue for the non-troubling character of this acknowledgment—if it appears to be so, it is just because we have a misleading view of either agency or morality, or both. In particular, in Chapter 8 I will deal with the significance and consequences of revising the Control Principle. In chapter 9 some related meta-theoretical issues, including variantism and revisionism about moral responsibility, will be addressed. Along this part I will also try to differentiate my position from other moral luck advocates.

Chapter 8 opens by retrieving Nagel's approach to moral luck. According to him, the very practice of moral judgment is *internally paradoxical*. We are implicitly committed to disregard things beyond the agent's control as

L

a basis for desert; but consequences, circumstances and constitutive elements of character, which after reflection we plainly see that are beyond one's control, are treated as relevant factors for determining moral responsibility. Indeed, the consistent application of the Principle of Control would involve undermining the vast majority of moral judgments that we ordinarily make, as we saw. For Nagel, this outcome—the paradoxical nature of judgments of moral responsibility—is a particular token of a more general skeptical problem. Paradoxes are essential, Nagel says, to human condition.³⁹

However, in Chapter 2, I defended the distinction between moral luck *proper* and the paradox about control; the latter having to do with our ordinary practices of moral judgment—particularly, with how things beyond the agent's control can make a difference in her moral desert—, and the former with radical skepticism, or what may be called *the skeptical worry about control*. This distinction has played, at least, a methodological role throughout this dissertation. In addition, the paradox, as put, might be rejected by a simple move; namely, by means of a distinction between *restricted* and *unrestricted control*. Actually, this is Michael Zimmerman's strategy.⁴⁰ Reformulated in terms of this distinction, Nagel's paradox becomes unconvincing. We cannot identify "something being under our control" with "something being *completely* under our control". In other words, it is a mistake to think that if x 's occurrence depends to some extent on something beyond my control, I then lack any control over x . Or again, to conclude, from the fact that my control upon anything is always vulnerable, that I do not have any control at all. Besides, that strong notion of control is not our ordinary notion, the one that

³⁹ Besides Nagel (1976), see Nagel (1986).

⁴⁰ Zimmerman (1987), pp. 217-21.

is really involved in everyday moral appraisal and, so, not the relevant one regarding the phenomenon of moral luck.

I intend to claim neither that moral luck and the skeptical worry are not related at all, nor that this move solves the problem of skepticism about control. (I refrain myself from deepening in this last topic in the present essay.) However, the preceding move does mean a rational justification for separating the skeptical worry from moral luck. Moreover, rejecting the necessity of understanding control in this strong sense allows me to dismiss the view of those who try to rebut the control principle on the basis of identifying control with such a strong notion.⁴¹

On the other hand, it is worth noting that in our debate no philosopher offers a positive argument for the Control Principle. The dialectics of the discussion is rather set up this way: some putative cases of moral luck are alleged, then deniers present a putatively unmasking interpretation of those cases or accept them as a right characterization of our practices but propose a revision, confirming thereby the Control Principle—whose defenders simply take as evident. However, it seems that some argument for it should be provided.

Maybe, the so-called Fairness Argument could supply a good positive argument in favor of the Control Principle. Recall, in particular, the Corollary to the Control Principle: two or more people ought not to be morally assessed differently if the only differences between them are due to factors beyond their control. And they ought not to be morally assessed differently because *it is unfair to blame or hold someone morally responsible for some wrong beyond her control*. Often, this is put in terms of a

⁴¹ This is Browne's strategy; see Browne (1992), pp. 347-8.

reproach: judging someone for something beyond her control is to be morally unjust or even callous.⁴²

Although the argument is not really strong—it might even be circular—, most philosophers have thought that it embeds so basic a principle that trying to find justificatory reasons or more basic premises might prove to be impossible.⁴³ Indeed, the Control Principle is allegedly the reflex of one of our “more commoving ideals: the ideal that human existence must be ultimately fair.”⁴⁴ However, the fact of relying on such a strong ideal is not sufficient for immunizing it against criticism. Actually, my results in Part I suggest that this high ideal turns out to be unequivocally illusory.

In addition, in this chapter I want to supplement my previous results with a series of experimental findings that point to the same upshot, as well as to question of the Fairness Principle as a moral aim worth achieving.

Concerning the former, I report some experiments that bring to light diverse asymmetries in folk attributions of intentional action and moral responsibility. *Prima facie*, it seems that people’s folk-psychological attribution of intentional action is independent of moral judgment: we *first* make an attribution of intentional action and *then* use this attribution as on input for forming a moral judgment. In parallel, we intuitively think that the folk mechanism of moral responsibility attribution consists of two distinct steps: we *first* settle the facts and the agent’s responsibility and *then* we determine the degree and polarity (blameworthiness or praiseworthiness) of that responsibility. However, a corpus of

⁴² See Domski (2004), pp. 463-4.

⁴³ See Sher (2005), pp. 180-1 for discussion on this.

⁴⁴ Williams (1985), p. 195.

experimental data—which makes up a sort of cognitive science of folk intuitions about attributions of intentional action and moral responsibility—appear to show, on the one hand, that people’s attributions of intentional action are often influenced by moral considerations (this is the celebrated *Knobe effect*) and, on the other hand, that the two steps of the mechanism for attributing moral responsibility are deeply entangled. In particular, I point out that some of those findings show that folk attributions of moral responsibility are influenced by considerations concerning the *actual* behavior’s moral polarity and the *actual* consequences of actions.

The aforementioned experiments include Joshua Knobe’s studies on asymmetries in attributions of praise and blame and intentional action for good and bad side-effects of one’s actions; as well as Malle and Bennett’s studies on asymmetries in assignation of praise and blame for pairs of good and bad actions and intentions (intentions received less praise and blame than their corresponding actions and bad intentions were blamed twice as much as good intentions were praised); also Pizarro, Uhlmann and Salovey’s findings, according to which people is more prone to praise than to blame agents when acting out of strong emotions; and eventually S. Walster’s experiment on the differences in assignation of moral responsibility in cases of severe and mild harm (it was shown that it is more likely that people consider morally responsible those agents who cause a severe harm than those who cause a mild harm, even if in both conditions the agent’s action is described as equally reckless. These results also agree with particular claims I argued for in Part II of this Dissertation.

Subsequently, I move on to discussing the Fairness Argument. In this section I seek to add some normative considerations to the primarily analytic or descriptive arguments in favor of moral luck offered in Part II.

As long as the Control Principle is vindicated as a moral aim by means of the Fairness Argument, it might happen that it clashed with other moral aims and, so, turn out that its sacrifice leads to an overall profit. Against the Control Principle, I argue for the moral superiority of Dependability and the virtue of *taking responsibilities*. Following M. U. Walker, I conceptualize the incompatibility between both principles in terms of the two different notions of agency assumed. *Pure agency* restricts agency and agents' responsibility to what strictly falls under their control; so pure agents will be freer and more independent regarding the diverse weight with which responsibility may burden an agent's shoulder. On the contrary, the notion of *impure agency* focuses on the moral significance of our causal link with the world and the importance of our response to those changes that we cause in it. No doubt, the impure agent will take more burdens on her shoulders, inexorably subject to luck, but it is just this acknowledgment of one's own vulnerability to luck that makes the agent dependably virtuous. Walker's conclusion is that a world without moral luck (of pure agents) would be a worse world to live in.⁴⁵

However, an objection to Walker's view tells us that she is unjustifiedly conflating two distinct senses of responsibility: the backward-looking sense and the forward-looking sense. The philosophical debate on moral responsibility focuses on desert-oriented responsibility, i.e. the sense of moral responsibility that implies *deserving* blame or praise, independently of any consequentialist benefit. Nonetheless, both my results in Part II and the preceding empirical data show that our practices of responsibility attribution work in a different way. Indeed, against that regimentation we can add the fact that our practices of praising and blaming, no doubt, depend on their effects in us. I mean neither that the aim of indi-

⁴⁵ Walker (1991), *passim*.

vidual practices and attitudes is always to influence particular agents in a certain way, nor that this should be the purpose of agents' judgments and reactions at all. Certainly, most of these attitudes and judgments are not individually so oriented. Instead, what I am pointing out is the overall aim of the responsibility system in influencing people for complying with moral norms and developing virtuous dispositions.⁴⁶

Walker's argument has a second part or a variant, which we may call the *Argument from Moral Integrity*. It runs as follows: if we accept the fact that moral integrity is a value worth having, so that a world with persons of integrity is better than one without them; and integrity is a quality of character "impossible to capture fully without reference to the vicissitudes of moral luck"⁴⁷ (i.e. the existence of moral luck is a necessary condition for integrity), then, a world with moral luck is much more desirable and morally worthy than one without it. Instead, a world when agents routinely and with no justification distance themselves from their actions' harmful or even cruel outcomes would be an unbearable place to live.

An objection to this argument is not difficult to see: what is more fundamental is that agents ought to form right intentions, attempt to behave correctly and not to be negligent; and this is what the criteria for judgment of moral responsibility should promote. So, Walker's insistence on agents' concern about actual harmful outcomes resulting from their actions misses the point; because it has the consequence of making agents to consider themselves free of guilt—in sum, the consequence of exempting lucky agents.⁴⁸ However, this criticism is unjustified, as long as moral

⁴⁶ See Vargas (manuscript) for recent work in this direction.

⁴⁷ Walker (1991), p. 241-2.

⁴⁸ For this objection, see Domski (2004), and also Domski (2005). Statman (2004) responds to Domski; my reply is in the same vein.

luck makes a difference as regards the increase in culpability for actual bad consequences, or by adding a new culpability to the initial culpability for the bad intention or negligence that carried those consequences. It does not follow either from acceptance of moral luck, or from the argument from moral integrity, that we should consider people free of culpability when negligence does not derive in harm; nor that people should not make the effort to deliberate and avoid reckless behavior. On the other hand, the demand that agents (moved by the wish to give restitution to the victim) should care about their actual victim in the very same degree or intensity than about their potential victim is striking, provided that the needs of each victim are different—in addition to being a highly revisionist and psychologically implausible proposal.

After rejecting the Control Principle, two subsequent problems arise: the lack of a limiting principle for attribution of moral responsibility and the remaining feeling of injustice before unfair moral judgments.

Firstly, it seems that if we accept that a person can be morally responsible for things beyond her control, the immediate problem is that we need an alternative limiting principle for moral responsibility, unless we were disposed to admit that a person can be considered morally responsible for anything. So, in order to alleviate this problem, I propose the following principle of Jurisdiction:

(J) A person can only be held morally responsible for things which she is appropriately connected with; and this *appropriate connection* is defined by normative considerations (moral norms) particular to the different practices of moral judgment.

This limiting principle defines the area of a person's moral responsibility; all those things that fall within this sphere make up the agent's *jurisdiction*, i.e. all that the agent is accountable for. It is worth saying that J is not a mere paraphrase of the control principle, provided that it is compatible with moral luck. With it, I intend to pick up the intuitive character of the idea that a person can only be morally responsible for things she is appropriately connected with, but without having to reject moral luck. No doubt, the principle is to be further specified, and some may think it is empty as well as disagree with its leaving the fixation of specific norms for determining moral responsibility to particular practices and contexts. However, I think that J can play the role of an abstract limiting principle, at the same time as be in accord with acknowledging the inefficacy of a substantive principle really able to fix in abstract the conditions for attributions of moral responsibility—that agrees better with how our ordinary practices actually work.

I finish this section by characterizing *the fragility of moral judgment* on the basis of the fuzzy limits of jurisdictions and the difficulty of reaching *definitive* (absolute and finally accurate) moral judgments. I point out that judging rightly demands moral wisdom and considerable effort, as well as that moral judgment does not necessarily aim at reaching a final *verdict* that unifies the various perspectives and aspects involved.

The second problem referred to was the remaining feeling of injustice before acceptance of moral luck. Particularly, I wish to discuss a moral-luck-friendly proposal consisting in the separation of moral judgment and punishment. The idea is that, in face of a morally wrong action, we should not react with anger and blame *against the agent*, but with a

sort of blame that does not include hostility and the will to punish.⁴⁹ More radically, some philosophers have argued for a plain distinction between blame or censure and the (retributive) reactive attitudes—such as indignation or resentment—, which usually come together, and have recommended an elimination of the latter.⁵⁰

In my reply, I argue that this proposal would imply giving up a basic means by which we react to wrongdoers, a means by which we ordinarily try to *do justice*—even though in an irremediably imperfect way. Indeed, it would suppose giving up the very possibility of doing justice at a moral level. Moreover, to say that criticism and reproaches do not embed *moral* blame involves again a strict regimentation of practices of moral judgment, far from our actual practices. Eventually, reactive attitudes (constitutive or not of moral judgment) play a fundamental role in our interpersonal relationships and in our own self-understanding, so that the idea of dropping ordinary praise and blame cannot be the right answer. To be sure, human action demands a wide-ranging variety of reactions that include a broad spectrum of emotional shades.

Likewise, the proposal of dissolving indignation towards the wrongdoer on the basis of the influence of luck would imply her dehumanization, as well as an arrogant paternalism. Acceptance of moral luck—in origin, in character-formation, in the circumstances we face and in the consequences of our actions—leads us to the acknowledgement of the contingency of our moral self and desert; which, eventually, instill irony into our reactive attitudes and judgments; an irony that makes forgiveness easier, but does not rebut those attitudes and judgments.

⁴⁹ Browne (1992), p. 350ff.

⁵⁰ See Slote (1994), Strawson (1986), pp. 117-120; and also Watson (1987), pp. 285ff.

In the last section of Chapter 8, I explore the possible cultural and social variation of the concept of moral responsibility and the practice of holding responsible. I track different cultural sources (both at the historical and geographical level). The issue is this: Is the concept of moral responsibility culturally variable? Or is it cross-culturally stable? To what extent? I attempt to yield the consequences for our debate that would follow from an answer to these questions. My conclusion is here particularly tentative and conditional. I suggest that the possible existence of substantive cultural differences in the concept of moral responsibility or in the practices of holding morally responsible would have repercussions for acceptance or rejection of moral luck. In particular, cultural variation would mainly be a worry for those who share a pure agency-based view of moral responsibility—characterized by the rejection of the idea of dependability.

Finally, Chapter 9 addresses some methodological and meta-theoretical questions concerning the moral luck issue. It begins with a characterization of the general discussion about folk intuitions and their place in theorizing. After laying out a picture of intuitions and a reconstruction of the dialectics between conceptualists (*armchair philosophers*) and experimentalists, some methodological desiderata about the role of intuitions in theorizing are offered. My proposal is intended to be uncontroversial for most philosophers: intuitions are the inevitable starting point for philosophizing, but experimental results may contribute to tell reliable from unreliable intuitions and detect conflicts of intuitions. Consequently, to the extent that folk intuitions regarding X are relevant to determine appropriate theories about X, it is important to find out what those intuitions are. So, I describe a group of experimental studies that have recently tested

LX

folk intuitions about the *Compatibility Question*. Afterwards, I apply the results to our issue.

The first experiments on folk intuitions about the Compatibility Question yielded the result that we are not *pre-theoretical* incompatibilists as most philosophers assume. For example, an experiment found that college students who were identified as determinists were no less punitive than indeterminists and no less likely to offer retributivist justifications for punishments. And another study showed a clear effect of identification in assignation of moral responsibility regardless of being in a determinist universe. However, Shaun Nichols and Joshua Knobe suspected that those results could be due to the effect of affective reactions, and conducted a series of experiments to explore whether participants would be more likely to report incompatibilist intuitions if the emotional and motivational factors were minimized. These studies yielded a surprising result: most people gave compatibilist responses to concrete cases, but incompatibilist responses to abstract cases.⁵¹ It seems that, eventually, there is no unique answer to the question whether we are pre-theoretically compatibilists or incompatibilist: we have different intuitions depending on how the issue is approached.

I think this is an interesting finding that can be transferred to the moral luck question. Here, the default position (in abstract) is that attributions of moral responsibility presuppose control, or are proportional to degree of control. But it is not difficult to think of concrete presentation of cases that—like in the case of the Compatibility Question—will yield the opposite results. In the main text I design the two sorts of vignettes (for abstract and concrete conditions) for testing experimentally those intuitions, although I have not conducted the experiments. What I intend

⁵¹ Nichols & Knobe (2007).

to show is the plausibility of projecting the experimental results about the Compatibility Question—variation according to the (abstract or concrete) presentation—to the moral luck issue. If such a parallelism is correct, it will provide (indirect) experimental support to the idea that moral luck is fundamentally a conflict between abstract and concrete intuitions.

However, Nichols and Knobe hypothesized (second phase of their experiment) that the compatibilist responses to concrete cases could be due to a prejudice or bias—an affective bias—and tried to test it. Specifically, their explanation of the effect is this: when people is faced with egregious violations of moral norms (concrete cases), they experience a strong affective reaction which makes them unable to apply their tacit theory for responsibility judgments correctly. But these last experimental results are objectionable. True, in one case the affective reaction is higher than in the other, but what lastly makes the difference is not the variance in affect but the *moral seriousness* of each behavior. The fact that the agent commits a rape, in contrast to a fraud, is what actually causes higher condemnation (and a compatibilist judgment)—although, doubtless, through triggering a higher affective reaction. But, ultimately, the crucial factor is not the affective reaction itself, but the judgment of moral seriousness in the behavior judged. Concrete descriptions bring moral seriousness to the fore, whereas abstract descriptions hide it. Indeed, the experiment's design prevents from discriminating between seriousness and affect in bringing about the effect, and so undermines Nichols and Knobe's conclusion.

We must conclude that the conflict between concrete and abstract intuitions cannot be decided in favor of the latter in virtue of the fact that responses in the concrete condition are due to a biased psychological mechanism. So, with no further experimental data available, this conflict

of intuitions cannot be solved (or dissolved) in a merely descriptive way. We are justified to think this is a genuine conflict.

Transferred to our issue, this upshot appears to reinforce Nagel's view, according to which moral luck is an irresolvable conflict—a real paradox. However, there is still an alternative way out: some kind of revision.

Throughout this dissertation moral luck has been understood as a clash of intuitions between (i), particular (or, now, concrete) intuitions about attributions of moral responsibility and (ii) a general (or abstract) intuition, i.e. the control principle. This involves that any possible way out would entail some sort of *revisionism* about (i) or (ii). Indeed, I have been offering different arguments in favor of keeping (i) and rejecting (ii).

Additionally, there can be different sorts and degrees of revisionism. On the one hand, we have to distinguish between revisionism about practices of attribution of moral responsibility (*practical revisionism*) and revisionism about our understanding of how those practices work (*theoretical revisionism*). On the other hand, three degrees of revisionism can be distinguished. Firstly, *radical revisionism* encourages a radical revision of either practices or theory; it constitutes an error theory about what it aims to revise. In the opposite side, *weak revisionism* is a sort of revisionism about *what people think they think*, i.e. people's failure to appreciate the nature of their own practices and conceptual commitments about attributions of moral responsibility. Eventually, in between of both sides we find *moderate revisionism*, i.e. revisionism about *what people think*.⁵²

⁵² See Vargas (2005a) for an interesting elaboration of a typology for revisionism I partially follow.

Consequently, among the views explored and rejected in preceding chapters we considered the Epistemic Argument (defended by Richards, Rosebury, Rescher, etc.), a view that simply advocated a weak revision, a reinterpretation of our ordinary practices of responsibility judgment on the basis of various distinctions that save both our actual practices and the control principle. But we also saw a more radical defense of the control principle. Particularly, Michael Zimmerman argued for an *error theory* about practices—a sort of *radical revisionism about practices*, according to the proposed classification. From the side of the moral luck defender, there was room for a sort of practical revisionism as well. B. Browne suggested dissociating reactive attitudes from punishment, and Michael Slote proposed a radical revision of morality entailing the elimination of moral blame. Nevertheless, a more natural response of those who acknowledge the phenomenon of moral luck is to claim for a kind of *theoretical revisionism*, according to which at least some of our beliefs and theoretical intuitions about moral responsibility are mistaken and should be revised if we want them to reflect actual practices. Bernard Williams, Michael Moore, Margaret Walker, Robert Adams and my own view are within this group.

The classical (Nagelian) characterization of moral luck in terms of a conflict between theory and practice, although right and useful, must be further specified. Indeed, at the end of Chapter 1 a preliminary three-piece distinction was presented. Three are the levels that make up the moral responsibility system susceptible of conflict: i) the level of *practices* of moral responsibility attribution, including judgments, psychological dispositions and attitudes associated (“reactive attitudes”) and further uses of these attributions (rewards, punishments, etc.); ii) the *tacit theory*, which regulates those practices, including folk intuitions, and is not always

transparent to our explicit beliefs; finally, iii) the level of abstract beliefs and principles that we intuitively think rule our practices. The relationship among the three levels is multidirectional. Generally, we think that our attributions of moral responsibility fit several general conditions, as well as that a series of intuitions work as their guide. But these intuitions often conflict with one another and, as result, in different particular cases we diverge regarding whether certain general conditions are sustainable or not. Moreover, there exists a process of overt feedback between the level of the tacit theory and the level of theoretical conceptions.

In the light of this threefold distinction, a new sort of revisionism must be added: *revisionism about the tacit theory*. So, my proposal—which focuses on the relationship between ii), i.e. the tacit theory, and iii), the theoretical conceptions—can now be defined as *moderate revisionism* concening directly to our theoretical conceptions and (provided that these influence the tacit theory) indirectly to our general intuitions. I have intended to show that neither practices of moral judgment, nor folk intuitions and beliefs referred to concrete cases—as long as they are not completely permeable to the direct influence of theoretical conceptions—are in need of revision. What is to be *re-interpreted* are our more abstract and theory-laden beliefs.

I think this proposal is advantageous as long as it entails an easier-to-assimilate sort of revisionism than that advocated by other proposals. In particular, two main reasons in favor of the greater practical plausibility of my proposal can be adduced. Firstly, requirements for moderate revisionism are less demanding than for radical revisionism. A proposal of radical revision must show its necessity and anticipate how things would be after such a revision. Recall that moderate revisionism intends to modify *what people think*. My view entails a modification of people's

general belief that control is a necessary condition for attribution of moral responsibility (or that attribution of moral responsibility is proportional to control). Indeed, it was part of my view that it is the idea of jurisdiction that actually rules attributions of moral responsibility, although people may not realize it and think that their judgments are committed, instead, to the Control Principle. Secondly, my proposal is directed to what people think about practices and not to the practices themselves. Although the recalcitrant character of our intuitions and practices involves a strong obstacle to any kind of revisionism, revising well-established social practices is always more difficult than revising general, theory-laden beliefs. Moreover, the endurance of these practices is a fact in their favor. Proposals of practical innovations, especially if radical, involve an important risk, because (if realizable) they can result in a great disaster. (In Chapter 3 I criticized Zimmerman's radical revisionism as being conceptually incoherent, and in Chapter 8 Browne's and Slote's, on the basis of their psychological implausibility—unattainability?)

To finish, I wonder whether my view could eventually be the result of a process of *reflective equilibrium*. In principle, this method fits well into the spirit of my position, which acknowledges the unavoidability of some sort of revision, but wants to be as conservative as possible. However, inasmuch as acknowledging moral luck entails rejecting the Control Principle, the outcome cannot strictly be a mutual readjustment. The frontal clash of practices and principle obliges us to opt for one or other side. The idea of *jurisdiction*—my replacement for the control principle—cannot be considered a revision or modification of the control principle, provided that it is compatible with acknowledging that luck can make moral differences. Nevertheless, once the control principle has been replaced by the idea of jurisdiction (that is not by definition incompatible

with luck making a moral difference), mutual readjustments between this idea and our ordinary practices of responsibility attributions, as a result of a process of reflective equilibrium, may be necessary.

Parte I

1. Las atribuciones de responsabilidad moral

Esbozo del sistema de la responsabilidad moral

- 1.1. Esquema básico, relación causal y agencia
- 1.2. Las actitudes reactivas y Agentes Moralmente Responsables
- 1.3. Voluntariedad y control
- 1.4. El sistema de la responsabilidad moral y sus tensiones

En este primer capítulo, y como preámbulo a la formulación del problema de la suerte moral, intentaré ofrecer una caracterización de la noción de responsabilidad moral y sus atribuciones, que quiere ser fiel a nuestras intuiciones generales al respecto. El fenómeno de la suerte moral vendrá precisamente a cuestionar esta caracterización de sentido común.

1.1. Esquema básico, relación causal y agencia

Cotidianamente, nos consideramos mutuamente responsables por, al menos, parte de lo que hacemos. Aplicamos el adjetivo ‘responsable’ a otras personas, atribuyéndoles con ello responsabilidad; así como también nos lo aplicamos a nosotros mismos, es decir, nos autoatribuimos responsabilidad. Ser responsable, en principio, puede entenderse como ser capaz de responder por los propios actos.¹ Considérese el siguiente esquema básico para la atribución de responsabilidad:

(AR) *A* es responsable de *O*, en base a *C*.

¹ No debe confundirse el sentido aquí relevante de ser responsable —la posesión de una capacidad— del que está presente en expresiones como “Manuel es responsable”, que constituye una valoración moral del agente, atribuyéndole una virtud, etc.

Donde *A* es el agente que recibe la atribución de responsabilidad; también denominado clásicamente el *locus* o *perpetrador*, que a menudo es un individuo, aunque parece que también puede serlo un grupo o colectivo. *O* es el objeto de responsabilidad, aquello por lo que se atribuye responsabilidad, sea una acción, un rol, un estado de cosas, otros seres, etc.; también se le ha llamado el *contenido* de la responsabilidad. Finalmente, *C* son las condiciones de atribución de responsabilidad a *A*.²

Un elemento que parece ser básico para la responsabilidad es la conexión causal con el fenómeno del cual, o por el cual, se es responsable. Si me caen de las manos las gafas y se rompen al dar con el suelo, soy por lo menos causalmente responsable del hecho de que se rompan. Pero esta responsabilidad causal no es suficiente, porque, por ejemplo, la sierra que corta los dedos de un carpintero es causalmente responsable de este hecho, pero no se puede decir que sea responsable en el sentido relevante. La responsabilidad causal o *superficial*, atribuible a objetos, sucesos, animales, etc., consiste meramente en ocupar un rol causal en la producción de un efecto —sin que se requiera la presencia de una persona o agente racional. Es, pues, necesaria, pero no suficiente, para la responsabilidad personal o normativa.³

² Como de inmediato veremos, estas condiciones son al menos de dos tipos. Las referidas directamente a *A*, y las condiciones ulteriores (aquí llamadas *C*), que están relacionadas con la atribución misma de responsabilidad. Por otro lado, a menudo también hablamos de un elemento *P*, es decir, de una *parte* respecto a la cual el agente es responsable: aquel o aquellos frente a quienes se ha de responder (*accountability*), bien sea la víctima, la familia de la víctima, la comunidad, la persona que delegó la responsabilidad, etc.

³ Es un lugar común que la responsabilidad *causal* es insuficiente para la responsabilidad *moral*. Ver, por ejemplo, Williams (1993b), p. 55. Obviamente, la condición causal puede problematizarse. Por ejemplo, ¿qué hay de las omisiones? Se puede ser moralmente responsable de una omisión (¡no desactivar la bomba!), sin que exista ninguna conexión causal “real”. No obstante, las omisiones son un tipo de acciones particularmente problemático, cuyo análisis detallado escapa a esta caracterización general del *sistema de la responsabilidad*; si bien una concepción razonable de la agencia y la responsabilidad no

La responsabilidad en el sentido relevante a que hacía referencia es la que sólo puede ser atribuida a agentes racionales o personas. Para que alguien pueda ser considerado un agente, en un sentido suficientemente fuerte del término, es necesario, por lo menos, que posea unas determinadas capacidades psicológicas, como es la posesión de las capacidades de comprensión y razonamiento, además de capacidad para formarse deseos, intenciones y planes, de deliberar y actuar por razones y la capacidad de tomar decisiones. Éstas serían las capacidades mínimas constitutivas de disfrutar de la condición de agente. Pero además, para que el agente pueda ser tenido como responsable éste debe disfrutar también de los llamados “poderes de autocontrol reflexivo”⁴, donde cabe distinguir entre el *poder reflexivo* y el *poder de autocontrol*. El primero tiene que ver con aquellas capacidades cognitivas que juegan un papel en la evaluación y selección de acciones, y el segundo comprende las capacidades que motivan y dirigen la propia conducta en virtud de la cognición relevante.

Así, para poder tener por responsable a alguien, es necesario que éste posea estas capacidades mínimas constitutivas de ser un agente y que disfrute de los poderes de autocontrol reflexivo. Además, para las atribuciones particulares se requiere la existencia de una conexión causal (en sentido amplio) con lo que hemos llamado objeto de responsabilidad.

Antes de continuar, hay que precisar que la atribución de responsabilidad no debe identificarse directamente con la *culpabilidad*. Un agente es susceptible de atribuciones de responsabilidad, con independencia de que sea una responsabilidad positiva, orientada así al reconocimiento, o una responsabilidad negativa, esto es, en la dirección de la cul-

puede excluir las omisiones, en tanto que intencionales y fruto de decisiones. En todo caso, la conexión causal ha de entenderse en un sentido amplio.

⁴ Véase Wallace (1994), p. 226; y Fischer y Ravizza (1998), pp. 36–37.

pabilidad. Incluso es plausible afirmar que uno puede ser responsable de algo, sin que esta responsabilidad tenga un signo positivo o negativo; puede ser una atribución moralmente neutra.

Por otro lado, el lugar del objeto de la atribución de responsabilidad, o posibles objetos en relación a los cuales podemos atribuir responsabilidad a un agente —el elemento *O* del esquema anterior—, puede ser ocupado por diversos entes. Típicamente, la responsabilidad se atribuye en relación a acciones, consecuencias o estados de cosas producidos por un agente, pero también en relación a intenciones, rasgos de carácter, reacciones emotivas, etc. Para una mayor concreción de la responsabilidad respecto a los diferentes objetos de atribución sería preciso remitirnos a concepciones particulares teóricamente orientadas de la responsabilidad, cosa que quiero evitar por el momento.

Además, las atribuciones de responsabilidad respecto a agentes pueden ser de diferentes tipos. Principalmente, encontramos la responsabilidad moral y la responsabilidad legal, pero hay también otros tipos, como es la responsabilidad intelectual o epistémica, o la responsabilidad adquirida por la posesión de un cargo o un lugar de trabajo (llamada técnicamente *role responsibility*), etc. En concreto, cada uno de estos tipos de responsabilidad se identifica con conjuntos distintivos de criterios —morales, legales, epistémicos, contractuales, etc.—, que pueden solaparse de maneras diferentes, y que actúan como guía de los respectivos tipos de atribución de responsabilidad. En particular, en el caso de la responsabilidad moral se requiere, en primer lugar, la presencia de un agente moral, caracterizado por poseer las condiciones anteriores, pero especialmente referidas a razones morales, así como de unos criterios de tipo moral que

regulen las posibilidades de atribución y los signos de las atribuciones particulares.

También se ha defendido que un requisito central para la condición de agente susceptible de atribución de responsabilidad moral es la *competencia normativa*;⁵ consistente, a grandes rasgos, en una capacidad compleja que permite que su poseedor atienda a consideraciones morales, se esfuerce por acceder a la información relevante para juicios morales particulares y establezca la conducta que se sigue de estos juicios. La competencia normativa pretende ir más allá de las capacidades mínimas anteriores y suele incluir elementos de contenido o corrección.

1.2. Las actitudes reactivas y Agente Moralmente Responsables

Por otra parte, cuando alguien realiza o deja de realizar una acción moralmente significativa, solemos pensar que se requiere un tipo particular de respuesta. Así, ser moralmente responsable de una acción, por ejemplo, comporta merecer un tipo particular de reacción por haberla llevado a cabo. (Lo que no implica que la reacción abierta esté siempre justificada, aunque muy probablemente sí una tendencia o disposición a actuar de este modo.) Pensamos que las personas que realizan buenas acciones y muestran cualidades admirables merecen elogio, aprobación y recompensa, mientras que aquellas personas que llevan a cabo acciones negativas merecen censura (*blame*), desaprobación y castigo. En otras palabras, la práctica de atribuir responsabilidad a las personas suele ir acompañada por el hecho de tenerlos como merecedores, en algún sentido del término,

⁵ Sobre la *competencia normativa*, véase Wolf (1990), pp. 121, 124 y 129; y Fischer y Ravizza (1998). Cfr. Watson (1987) pp. 263-5, sobre “comprensión moral”. Discutiré directamente estas cuestiones, en conexión con la cuestión de la suerte formativa, en el capítulo 5.

de censura o aprobación, de premio o castigo, de gratitud o resentimiento; respuestas, todas ellas, que forman parte de la compleja red de conceptos que circunda la idea de responsabilidad moral.

Como es sabido, Peter Strawson llamó a las respuestas de esta clase *actitudes reactivas*.⁶ Cotidianamente, concedemos una gran importancia a las actitudes e intenciones que los demás seres humanos tienen hacia nosotros, lo cual está estrechamente relacionado con la gran dependencia que nuestros sentimientos y reacciones personales tienen con respecto a nuestras creencias acerca de aquellas actitudes y creencias. Nuestras relaciones interpersonales constituyen una red de sentimientos y actitudes, tales como la indignación, el resentimiento, el respeto, el agradecimiento, la aprobación, la censura, el amor, el perdón, etc., que nos permiten tomar distintas perspectivas para con las personas. Y como puede verse, estas actitudes no constituyen sólo reacciones *negativas*, como la indignación, el resentimiento, la desaprobación, etc., sino que también incluyen reacciones *positivas*, como el respeto, la admiración, la gratitud, el reconocimiento, etc.

Cabe precisar que puede que no todas estas actitudes reactivas sean igualmente relevantes a la hora de tener a alguien por moralmente responsable. En concreto, R. Jay Wallace considera que para la censura moral son particularmente relevantes el resentimiento, la indignación y la culpa, las cuales tienen en común un mismo contenido proposicional. Cuando experimentamos alguna de esas tres emociones reactivas, respondemos a la creencia de que un agente ha violado una obligación moral.⁷

⁶ Strawson (1974).

⁷ Wallace (1994), p. 11.

Desde esta perspectiva general, podemos definir así al Agente Moralmente Responsable:

(AMR) A es un agente moralmente responsable en tanto que es un candidato apropiado para las *actitudes reactivas*, o por lo menos para algunas de ellas (resentimiento, indignación y culpa), sobre la base de su conducta (o, por lo menos, parte de sus acciones), y, quizá, también sobre la base de sus rasgos (o por lo menos algunos de ellos).⁸

El “ser un candidato apropiado para las actitudes reactivas...” sería así una condición necesaria y suficiente para la atribución de responsabilidad moral. Pero para ser un *candidato apropiado*, además de cumplir las condiciones anteriores referidas a A, o a la constitución del agente, parece que se requiere también satisfacer ulteriores condiciones, que detallaré a continuación. En todo caso, si un agente no cumple las condiciones referidas a A, no cumplirá AMR y estaría *exento* de la adscripción de responsabilidad moral. En particular, un individuo está *exento* cuando su condición (o estado) hace inapropiado imponerle demandas morales; ejemplos típicos de ello son los niños y aquellos que sufren de trastornos mentales severos.⁹

⁸ Esta caracterización se inspira en Fischer y Ravizza (1998), pp. 6-7. También Wallace (1994) ha destacado como defensor de una caracterización semejante. Huelga decir que ambas propuestas, además de diferir de la posición original de Strawson, también divergen entre ellas en aspectos importantes.

⁹ Cabe distinguir la *exención* de la responsabilidad, es decir, el no cumplir AMR, de las *excusas*. Estas últimas se dan cuando aquellos que serían normalmente tenidos como moralmente responsables (esto es, cumplen con AMR) actúan, no obstante, en circunstancias en las que no es razonable o sería injusto aplicarles, en esas circunstancias, exigencias morales; los casos de coerción son aquí ejemplos familiares. La presente distinción tiene su origen en Strawson (1974), pp. 59-81. Ver también Watson (1987), pp. 119-49; y Wallace (1994) p. 118.

Cabe remarcar que esta es una concepción, digamos, de tipo *social* de la responsabilidad moral que creo que encaja bien con el que será nuestro objeto de estudio, la suerte moral.¹⁰

1.3. Voluntariedad y control

Una vez tenemos una caracterización general de los agentes susceptibles de ser considerados como moralmente responsables, hay que concretar ahora la aplicación respecto de las situaciones individuales o particulares; es decir, es necesario establecer las condiciones ulteriores bajo las cuales un agente puede ser, *adecuadamente*, objeto de una atribución de responsabilidad moral particular. Se trata de lo que he llamado elemento *C* en el esquema anterior.

Hay que decir que la responsabilidad moral ha sido definida desde antiguo en términos de lo voluntario, y la voluntariedad, a su vez, ha sido definida en términos de control. Una manera, que puede resultar útil para caracterizar la responsabilidad moral, y que ya utilizó Aristóteles, es la de considerar cuándo no es correcto atribuir responsabilidad a un agente. Así, siendo la posición por defecto que un agente es moralmente responsable siempre que sea un candidato apropiado para las actitudes reactivas, cabrá definir las condiciones *negativas*, restrictivas, de la responsabilidad moral, por las cuales el agente quedaría *excusado*.

Aristóteles mantuvo que la voluntariedad estaba ausente cuando intervenía la *ignorancia* o la *fuerza*. En la *Ética a Nicómaco*, afirma:

¹⁰ Cabe hacer la siguiente precisión. Strawson presentó esta teoría de las actitudes reactivas como un tipo de solución compatibilista *sui generis* al problema de la libertad y la responsabilidad moral. Esto contrasta con el uso que de ella se hace aquí, a saber, como caracterización general plausible de nuestras actitudes psicológicas en las relaciones humanas, lo cual no implica compartir el papel de esta caracterización como solución a aquel problema.

Parece, pues, que son involuntarias las cosas que se hacen por fuerza o por ignorancia; es forzoso aquello cuyo principio viene de fuera y es de tal índole que en él no tiene parte alguna el agente o el paciente, por ejemplo, que a uno lo lleve a alguna parte el viento o bien hombres que lo tienen en su poder.¹¹

Para esta concepción, una acción responsable tiene que ser voluntaria en algún sentido, y para que sea voluntaria un primer requisito es que no haya sido forzada. Sin embargo, Aristóteles considera que hay veces en las que no es claro si un acto es voluntario o involuntario, como en el caso de un tirano que obliga a alguien a llevar a cabo una acción ignominiosa reteniendo a sus padres o hijos, o cuando en una tempestad alguien se ve obligado a tirar por la borda las mercancías que transporta. También respecto a la ignorancia hay multitud de casos difíciles.

Modernizando la terminología, convertiré el requisito de *no-ignorancia* en “condición epistémica” y el de ausencia de *fuerza* en “condición de control”. Con lo que podemos reformular las dos condiciones de la manera siguiente. En primer lugar, la *condición epistémica* afirma lo siguiente:

(CE) Un agente susceptible de responsabilidad moral ha de tener el conocimiento relevante y en el grado apropiado de las circunstancias en las que actúa.

Es decir, no debe ignorar lo relevante de la situación en la que tiene lugar su conducta. Por otro lado, la *condición de control* quedaría así:

¹¹ Aristóteles, EN, libro III, epígrafe 2; 1110a35. Se cita la traducción de María Araujo y Julián Marías; Aristóteles (1960), p. 32.

(CC) Un agente susceptible de responsabilidad moral ha de controlar su conducta en el grado apropiado y en los aspectos relevantes al caso.

También —en la versión negativa— podríamos decir que no debe faltar el control de la propia conducta en el grado apropiado y en los aspectos relevantes al caso. En general, éstas parecen ser dos condiciones básicamente correctas para la atribución de responsabilidad moral.¹²

Hay que aclarar que el debate contemporáneo se ha centrado de manera abrumadora en la discusión del *control* —o, por lo menos, en el término ‘control’. Pero esto no significa que la condición de no ignorancia, o CE, haya sido rechazada. De hecho, CE puede también entenderse como una forma o dimensión del control —el control cognitivo.¹³ Por ello, podemos hablar con igual propiedad de dos condiciones diferentes, con dos nombres, o subsumirlas ambas bajo una misma etiqueta (el control, en sentido amplio). En realidad, distinguir o no entre ambas condiciones responde, a mi juicio, exclusivamente a razones contextuales, sin que haya nada fundamental en ello. En lo que sigue, las incluiré a ambas en mi uso del término ‘control’. Cuando siga el uso contrario, debido a que la distinción se torne relevante, lo haré notar.

¹² Si seguimos con Aristóteles, habría que distinguir en relación a CE entre acciones que cumplen esta condición (‘voluntarias’), de las que no lo hacen (‘involuntarias’), y aquellas en las que el agente no cumple con CE y al mismo tiempo no experimenta ningún pesar (‘no voluntarias’). Respecto a CC, tendríamos también acciones ‘voluntarias’, en los casos en que se satisface la condición; acciones ‘involuntarias’, cuando no se satisface; y acciones ‘mixtas’ cuando uno actúa condicionado y no está claro en qué sentido podría actuar de otro modo. Así, parece que las acciones tendrían un triple valor respecto de cada una de las condiciones formuladas. En todo caso, esta clasificación no está libre de problemas, que no pueden ser tratados aquí. Sobre la cuestión de la voluntariedad, en general, véase EN, 1109b30-1111b5.

¹³ No obstante, hay también una literatura importante e interesante en relación a la cuestión de la condición epistémica: sobre ignorancia culpable, condiciones cognitivas, conocimiento moral, etc. En los capítulos 5 y 7, abordaré directamente algunas de estas cuestiones.

Podemos, ya, formular el siguiente Principio de Control:

(PC) Un agente es moralmente responsable de x sólo si tiene (un grado apropiado de) control sobre x .

La tesis actual más sostenida es que la responsabilidad moral está básicamente conectada con algún tipo de control. El problema reside en precisar en qué consiste este control, o el grado apropiado de este control.

En particular, algunos filósofos han considerado que este control incluye, o debe ir acompañado, del requisito de las posibilidades alternativas: el agente selecciona entre diversas posibilidades alternativas genuinamente abiertas. La tesis es que una persona es moralmente responsable de lo que ha hecho sólo si pudiera haber actuado de otro modo. No obstante, este requisito de posibilidades alternativas recibió la réplica de Harry Frankfurt, quien presentó ejemplos a favor de la compatibilidad de la responsabilidad moral con la ausencia de posibilidades alternativas.¹⁴ Desde entonces, los *casos Frankfurt* han proliferado, aunque a muchos les sigue pareciendo que la posibilidad de alternativas es (pre)condición del control. La cuestión sería si podemos dar sentido al control sin descansar en alternativas. Por ejemplo, Fischer y Ravizza afirman que el *control guía*, esto es, el control dentro de la secuencia realizada, es suficiente para la libertad y la responsabilidad moral; con lo que las posibilidades alternativas no son necesarias.¹⁵

¹⁴ El texto seminal al respecto es Frankfurt (1969).

¹⁵ Fischer y Ravizza confrontan el *control guía* al *control regulativo*. Mientras que el primero sólo requiere que el agente pueda realizar libremente una acción, el segundo comporta “un poder *dual*: por ejemplo, el poder libremente realizar un acto A, y el poder libremente hacer otra cosa” (1998, p. 32). Sólo el primero sería necesario para la responsabilidad moral. Por otro lado, el control guía incluye la posesión por parte del agente de un mecanismo de deliberación y toma de decisiones que responda a razones, el cual ha de pertenecer, en un sentido importante y de una manera apropiada, al agente. Véase

Pero, más allá de perseguir la evolución del debate en torno al control y las posibilidades alternativas, o sobre la libertad y la responsabilidad moral en general, lo que me interesa aquí es presentar el marco en el que cobra sentido el problema de la suerte moral. Como veremos, éste último tiene como uno de sus elementos centrales de discusión el control o, como lo hemos bautizado, el Principio de Control, y, en realidad, supone un desafío importante de éste. Pero creo que no es directamente relevante ninguna concepción concreta o determinada del control —en el sentido anterior. El desafío se extiende a cualquier tipo de control, incluso al control menos exigente, como ciertos de sus sentidos más cotidianos. Por ejemplo, hay una diferencia notable entre tú y yo, cuando ambos vamos en un coche y tú *controlas* el volante mientras que yo voy sentado a tu lado. En esta situación, tú dispones de un control que a mi me falta totalmente. Creo que sólo con esta concepción débil, o relativa, del control ya podemos empezar a hablar del problema de la suerte moral.

Así pues, el sentido de control que habrá que tener presente, en primer lugar, no hace falta que sea fuerte —o metafísico, o último, o como queramos decirlo—, sino que, en muchos casos, es suficiente con un sentido meramente volitivo y cognitivo de control, como es el caso del ejemplo anterior. Esto, no obstante, no excluye que avancemos hacia la consideración, o desestimación, de tipos o sentidos más fuertes —o últimos, o metafísicos— de control. Y, en todo caso (y esta es una lección que aprenderemos de la cuestión de la suerte moral), el no tratar con sentidos más metafísicos de control no hará, en absoluto, la discusión menos relevante o accesoria; en ocasiones será más bien al contrario.

Fischer y Ravizza (1998), esp. cap. 2 y 3, y 6 y 7. Para una exposición detallada e importante crítica de esta posición, ver Moya (2007), cap. 3, § 6. Para una útil presentación de la bibliografía reciente sobre el requisito de posibilidades alternativas para la responsabilidad moral, véase Fischer (1999).

1.4. El sistema de la responsabilidad moral y sus tensiones

Hasta ahora, he tratado de delinear la que puede ser considerada como tarea inicial de una teoría de la responsabilidad moral. En primer lugar, he abordado la cuestión del establecimiento de las condiciones para ser un agente moral o susceptible de adscripciones de responsabilidad moral. Me he referido a la práctica de atribuir responsabilidad a las personas, y a las actitudes que suelen ir unidas a ella. Había que establecer, después, los criterios ulteriores bajo los cuales puede ser aplicado con propiedad el concepto de responsabilidad moral en relación a un caso o circunstancia particular. Hemos visto que, además de las condiciones mínimas que uno tiene que reunir para poder ser considerado un agente moral, la discusión toma cuerpo en relación a si es o no necesaria la voluntariedad o control, y de qué tipo.

Pero más allá de esta tarea, en nuestra caracterización de lo que podemos llamar el sistema de la responsabilidad, cabe también atender directamente a las *prácticas* mismas de atribución de responsabilidad moral y las *creencias* de los sujetos participantes en ellas. En concreto, tenemos que prestar atención a cómo, de hecho, atribuimos responsabilidad moral —los contextos de uso del predicado ‘moralmente responsable’—, y a aquellas ideas o principios que los agentes participantes consideran que regulen estas prácticas. Además, dentro de este último nivel, cabría discriminar entre las intuiciones ordinarias, o creencias de sentido común, sobre la responsabilidad moral o aplicación de ‘moralmente responsable’ —que incluyen principios que nos parecen intuitiva o prerreflexivamente autoevidentes—, y las concepciones teóricas que tenemos sobre la responsabilidad moral, solidarias de aquellas teorías ético-normativas que

sostenemos a nivel más abstracto. Quedaría, así, un esquema como el que sigue:

1. Usos o aplicaciones: atribución de responsabilidad moral. Prácticas (de juicio y reactivas).
2. Contexto regulativo: creencias acerca de la atribución de responsabilidad moral y su justificación.
 - a. Teoría tácita y creencias y principios de sentido común o intuiciones.
 - b. Concepciones teóricas abstractas sobre la responsabilidad moral.

La distinción entre diferentes niveles dentro del sistema de la responsabilidad moral nos ayudará a su comprensión.

En general, pensamos que nuestros usos del término ‘responsable’ satisfacen diversas condiciones o descripciones generales, además de una serie de intuiciones que parecen funcionar como guía. Por ejemplo, parece que las personas se ven generalmente como responsables de aquello que hacen que es resultado de sus rasgos de carácter o tendencias dentro de un espectro normal y con causas normales. Pero sucede que a menudo encontramos desacuerdos destacados, como respecto a si somos o no responsables de aquello que hacemos como resultado, en parte, de serias privaciones. Es fundamental darse cuenta del hecho de que estas intuiciones entran a menudo en conflicto, y que en casos particulares diversos divergimos en cuanto a considerar si ciertas condiciones generales son o no sostenibles. Esto es, nuestras intuiciones más cotidianas pueden divergir.

Por otro lado, la distinción entre los niveles anteriores no es tan fácil, ya que entre nuestras creencias de sentido común sobre la responsabilidad moral y las concepciones teóricas tiene lugar un proceso de retroalimentación claro. Quizá, cabría distinguir aún dos niveles en relación a

las creencias de sentido común. Podemos identificar, por un lado, creencias de sentido común de tipo *general*, como ‘es necesario que uno controle, en alguna medida, su acción para que sea responsable por ella’, que suelen ser más solidarias de las concepciones teóricas; y creencias *particulares*, es decir, creencias o intuiciones que tenemos ante casos particulares, que están más estrechamente conectadas con las prácticas cotidianas de juicio moral, y que a menudo chocan con las generales.

Estas consideraciones son relevantes en tanto que el problema de la suerte moral, que presentaré a continuación, suele ser descrito como un conflicto entre una práctica y una intuición o creencia que, al margen de la práctica particular, consideremos sólida. Más directamente, también puede caracterizarse como el choque entre una creencia general y diferentes creencias particulares.

Dónde estamos y adónde vamos

En este primer capítulo he comenzado caracterizando a grandes rasgos nuestras prácticas de juicio moral, en concreto nuestras atribuciones de responsabilidad moral, considerando cierta bibliografía sobre la voluntariedad y el control. La meta ha sido ofrecer una caracterización lo más neutral posible, que nos ayude a enmarcar nuestro tema. He destacado que no se requiere ni una noción fuerte ni precisa de control para que la cuestión de la suerte moral se plantee. Finalmente, he propuesto ciertas distinciones dentro de nuestra noción cotidiana de responsabilidad moral, las cuales sin embargo no adquirirán un significado preciso para el lector hasta la parte final de esta investigación. Pero pasemos ya a ver en qué consiste exactamente eso de la suerte moral y qué tipo de problemas suscita.

2. La suerte moral

Nagel y Williams sobre la suerte moral

- 2.1. ¿Qué es la suerte moral?
 - 2.1.1. Nagel: control, juicio y paradojas
 - A. El principio y las prácticas.
 - B. La naturaleza paradójica del juicio moral
 - 2.1.2. Williams: justificación retrospectiva y el pesar-del-agente
 - A. La justificación retrospectiva.
 - B. El-pesar-del-agente.
 - C. Contra la primacía de lo moral
- 2.2. Casos y variedades de la suerte moral
 - 2.2.1. Tipos de casos y el corolario de PC
 - 2.2.2. Los cuatro tipos nagelianos discutidos (y modificados)
- 2.3. Contraste de los planteamientos de Nagel y Williams
- 2.4. La articulación del debate: ¿prácticas o principio?

Este segundo capítulo es una presentación detallada de en qué consiste el llamado fenómeno de la suerte moral y cuáles son las particularidades de sus diversas formulaciones.

2.1. ¿Qué es la suerte moral?

Lo primero que debemos saber es en qué consistiría el supuesto fenómeno de la suerte moral. Pues bien, puede decirse que un caso de suerte moral tendrá lugar cuando un agente sea moralmente juzgado — esto es, tratado como objeto de juicio moral—, de modo correcto, con independencia de que un aspecto significativo de aquello por lo que es juzgado dependa de factores que escapen a su control. Las palabras exactas con que Nagel define la suerte moral son las siguientes:

Cuando un aspecto significativo de lo que alguien hace depende de factores que están más allá de su control, y continuamos tratándole a este respecto como objeto de juicio moral, a eso podemos llamarlo suerte moral.¹

Por su parte, Williams afirma que, en general, lo que se discute son ejemplos de determinación por los hechos, es decir, de determinación del juicio de las decisiones de un agente, que éste recibe, por aquello que de hecho ocurre, más allá de su voluntad.² El problema, en definitiva, consistiría en la tensión planteada entre, por un lado, el hecho de que parece que la responsabilidad moral, la justificación, la censura moral, etc., no pueden ni deben ser afectados por la suerte y, por el otro, la posibilidad real de que la suerte juegue un papel importante, o, quizá, incluso fundamental en ellos.

La cuestión central sería: ¿puede la suerte llegar a marcar una distinción moral? En efecto, sólo estaremos ante un caso de suerte moral si la suerte es capaz de *marcar una distinción moral*; la cual puede ser de distintas clases: bien una distinción respecto a aquello por lo que una persona es moralmente responsable, bien que la suerte afecte a su justificación moral, o que afecte en general a su consideración moral.³ Podemos empezar a vislumbrar, pues, que la cuestión no se plantea de un modo unívoco.

La literatura formal sobre la suerte moral comenzó con sendos artículos de Bernard Williams y de Thomas Nagel, titulados ambos “Mo-

¹ Nagel (1979), p. 26; mi traducción. Se entenderá en adelante que todas las citas procedentes de fuentes que no están en castellano son traducción mía.

² Williams (1981), p. 30.

³ Por supuesto, la suerte puede ser buena o mala. Al hablar de suerte, no presupongo la asimilación ordinaria de suerte con buena suerte, como en expresiones del tipo: ‘¡Qué suerte has tenido!’, donde suerte significa buena suerte. La mera presencia de la suerte es todavía incalificada.

ral Luck”, y leídos en el mismo simposio.⁴ Sin embargo, es evidente que la problemática en cuestión ha sido tratada desde antiguo, aunque no en los términos precisos de la discusión contemporánea.⁵ De lo que no hay duda es que inventaron el término *suerte moral*. En este capítulo, expondré los diferentes planteamientos que tanto Nagel como Williams han hecho de la cuestión: cómo la motivan, cómo llegan a sus respectivas formulaciones y, finalmente, cuáles son sus diagnósticos. Ulteriormente, intentaré poner en claro las semejanzas y divergencias de ambos planteamientos, y las diferentes estrategias que se abren para afrontar la cuestión. Asimismo, me ocuparé de una cuestión fundamental para el resto de esta investigación: los tipos posibles de suerte moral y su clasificación.

2.1.1. Nagel: control, juicios y paradojas

A. El principio y las prácticas

Debe reconocerse que, en principio, parece perfectamente razonable acogerse a la intuición de que la suerte no puede variar nuestras valoraciones morales, ni puede alterar en absoluto la consideración moral de una persona, ni influir en su responsabilidad moral. Por mucho que la suerte juegue un papel importante en nuestras vidas, afectando a nuestro éxito y a nuestra felicidad, pensamos que la moralidad es la única esfera donde la suerte no tiene ningún poder. No obstante, parece que nuestros juicios morales ordinarios a menudo tienen como objeto, o toman en con-

⁴ Aparecieron originalmente en *Proceedings of the Aristotelian Society* (supplementary volume), I, 1976. Nagel (1976) y Williams (1976). Versiones revisadas se han incluido después en Williams (1981) y Nagel (1979), reeditadas después en la antología compilada por Daniel Statman (1993); véase la introducción del compilador.

⁵ Véase el impresionante estudio de Martha Nussbaum (1986) de las conexiones entre suerte y ética en la filosofía y literatura griegas antiguas. Por otro lado, cabe subrayar a Joel Feinberg y Peter Winch como claros precedentes del debate contemporáneo. Véase, en particular, Feinberg (1962) y (1970); y Winch (1972), esp. cap. 7.

sideración, de manera significativa, elementos o factores que están más allá del control del agente juzgado.

Por ejemplo, imaginemos el caso de una persona que tiene el propósito de asesinar a alguien, pero cuando se dispone a hacerlo sucede algo, pongamos que falla el tiro, y no logra su objetivo. Seguidamente, se ve obligado a huir sin poder reintentarlo. Contrastémoslo con el caso de otra persona que tiene también el propósito de asesinar a alguien y, como el primero, realiza todas las acciones necesarias conducentes a este fin, pero sin fallar en el disparo. Así, uno acaba matando a alguien, convirtiéndose en un asesino, mientras que el otro no mata a nadie, y no es un asesino. Hay un sentido, en el que parece que nuestro juicio variará, y evaluaremos más severamente al asesino exitoso que al asesino frustrado, y eso como resultado de que la bala llegue o no a impactar en la víctima y le produzca la muerte, algo que no depende de ninguno de los dos agentes.

Parece que, en general, se produce la colisión entre una intuición y una práctica:

Intuición: pensamos que sólo somos moralmente evaluables por aquello que está bajo nuestro control.

Práctica: de hecho, juzgamos a las personas incluso por aquello que no está bajo su control.

No obstante, podríamos hablar también de un choque entre una ‘intuición general’, o *principio*, y una o unas intuiciones particulares, que funcionan en los casos concretos haciendo caso omiso del principio o intuición general.

En este planteamiento, el fenómeno de la suerte moral apunta, primariamente, a la responsabilidad moral. De hecho, la apelación a un

elemento como la suerte, una suerte que se pretende moral, supone un importante reto para la coherencia de nuestro concepto de responsabilidad moral —y, en concreto, de los *juicios* de responsabilidad moral. Como vimos en el capítulo anterior, parece que *prima facie* el siguiente principio, al que llamamos Principio de Control (o PC), es condición necesaria para la atribución de responsabilidad moral:

(PC) Un agente *A* es moralmente responsable de *x* sólo si tiene (un grado apropiado de) control sobre *x*.

Es decir, para que un agente pueda ser considerado moralmente responsable por algo éste tiene que tener control sobre aquello por lo que se le atribuye la responsabilidad moral, y hacerlo en el grado apropiado y en relación a los aspectos relevantes del caso. Lo cual incluye la posesión de un conocimiento suficiente de los hechos, así como la obligación de haber adquirido unas creencias morales mínimas y otras cuestiones de alcance más general (que vimos en 1.1.).

Sin embargo, a pesar de la validez intuitiva de este principio, resulta que en los casos de suerte moral juzgamos la responsabilidad moral de un agente en relación a algo que no está bajo su control (al menos, en el grado apropiado). Puede decirse que el agente *A* es juzgado como moralmente responsable de la acción *x*, aunque *x* o algún aspecto relevante de *x* no estuviera bajo su control. Con ello, llegamos a la definición siguiente de los juicios morales propios de los casos en los que la suerte moral está inmersa:

(JSM) Un agente *A* es moralmente responsable de *x*, aunque *x* depende significativamente de factores que están fuera del control de *A*.

O, para ponerlo de un modo en el que se aprecie mejor su contraste con el Principio de Control:

(JSM*) Un agente *A* es moralmente responsable de *x*, aunque *x* no tenga (un grado apropiado de) control sobre *x*.

Así, y dado que los Juicios de Suerte Moral parecen ser habituales en nuestra práctica cotidiana de juicio moral, el resultado es el choque entre PC y JSM.

Más arduo es decir qué debemos pensar ante este resultado. De hecho, este será el punto principal de discusión en los siguientes capítulos de esta tesis.

Una conclusión más o menos inmediata parece ser que esta colisión constituye una contradicción; ante la cual cabría desechar uno u otra de las proposiciones en conflicto. No obstante, Nagel niega que estemos realmente ante una contradicción; pues considera que si lo estuviésemos —*A* o no *A*—, podríamos, y deberíamos, deshacernos o bien de la intuición (principio) o bien de la práctica. Sin embargo, esto no le parece posible. Según Nagel, la intuición es correcta, y reside en el núcleo mismo de la noción de moralidad; pero, al mismo tiempo, resulta imposible impedir que la suerte influya en los juicios de responsabilidad moral de una persona. Este conflicto entre el principio de control y los juicios morales ordinarios, dado que es real e insalvable, constituiría una paradoja.⁶

⁶ Nagel (1979), p. 26.

B. La naturaleza paradójica del juicio moral

Para Nagel, la misma práctica de juicio moral es internamente paradójica.⁷ Por un lado, es condición del juicio moral que lo que está más allá del control de una persona no puede ser el fundamento de lo que merece; implícitamente nos comprometemos a ver las cosas fuera de nuestro control como más allá del alcance de la responsabilidad. Pero, por otro lado, consecuencias, circunstancias y factores constitutivos del carácter, que la reflexión nos dice que están más allá de nuestro control, son tratados como relevantes para determinar nuestra responsabilidad moral. Y lo que es más, la aplicación consistente de la condición de control nos llevaría a socavar la mayor parte de las evaluaciones morales que consideramos normal hacer.

Es importante resaltar que Nagel no piensa que estemos ante un problema atribuible sólo a situaciones rocambolescas, sino que la tensión, que finalmente lleva a la paradoja, surge de la mera reflexión sobre nuestros juicios morales más corrientes. Es parte de nuestras prácticas el ver la ausencia de coerción, la ignorancia o el movimiento involuntario como condiciones para la responsabilidad moral. Así, partiendo de las evidentes limitaciones externas ordinarias de nuestro control o poder de acción, el movimiento escéptico consiste en descubrirnos nuevas limitaciones, internas y menos evidentes. Se trata de una problematización de la noción de juicio de responsabilidad moral a partir de sus mismas condiciones cotidianas. Y es precisamente esto, el ser una consecuencia natural de la idea cotidiana de evaluación moral, lo que convierte a la suerte moral en un problema filosófico.

Cabe remarcar que Nagel concibe este resultado como una muestra particular, en el ámbito de la moralidad, del problema escéptico gene-

⁷ Nagel (1979), p. 27ss.

ral. Se trata de un movimiento estructuralmente semejante al problema escéptico en el ámbito del conocimiento, en el que a partir de demandas perfectamente naturales, en condiciones normales, sobre las afirmaciones de conocimiento, se llega a una situación de amenaza generalizada, si las demandas son consistentemente aplicadas.⁸ Para Nagel, los argumentos escépticos siguen una misma pauta: no son el resultado de la imposición arbitraria de estándares exigentes de conocimiento, resultante de incomprendimientos filosóficos o teóricos; sino de la aplicación consistente de estándares cotidianos para la evaluación de los juicios.

Así, y volviendo a la esfera moral, Nagel concluye que:

La concepción de que la suerte moral es paradójica no es un *error*, ético o lógico, sino la percepción de una de las maneras en las que las condiciones intuitivamente aceptables del juicio moral amenazan con socavarlo todo.⁹

La tensión implícita en nuestras prácticas de juicio moral se vuelve una paradoja paralizante cuando una “comprensión más completa y precisa de los hechos” se introduce en nuestras deliberaciones reflexivas. Para Nagel, la paradoja resulta ineludible.

Cabe destacar que para Nagel vivir en esta paradoja es parte de la condición humana, forma parte de nuestra autocomprensión. De hecho, en *The View from Nowhere*, Nagel se ha dedicado a diagnosticar diferentes problemas filosóficos que —en cuanto descubrimos la indispensabilidad de la propia perspectiva sobre nosotros mismos, y su incompatibilidad con la perspectiva objetiva— nos conducen ineludiblemente a resultados

⁸ Para más sobre las interesantes analogías entre suerte moral y suerte epistémico y los escepticismos asociados a cada una de ellas, véase Statman (1991).

⁹ Nagel (1979), p. 27.

paradójicos. Y este choque, entre las perspectivas interna y externa, no admite solución, para Nagel.¹⁰

Inciendiando en esta idea, Nagel va más allá e intenta formular de una manera más fuerte y desafiante el problema. Como se ha dicho, es obvio que la clara ausencia de control, a causa del movimiento involuntario, la fuerza física o la ignorancia de las circunstancias, excusa de la acción cometida; pero sucede que lo que hacemos depende de maneras muy diversas de cosas que no están bajo nuestro control y, al mismo tiempo, las influencias externas, en un sentido amplio, no suelen considerarse disculpas de lo que alguien ha hecho, sino que forman parte del mismo juicio moral, sea positivo o negativo. La conclusión es que: “[e]sta forma de determinación moral por lo real es también paradójica, pero podemos empezar a ver lo profundamente incrustada que está la paradoja en el concepto de responsabilidad moral.”¹¹ Más adelante, añade todavía:

Una persona puede ser moralmente responsable sólo de lo que hace; pero lo que hace resulta en gran medida de lo que no hace; por lo tanto, no es moralmente responsable de aquello por lo que es y no es responsable. (Esto no es una contradicción, sino una paradoja.)¹²

Convendrá el lector que este argumento, y en especial su conclusión, es un tanto oscuro. Pienso que podemos intentar ponerlo en claro reformulándolo del modo siguiente:

1. Una persona puede ser moralmente responsable sólo de lo que hace;

¹⁰ Véase Nagel (1986), pp. 125ss. Más sobre esta distinción en la sección 2.3.

¹¹ Nagel (1979), p. 25.

¹² Nagel (1979), p. 34.

2. pero lo que hace depende en gran medida de cosas que él no ha hecho;
3. *por lo tanto*, no es moralmente responsable de aquellas cosas de las que es responsable y de las que no es (directamente) responsable.

Añadiendo el matiz *directamente*, el razonamiento se vuelve algo más claro. Pero aún cabe exigir más. Michael Zimmerman ha reformulado el argumento en estos términos:

1. Una persona *P* es moralmente responsable de la ocurrencia de un suceso *s* sólo si la ocurrencia de *s* no fue cuestión de suerte.
2. Ningún suceso es tal que su ocurrencia no es nunca una cuestión de suerte. Por lo tanto,
3. ningún suceso es tal que *P* sea moralmente responsable de su ocurrencia.¹³

Como las premisas parecen verdaderas y el razonamiento válido, pero la conclusión falsa, tendríamos —de acuerdo con la interpretación de Nagel— una paradoja genuina. El problema reside en que aquello por lo que una persona *P* es moralmente juzgada está parcialmente determinado por la intervención de la suerte, es decir, por factores externos al control de *P*, cuando parece que no debería estarlo.

Sin embargo, creo pueden distinguirse aquí dos cuestiones distintas, que apuntan a problemas parcialmente diferentes, pues una cosa es la putativa paradoja en relación a la práctica ordinaria del juicio moral y su conexión con el principio de control, y otra cosa es esta nueva formulación más estrechamente relacionada con el problema escéptico del control del agente sobre sus acciones. Creo que es fundamental distinguir entre ambas cuestiones, de las cuales sólo la primera se identifica con el pro-

¹³ Véase Zimmerman (1987), p. 217.

blema de la suerte moral; si bien, como es obvio, ambas tienen aspectos comunes y puede que, en última instancia, una derive de la otra. Aun así, creo justificado y útil separarlas. Entiendo que esta distinción no queda muy clara aquí, pero ante ello sólo puedo remitir al capítulo 9, donde desarrollaré esta propuesta.

¿Pero que puede decirse ante el *descubrimiento* de semejante paradoja (o paradojas)? En una de sus concepciones más populares, se entiende que una paradoja es un razonamiento que, desde premisas aparentemente verdaderas, conduce, por medio de pasos argumentativos aparentemente inocuos, a una conclusión contradictoria o absurda. Por ejemplo, en la conocida Paradoja del Mentiroso, un cretense afirma ‘Todos los cretenses mienten (siempre)’ —o, yo mismo: ‘Lo que estoy diciendo ahora es falso’. Sea la premisa de que lo que esta persona dice verdadera o falsa, la conclusión será igualmente tanto verdadera como falsa. El interés filosófico de las paradojas reside, cuanto menos, en que nos fuerzan a cuestionarnos algunas de nuestras creencias más fundamentales, y es aquí donde estaría la importancia inicial del *descubrimiento* de Nagel. Sin embargo, cuando nos encontramos con algo que se presenta como una paradoja a menudo nuestra primera intención es tratar de ver si realmente estamos ante una paradoja o no, y, si no lo estamos, intentar desenmascararla. Y si realmente se trata de una paradoja, habrá al menos que explicarla o explicitarla.

Por ello, se trate de una contradicción o de una paradoja, parece que la manera de proceder, en todo caso, será intentar ver si alguno de los extremos en colisión (PC o JSM) es falso. Con ello, los resultados posibles a los que llegar serán tres y no dos: el rechazo de PC, el de JSM, o la imposibilidad de deshacernos de ninguno de ellos y, así, la afirmación de

la paradoja como insoslayable. (Lógicamente también cabría la posibilidad de deshacernos de ambos, pero esto no parece nada razonable.)

Así es cómo básicamente Nagel presenta el fenómeno de la suerte moral. Es cierto que hay otros elementos importantes en su caracterización, pero los veremos más adelante. Paso ahora a presentar los aspectos fundamentales que Bernard Williams, por su parte, asocia a este fenómeno.

2.1.2. Williams: justificación retrospectiva y lamentación del agente

A. La justificación retrospectiva

El planteamiento que Williams realiza del problema de la suerte moral difiere un tanto del anterior. Su meta inmediata es tratar de mostrar el choque entre —como él las llama— la justificación racional y la justificación moral, lo cual haría inviable el seguir considerando el valor moral como el valor supremo e insuperable. Cabe aclarar que, en tanto que crítica de la *concepción kantiana de la moralidad*, como veremos, Williams asume la conexión ineludible que las nociones de moralidad, racionalidad, justificación y valor supremo, tienen en Kant y en la concepción de la moralidad de herencia kantiana.¹⁴

Con esa finalidad, nos presenta el caso de la decisión de Paul Gauguin; un experimento mental que hay que distinguir del Paul Gauguin real. El Gauguin de Williams debe decidir entre continuar con su familia,

¹⁴ Dice Williams: “La concepción kantiana conecta, y afecta a, un amplio espectro de nociones: moralidad, racionalidad, justificación y valor último o supremo. La conexión entre estas nociones, en la concepción kantiana, tiene un gran número de consecuencias para la evaluación reflexiva por parte del agente de sus propias acciones —por ejemplo, que, en el nivel último y más importante, no puede ser cosa de suerte que estuviese justificado a hacer lo que hizo. Es esta área la que quiero considerar. De hecho, diré muy poco hasta el final acerca de lo moral, concentrándome más bien en aspectos de la justificación racional.” (1981, p. 22.)

de la que se siente responsable y con la que vive feliz, o trasladarse a una isla de los Mares del Sur, donde piensa que podrá desarrollar su capacidad artística y convertirse en un gran pintor. Finalmente, se decidirá por la segunda opción, abandonando a su familia para emigrar al Pacífico. La cuestión que se plantea es: ¿podemos decir que la decisión que toma Gauguin está justificada? Esto es, ¿fue racional para él —dados sus intereses— actuar como actuó?¹⁵

Como es obvio, no es posible saber con antelación si la empresa tendrá éxito. La primera tesis de Williams es que, dada esta situación, la única cosa que justificará su decisión será el éxito mismo. Si falla, habrá actuado incorrectamente, “en el sentido de que haber hecho lo incorrecto en aquellas circunstancias no le da un fundamento para pensar que estaba justificado a actuar como actuó. Si tiene éxito, sí que tiene el fundamento para este pensamiento.”¹⁶ Así, la justificación, si la hay, será *esencialmente retrospectiva*.

Pero este éxito o fracaso de su proyecto está irremediamente sujeto, aunque sólo sea parcialmente, a la suerte. De modo que también en su justificación se verá implicada la suerte, por bien que no cualquiera tipo de suerte. Williams distingue entre la suerte intrínseca al proyecto y la suerte extrínseca. Ambas son necesarias para su éxito, y así para la justificación, pero solamente la suerte intrínseca está conectada con la *injustificación*. En el caso concreto del ejemplo de Gauguin, la suerte intrínseca se identifica con el tipo de suerte que se origina a partir de sus decisiones y acciones, pero no es preciso que ello sea siempre así. La suerte intrínseca al proyecto puede estar parcialmente fuera del agente. Es lo que

¹⁵ Precisando más, entenderé que la cuestión es si Gauguin estaba racionalmente (epistémicamente) justificado para pensar que actuar como lo hizo incrementaría sus opciones de convertirse en un gran pintor. Sobre esta cuestión, véase Latus (2001), §1.

¹⁶ Williams (1981), p. 23.

sucede en el caso de otro ejemplo que Williams nos propone: el de Ana Karenina, la protagonista de la novela de Tolstoi. Ana, tras ocho años de matrimonio, descubre su auténtico amor en un oficial del ejército llamado Vronski y decide dejar a su marido por él. La decisión podría haberse tornado errónea si, por ejemplo, Vronski se hubiera suicidado; sería un factor de suerte intrínseca —interna al proyecto— aunque no interna a la misma Ana.¹⁷ El fallo debido a la suerte extrínseca meramente supone, por así decirlo, el haber perdido; pero el fallo debido a la suerte intrínseca arruina la misma justificación del proyecto, y convierte a su protagonista en alguien roto. El “coste moral” es incomparable.

Con ello, Williams trata de mostrar que la justificación racional es en parte cuestión de suerte. Pero, ¿por qué hemos de aceptar que la justificación es “esencialmente retrospectiva”, aun reconociendo que ciertamente el éxito o fracaso del proyecto de Gauguin sólo puede conocerse retrospectivamente? Es decir, ¿por qué identificar justificación y resultado? Parece que hay un sentido por el cual una decisión está, sin más, justificada o no en el momento en que se toma —satisfechos los requisitos que una correcta deliberación exige. Sin embargo, el punto central de la cuestión es el de la fuerza que tiene en nuestras vidas la manera como resultan las cosas; y para tratar de mostrar esto Williams apela a la noción del “pesar-del-agente”.¹⁸

¹⁷ Véase Williams (1981), pp. 26-7.

¹⁸ Prefiero esta expresión a la de “lamentación del agente” (usada en las diferentes traducciones al castellano), como traslación de *agent-regret*. No obstante, no tengo ninguna razón de peso para esta preferencia. Más importante es caer en la cuenta de que el término inglés *regret* puede traducirse al castellano tanto por ‘lamento’ o ‘pesar’, como por ‘arrepentimiento’. La traducción más correcta, para la cuestión que nos ocupa, es sin duda ‘pesar’ (o ‘lamentación’), pero puede que Williams aproveche en su argumento la estrecha conexión de los dos sentidos del término inglés.

B. El pesar-del-agente

El “pensamiento constitutivo” del pesar es que “sería mejor que las cosas hubiesen sido de otro modo”¹⁹, en contraposición a como de hecho han sido. En concreto, el pesar-del-agente es un tipo de pesar que una persona sólo puede sentir en relación a sus propias acciones; consiste en el lamento que un sujeto siente con respecto a sus propias acciones, en tanto que *partícipe*. Se trata de asumir, de hacerse cargo, de la responsabilidad de una acción, con consecuencias negativas, que uno ha cometido y de tener la voluntad de enmendar la situación. Un punto fundamental aquí es captar la gran diferencia que se abre entre el pesar del propio agente y el de un mero espectador, en el sentimiento y en la expresión. La conciencia relevante de haber realizado un daño es básicamente la de algo que ha sucedido como consecuencia de los propios actos.

El ejemplo es el de un camionero; un camionero que va conduciendo su camión y que no se percata de estar atropellando a un niño que jugaba en la calle. Al conocer el suceso, el conductor sentirá una clase de pesar por la muerte del niño que nadie más puede sentir. Fue él mismo, al fin y al cabo, quien causó la muerte del niño. Ante este hecho, consideraríamos una reacción negativa, incluso inmoral, que el camionero no reaccionase con una profunda tristeza y con un fuerte pesar por la acción que ha cometido. Una reacción *sana* —según la terminología de Williams— consistiría en la sospecha por parte del agente de que había alguna cosa más que él podría haber hecho, que podría haber salvado la vida del niño: si hubiese estado un poco más alerta, o conducido más despacio; si hubiese visto al niño jugando cerca del camino, etc.

La cuestión reside, pues, en el reto que supone para la evaluación moral la inclusión de la visión retrospectiva. Es decir, si nos limitamos a

¹⁹ Williams (1981), p. 27.

considerar exactamente lo que es el caso cuando tomamos nuestras decisiones sin dejar sitio a las consecuencias, y la suerte que con éstas se introduce, entonces el camionero no tiene el *deber* de experimentar el pesar-del-agente, sino que simplemente tendría que decirse a sí mismo que hizo todo lo que pudo; pero así nos quedaríamos con una concepción *insana* de la racionalidad. Creo que este énfasis en la importancia de los resultados y las reacciones posteriores para la evaluación y la justificación, es correcto; y, ciertamente, ello atenta contra la primacía de la intención. Pero Williams va más allá y afirma que la misma idea de involuntariedad es superflua: el pesar no se limita a las acciones voluntarias, pues “[p]uede extenderse bastante más allá de lo hecho intencionalmente hasta prácticamente cualquiera cosa de la que se es causalmente responsable en virtud de alguna otra cosa que se hizo intencionalmente.”²⁰ Esperamos, afirma, que el pesar-del-agente se dé hasta en casos en los que no pensamos que el agente haya cometido falta alguna. Si pensamos que el conductor no podía haber hecho nada más para impedir la muerte del niño, intentaremos consolarlo; pero esto no quita que tuviésemos una visión negativa del conductor si él no mostrase ningún tipo de lamento, y se limitara a afirmar que ‘es un hecho terrible que haya sucedido, pero yo hice todo lo que pude por evitarlo’.

No obstante, parece que aquí se confunde *involuntario* con hacer algo sin falta, cuando alguien puede cometer una falta aun de manera involuntaria. En este sentido, quien lleva a cabo una acción y produce unas consecuencias negativas, incluso sin cometer falta alguna, parece que ha de lamentarse por el daño que ha cometido, pero su situación es bastante menos desdichada que la de quien se lamenta por una acción producto de una falta. La suerte afecta, en efecto, tanto a uno como a otro, y por eso

²⁰ Williams (1981), p. 29.

estoy de acuerdo en que no podemos “disociarnos completamente de los aspectos no intencionales de nuestras acciones”; pero distinciones como la de acciones que son fruto de una falta y acciones sin falta juegan un papel importante.

C. Contra la primacía de lo moral

En suma, la existencia del pesar-del-agente mostraría, según Williams, que hay un componente retrospectivo en la justificación racional, y ello impide que la justificación pueda ser inmune a la suerte. Pero sucede que ésta contrasta con la imagen recibida de la justificación moral, que la supone inmune a la suerte. En el caso de Gauguin, el choque entre estos dos tipos de justificación parece dirimirse a favor de la justificación racional. Si mostramos nuestra gratitud para con la decisión de Gauguin — que le permitió convertirse en un genio de la pintura, y ahora nos posibilita que gocemos de su arte—, ello implica que la moralidad no siempre prevalece. Estaríamos ante un caso en el que los valores morales han sido tratados como unos valores entre otros, y no como los valores incuestionablemente supremos.

En realidad, Williams nos pone ante un dilema: o el valor moral no es el tipo supremo de valor, o (a veces) es una cuestión de suerte. Ambas son opciones igualmente indeseables para la noción imperante de moralidad, pues ambas la ponen en cuestión.²¹ Si el valor moral no es completamente aislable de la suerte, o no es el tipo supremo de valor, entonces no puede constituir la forma última de justicia, y realizar el papel de “consuelo para un sentido de la injusticia del mundo”, que Kant le atribuía. Éste es un dilema importante al que es difícil escapar, y al que pare-

²¹ Sobre este punto véase Mendus (1988), pp. 334-5, y Latus (2001).

ce que todo escéptico acerca de la suerte moral habrá de responder. (Volveré sobre estos temas, con más detalle, en el capítulo 7.)

Con esto, creo que tenemos ya una caracterización amplia del fenómeno. No obstante, los aspectos expuestos no agotan aún la cuestión. Para alcanzar una comprensión más completa, antes de pasar a confrontar y explicitar las divergencias entre ambos planteamientos, en la sección siguiente expondré los tipos de casos en los que se pueden agrupar los ejemplos típicos de suerte moral y aclararé las diferentes maneras en que ésta se nos puede presentar.

2.2. Casos y variedades de la suerte moral

Hemos visto que Williams nos proponía tres ejemplos principales en su exposición de la suerte moral. El de Gauguin, sobre el carácter retrospectivo de la justificación y el conflicto entre justificación racional y justificación moral; el de Ana Karenina, presentado como una variación del anterior; y, finalmente, el del camionero que atropella a un niño, como paradigmático del pesar-del-agente. Por su parte, Nagel presenta también sus ejemplos, los cuales quieren mostrar que “[s]i tenemos éxito o fallamos en aquello que intentamos hacer casi siempre dependerá en alguna medida de factores que están más allá de nuestro control”.²² En esta sección acabaré de exponer los ejemplos clásicos, que se han convertido en paradigmáticos para la discusión del fenómeno, avanzando algunas claves de su clasificación; para, a continuación, ocuparme de la distinción entre los diferentes tipos de suerte moral.

²² Nagel (1979), p. 25.

2.2.1. Tipos de casos y el corolario de PC

El ejemplo que ha adquirido mayor fama es del conductor que, tras ingerir alcohol, decide conducir su coche, pierde el control e invade la acera. Dependiendo de si se encuentra o no a algún peatón en su camino, será moralmente afortunado o desafortunado. Si se lleva a alguien por delante, será juzgado más severamente, censurado por provocar una muerte y perseguido por homicidio (por imprudencia). Pero si no atropella a nadie, por mucho que su imprudencia sea la misma, será culpable de una ofensa legal mucho menos seria y recibirá una censura o reproche mucho menos severo. En todo caso, la única diferencia entre ambos, que *hace* más culpable al primero que al segundo, es el hecho de que alguien pasaba por allí en aquel preciso instante.²³ (Hay que notar que Nagel aporta este ejemplo comparándolo con el del camionero de Williams, al que niega la condición de caso de suerte moral, como veremos.)

Otro ejemplo *legal* lo constituye el hecho de que la pena por intento de asesinato es menor que la pena por asesinato, a pesar de que en los dos casos las intenciones y motivos puedan ser los mismos. El grado de culpabilidad dependería pues, al parecer, de si la víctima usaba chaleco antibalas o no, o si un pájaro desvió la trayectoria de la bala mientras volaba; es decir, de cosas que están fuera del control del agente.²⁴

Un último ejemplo de este tipo que citaré es el de una madre que negligentemente deja a su bebé dentro de la bañera con el grifo abierto. Cuando se percata de que su hijo puede haberse ahogado, comienza a subir corriendo las escaleras en dirección al baño. Al llegar, quizá el bebé se haya ahogado, y, así, lo que habrá hecho es moralmente terrible; mientras que si continúa sano y salvo, su descuido no será tan grave.

²³ Nagel (1979), p. 29.

²⁴ Nagel (1979), p. 29.

Hay que reparar en el hecho de que un elemento importante de estos ejemplos es que comparan la actuación de dos agentes (o de un mismo agente en situaciones diferentes) que, a pesar de llevar a cabo las mismas acciones, que son manifestación de las mismas intenciones, lo que acaba sucediendo es diferente a causa de factores más allá de su control; y, sin embargo, los tratamos o juzgamos diferentemente, y nos parece que estamos justificados a ello. Resulta relevante tener en cuenta el siguiente corolario de la que llamamos Principio de Control (PC), para el análisis de estos casos:

(Corolario de PC) Dos o más personas no deben ser moralmente evaluadas de manera diferente si las únicas diferencias entre ellas se deben a factores fuera de su control.

Cabe remarcar que es indiferente para nuestro tema si consideramos que los casos implican a dos personas similares en situaciones similares, o a la misma persona en diferentes momentos temporales, o se trata de dos situaciones diferentes contrafácticamente posibles.

Por otro lado, tenemos los casos comunes de “decisión bajo incertidumbre”. Son casos del tipo del de Ana Karenina al escaparse con Vronsky, o el de Gauguin cuando abandona a su familia, Chamberlain firmando el acuerdo de Munich, los decembristas rusos persuadiendo a las tropas bajo su mando para sublevarse contra el Zar, las colonias americanas al declarar la independencia a Gran Bretaña, o el de alguien que presenta a dos personas intentando que se emparejen.²⁵ Son todos ellos casos en los que los sujetos tienen que afrontar una elección difícil debido a que el resultado no puede preverse con certeza. Pero veamos más detenidamente algunos de estos casos.

²⁵ Nagel (1979), p. 29-30.

Pensemos en los líderes de un grupo revolucionario que toman una decisión y lanzan un ataque contra un dictador. Es claro que esta decisión no puede basarse en el conocimiento cierto de cómo acabará todo; es una decisión que siempre implica un cierto riesgo. Si la insurrección tiene éxito, se obtendrá la libertad y los insurrectos pasarán a ser héroes nacionales; pero si, por el contrario, el intento queda frustrado y los dirigentes del régimen tiránico responden con una represión sangrienta que provoca un coste amplio en vidas y sufrimiento, los revolucionarios recibirán un juicio moral bien distinto. Así, en muchos casos, la elección es muy difícil a causa del hecho de que el resultado no puede ser previsto con certeza. Nagel considera que:

Es tentador en todos estos casos sentir que alguna decisión debe ser posible, a la luz de lo que se conoce en aquel momento, que hará el reproche inapropiado sin importar como resulten las cosas. Pero eso no es verdad... cómo resultan las cosas determina lo que él ha hecho.²⁶

Dentro de ésta misma categoría se sitúa el fracaso de los decembristas rusos en su esfuerzo por derrocar a Nicolás I para establecer un régimen constitucional, que podemos contrastar con el éxito de la revolución americana al conseguir independizarse de los británicos. Si comparamos ambos casos, vemos que la derrota de los primeros comportó los terribles castigos infligidos por las autoridades zaristas sobre las tropas insurrectas, de lo cual fueron responsables los dirigentes decembristas que instigaron a la rebelión; mientras que Jefferson, Franklin, Washington y compañía tuvieron éxito en su objetivo y se convirtieron en héroes nacionales, siendo muy otra su responsabilidad.

²⁶ Nagel (1979), pp. 29-30.

Un tipo diferente de casos es aún el representado por el ejemplo de dos alemanes con simpatías nazis. De estos dos alemanes, uno permanece en Alemania durante el período nazi y desarrolla sus simpatías, hasta el extremo de llegar a convertirse en oficial de un campo de concentración; mientras que el otro se vio arrastrado a emigrar a Argentina, por motivos económicos, antes de la II Guerra Mundial y no tuvo la oportunidad de hacer efectivas sus simpatías.²⁷ Podemos suponer que si éste hubiese permanecido en Alemania podría también haber llegado a convertirse en oficial de un campo de concentración como su álgter ego. Sin embargo, ¿queremos decir que deberíamos juzgar al expatriado tan duramente como al otro?

Una variación de este ejemplo apunta a los ciudadanos corrientes de la Alemania nazi. Éstos tuvieron la oportunidad de comportarse heroicamente oponiéndose al régimen, pero también de colaborar; y muchos de ellos se convirtieron en culpables por no superar esta prueba. Sin embargo, como dice Nagel, es una prueba muy exigente a la que sólo algunas personas han tenido la mala suerte de ser sometidas.²⁸

Una cuestión ulterior es la planteada por la posesión de unos determinados rasgos temperamentales y no otros, con sus implicaciones en la conducta moral. Hay personas con una afección moral fuerte, que atienden a los requerimientos morales sin ningún esfuerzo. Ello se hace especialmente evidente en comparación con personas que poseen una sensibilidad moral escasa y que no sufren mucho por la injusticia hacia los demás, aunque ellos mismos sean los causantes. Podemos confrontarlo también con el caso de una persona de voluntad débil, que falla a menudo en el desempeño de sus intenciones. Asimismo, cabe considerar el caso de

²⁷ Nagel (1979), p. 26.

²⁸ Nagel (1979), p. 34.

personas envidiosas, cobardes, frías, poco amables, presumidas, que sin embargo logran comportarse perfectamente a causa de un enorme esfuerzo de su voluntad. En definitiva, en la medida en que hay un componente innato en el temperamento o rasgos de carácter de una persona, hay algo que ésta no controla pero por lo que sigue siendo moralmente evaluada.²⁹

En suma, hay una gran variedad de casos implicados en el fenómeno de la suerte moral. Dada esta diversidad, parece que alguna clasificación es necesaria, tanto para proceder más ordenadamente como para detectar pautas en las diferentes maneras en las que la suerte puede (supuestamente) inmiscuirse en el juicio moral. En la siguiente subsección, discutiré la clasificación nageliana de los tipos de suerte moral, que ha conformado el marco general del debate, y propondré una modificación que la matiza y hará más clara la discusión subsiguiente.

2.2.2. Los cuatro tipos nagelianos; discusión y modificación

Tanto Nagel como Williams se refieren a diferentes maneras por las que el objeto de nuestra valoración moral puede quedar sujeto a la suerte. Los cuatro tipos originalmente descritos por Nagel son los siguientes:

Suerte moral resultante (o consecuencial): es la suerte moral envuelta en cómo resultan las cosas; o, también, la suerte moral que marcará el resultado de las acciones y proyectos de un agente. El ejemplo paradigmático aquí es el de los dos conductores ebrios, donde se pone en juego el factor de la *negligencia*; aunque tenemos también el caso del asesino frustrado y

²⁹ Nagel (1979), pp. 32-3.

el asesino exitoso. Por otro lado, están los casos de decisión bajo *incertidumbre*, cuyo ejemplo central era el del acto revolucionario.

Suerte moral circunstancial (o situacional): es la suerte moral debida a las circunstancias en las que alguien tiene que actuar; o, también, la suerte de estar en el lugar correcto o incorrecto en el momento correcto o incorrecto, lo cual puede tener profundos efectos en la manera en que somos moralmente evaluados. El ejemplo destacado es el del contraste entre los dos alemanes pro-nazis. Se trataría de un caso de suerte moral circunstancial porque es el hecho de estar en unas circunstancias y no en otras lo que parece marcar la diferencia moral.

Suerte moral constitutiva: la suerte moral de ser quien se es, de tener las disposiciones que se tienen. Es la suerte involucrada en el hecho de que una persona posea las “inclinaciones, capacidades y temperamento” que de hecho posee. Tiene que ver con el propio carácter que una persona tiene, y que no depende completamente de ella. Nagel afirma que ciertos rasgos del temperamento de una persona, como la simpatía o la frialdad, constituyen un trasfondo respecto al cual la respuesta a los requerimientos morales será más o menos difícil. El carácter de una persona se constituye, en cierto grado, por calidades temperamentales innatas; pero también, en algún otro grado, son fruto de elecciones anteriores, y pueden ser susceptibles de cambio por las acciones presentes. En cualquier caso, el haber recibido unas cualidades temperamentales y no otras es fruto de la suerte. Al igual que hay quien tiene una dotación genética que lo hace, en principio, más apto para el trabajo intelectual, parece que hay quien tiene un temperamento que lo hace más sensible moralmente o que le predispone a dar respuestas moralmente más adecuadas.

Suerte moral causal: la suerte en el modo en que alguien es determinado por las circunstancias antecedentes. Es el tipo de suerte moral que, tanto Nagel como otros pensadores, han asimilado al debate tradicional sobre el libre albedrío y el determinismo. El problema, tal y como Nagel lo refiere, surge porque a menudo parece que nuestras acciones —incluso los mismos “actos de nuestra voluntad”— son consecuencia de lo que no está bajo nuestro control. Si eso es así, ni nuestras acciones ni nuestras voliciones son libres. Y si se piensa que la libertad es necesaria para la responsabilidad moral, no podemos ser moralmente responsables ni tan siquiera de lo que queremos. Hay que destacar, no obstante, que esta categoría de la suerte causal parece redundante, pues su área queda ya cubierta con la combinación de la suerte circunstancial y la suerte constitutiva. De hecho, es el tipo de suerte menos tratado en la bibliografía específica de la suerte moral, costumbre que será aquí continuada.³⁰

En realidad, Nagel identificó estos cuatro tipos de suerte moral, pero sin dar nombre a todos ellos. De hecho, empleó sólo la etiqueta de “suerte constitutiva”, que toma de Williams;³¹ quien, por otro lado, le concede un alcance más amplio, al incluir la influencia de la suerte en los

³⁰ Véase el Apéndice a la Parte I, sobre las conexiones entre la suerte moral y el debate del libre albedrío y la responsabilidad moral.

³¹ Los nombres de suerte *circunstancial* y suerte *causal* son originarios de Statman (1993), p. 11; mientras que el de suerte *resultante* procede de Zimmerman (1993), p. 219, que también ha sido llamada ‘suerte consecuencial’. La suerte circunstancial y la suerte causal han recibido los nombres alternativos de ‘suerte situacional’ —Walker (1991), p. 235— y ‘suerte determinante’—Mendus (1988), p. 334—, respectivamente. Tomo esta información de Latus (2001), nota 10.

motivos, las intenciones y la personalidad, y contraponerla con la que llama suerte *incidental*.³²

De esta última clasificación, se encuentra muy próxima la utilizada por Michael Zimmerman, quien agrupa los diferentes tipos primitivos, o nagelianos, en dos clases principales.³³ Identifica, por una parte, la suerte *situacional*, es decir, la suerte con respecto a las situaciones afrontadas, incluida la naturaleza del propio carácter (inclinaciones, capacidades, etc.) en tanto que formada; que incorporaría, pues, la “suerte en las circunstancias” (o circunstancial), la “suerte en cómo alguien es determinado por las circunstancias antecedentes” (o causal) y, en parte, la “suerte constitutiva”. Y, por otro lado, la suerte *resultante*, descrita por Nagel como la “suerte en la manera como resultan las acciones y proyectos de una persona”; es la suerte con respecto a lo que resulta de las decisiones, acciones u omisiones de alguien.

NAGEL	WILLIAMS	ZIMMERMAN	DWORKIN/HURLEY
Causal	Constitutiva	Situacional	Bruta / en las causas
Constitutiva		Resultante	Opcional / en los efectos
Circunstancial	Incidental		
Resultante			

Tabla 1. Tipos de suerte

³² Hay que notar que esta distinción difiere de aquella otra que veíamos más arriba, entre suerte intrínseca y suerte extrínseca al proyecto.

³³ Véase Zimmerman (1987), p. 219, n. 7; y (2002), p. 559.

Ulteriormente, hay una agrupación alternativa de los tipos de suerte, formulada por Susan Hurley, que distingue la suerte *en las causas* y la suerte *en los efectos*. La primera abarca los tipos nagelianos de la suerte causal y la suerte constitutiva, y la segunda comprende la suerte circunstancial y la suerte resultante.³⁴ (Véase en la Tabla 1 la comparación entre los distintos tipos de suerte establecidos por los principales autores considerados.) Pero todavía podríamos intentar hilar más fino e identificar nuevos tipos de suerte, con ulteriores subdivisiones. Michael Moore, por ejemplo, introduce la distinción entre “suerte en la ejecución” y “suerte en la planificación”.³⁵

Por otro lado, no debemos caer en el error de pensar que estamos ante modos totalmente independientes y claramente distintos en que la suerte puede influir en nuestros juicios morales. Consideremos, por ejemplo, el caso de los dos pro-nazis. Este caso suele aducirse como ejemplo paradigmático de suerte circunstancial; pero según el punto de vista desde el que lo consideremos, esta suerte circunstancial puede parecerse más a la suerte constitutiva. Lo que supuestamente marca la diferencia en el ejemplo es el hecho de que ambos contrapartes vivan en ambientes diversos, lo que puede ser interpretado como suerte constitutiva, en sentido amplio (o suerte antecedente; en particular, suerte en la formación) o como suerte circunstancial (porque lo que marca la diferencia serían las cir-

³⁴ Véase Hurley (2003), cap. 4, §2. Esta distinción está estrechamente relacionada con la conocida distinción de Ronald Dworkin entre *brute luck*, o suerte que está por completo más allá del control del agente —un tipo de suerte que tiende a ser o constitutiva o circunstancial—, y *option luck*, o suerte originada en el contexto de lo que el agente hace, o podría hacer. Ejemplos destacados de la primera serían el país donde se nace o los padres que nos tocan, ser ciego o sordo de nacimiento, etc.; y de la segunda, ganar la lotería, tener buena fortuna en el amor, escoger una carrera que resulta ser satisfactoria, etc. Véase Dworkin (2000), p. 73.

³⁵ Véase Moore (1997), p. 235.

cunstances, en oposición a sus constituciones originales). Asimismo, otros ejemplos de suerte circunstancial, como por ejemplo el del hombre que pretende matar a alguien pero no llega a disparar porque su pistola se encasquilla (aducido como tal por Zimmerman), más bien parece un caso de suerte resultante. En general, hay modos muy distintos por los que la suerte puede presuntamente inmiscuirse en nuestros juicios morales, y en muchos casos los tipos particulares identificados dependerán del punto de vista y propósito desde el que describamos una misma acción o situación. En particular, cuando consideramos un escenario desde un punto de vista estático, todo aquello con lo que el sujeto llega a ese escenario suele ser descrito como suerte constitutiva —o mejor, *suerte antecedente*—; aquellos factores que interactúan con el agente en dicho escenario constituyen la suerte circunstancial, y finalmente aquello que resulta de las acciones del agente será la suerte resultante. Pero si atendemos la historia personal de un agente, entonces la suerte constitutiva será sólo aquello que pertenece a su constitución original, mientras que las experiencias que va acumulando, que no son más que las circunstancias a las que se va enfrentando, constituirán un factor de suerte circunstancial —o mejor, de *suerte formativa*.³⁶

Creo que de las anteriores consideraciones podemos extraer dos cosas: una conclusión y una nueva clasificación. La conclusión es que la distinción entre los diferentes tipos de suerte es en gran parte contextual —si bien cabe reconocer que la suerte resultante parece ser el tipo más independiente. La nueva clasificación que quiero proponer, y que seguiré a lo largo de este trabajo, distingue como tipos más básicas: la suerte

³⁶ Este tipo ha sido introducido por Nafsika Athanassoulis; véase Athanassoulis (2005), p. 22 y p. 173, n35. Ella utiliza la expresión “suerte en el desarrollo” (*developmental luck*), que incluye todos los factores que influyen en el desarrollo moral del agente y que escapan a su control.

constitutiva, la suerte formativa, la suerte circunstancial y la suerte resultante. Estos tipos pueden agruparse del siguiente modo:

1. *Suerte antecedente* (o *suerte constitutiva* en sentido amplio).
 - a. *Suerte constitutiva 'estricta'*
 - b. *Suerte formativa*
2. *Suerte circunstancial* ('estricta')
3. *Suerte resultante*

Cabe precisar que también podría seguirse esta clasificación alternativa: la suerte circunstancial 'amplia' englobaría a la suerte formativa más la suerte circunstancial 'estricta'. En cualquier caso, habrá cuatro tipos básicos de suerte moral, en virtud de los cuales organizaré la discusión subsiguiente. No obstante, dentro de cada tipo, es aun posible distinguir entre casos significativamente diferentes; esto es particularmente destacable con respecto a la suerte circunstancial, entre casos en los que la acción del agente está muy próxima, que se parecen más a casos de suerte resultante, y casos de suerte circunstancial más remota, que más bien parecen casos de suerte antecedente o *formativa*.³⁷ Por otro lado, esta nueva clasificación no altera significativamente la discusión, pues no rompe con la distinción original de Nagel sino que sólo la reorganiza. (Omito las definiciones, que ya han quedado claras, y excluyo la suerte causal por las razones anteriores.) La tabla siguiente (Tabla 2) contrasta mi clasificación con las de Nagel y Williams.

³⁷ Comparar con la quintuple distinción de Enoch y Marmor (2007). Ellos añaden una distinción entre dos tipos de suerte constitutiva (entre aquella referida a rasgos de carácter fácilmente modificables y controlables y aquella referida a rasgos de carácter sobre los que no tenemos control). Es una distinción también presente en otros autores (véase, especialmente, Zimmerman, 1998, 2002) y a la que irremediablemente llegaré en el transcurso de mi discusión de la suerte constitutiva, en el capítulo 4.

NAGEL	WILLIAMS	MI PROPUESTA	
Constitutiva	Constitutiva	Antecedente	Constitutiva
Circunstancial			Formativa
Resultante		Circunstancial (estricta)	
	Incidental	Resultante	

Tabla 2. Tipos de suerte: mi propuesta, en contraste

2.3. Divergencias entre Nagel y Williams

Una vez vistos los principales elementos que articulan el debate en torno a la cuestión, paso ahora a contraponer los planteamientos de Nagel y Williams, con la meta de sacar a relucir los diferentes énfasis e intereses de ambos. Con esto, podremos acabar de componer el escenario del problema.

Bernard Williams ha afirmado que:

cuando primero introduje la expresión *suerte moral*, esperaba sugerir un oxímoron. Hay algo en nuestra concepción de la moralidad, con lo que Tom Nagel concuerda, que despierta la oposición a la idea de que la responsabilidad moral o el mérito moral o la censura moral hubieran de estar sujetos a la suerte.³⁸

De hecho, si alguien considera que la moralidad y la suerte son cosas anti-téticas entenderá que “suerte moral” es una noción autocontradictoria; le parecerá, como dice Nagel, algo absurdo e irracional.³⁹

³⁸ Williams (1993), p. 251.

³⁹ Véase Nagel (1979), p. 31.

Por otro lado, tanto Williams como Nagel sitúan como figura destacada de sus disquisiciones la concepción de la moralidad atribuida a Kant. Williams es, sin embargo, un crítico más severo. Su objetivo es desacreditar la noción misma de moralidad, en tanto que identificada con ciertos elementos nucleares de la concepción kantiana. Williams destaca que ciertas posiciones filosóficas importantes —cuyo paradigma son los estoicos— muestran una tendencia a aislar el dominio propio del sujeto de lo que está fuera de su control, huyendo de la sujeción a la suerte y a las contingencias, enemigas de la tranquilidad. Aunque la idea de que la vida entera pueda ser inmune a la suerte, no ha prevalecido, su lugar habría sido ocupado por una idea de su misma especie, procedente de Kant y poderosamente influyente, según la cual hay una forma básica de valor, el valor moral, que es inmune a la suerte e “incondicionado”. Para Kant, el juicio debe fijar su objeto en el producto de la voluntad incondicionada; sólo así puede el kantismo auto-presentarse como “consuelo a un sentido de la injusticia del mundo”, ofreciéndonos una justicia última donde asirnos.⁴⁰

Nagel, por su parte, toma a Kant como principal valedor del principio de control —uno de los extremos de su paradoja—, en una forma particularmente estricta. Kant creía, nos dice Nagel, que la buena o mala suerte no debía influir en nuestro juicio moral de una persona y de sus acciones, ni en su autoevaluación moral.⁴¹ Y añade: “Esta opinión parece

⁴⁰ Véase Williams (1981), pp. 21-2. En realidad, la crítica de Williams al kantismo ha sido muy recurrente en toda su obra, y ésta no es más que una parte de ella —y ello no le ha impedido ser también un crítico destacado del utilitarismo. Véase, en particular, Williams (1985).

⁴¹ De hecho, Nagel empieza su artículo citando el famoso pasaje de la *Fundamentación* en el que Kant caracteriza la buena voluntad: “*La buena voluntad* no es buena por lo que efectúe o realice, no es buena por su adecuación para alcanzar algún fin que nos hayamos propuesto; *es buena sólo por el querer, es decir, es buena en sí misma*. Considerada por sí misma, es, sin comparación, muchísimo más valiosa que todo lo que por medio de

estar equivocada, pero surge en relación al problema fundamental de la responsabilidad moral para el cual no tenemos una solución satisfactoria.”⁴²

A pesar de la coincidencia, hay un matiz que cabe señalar. Para Williams, Kant, o más concretamente la concepción kantiana de la moralidad, es algo que hay que combatir hasta su descrédito. Mientras que Nagel —aunque la considere, en general, falsa— le reconoce algo que considera intuitivamente irrenunciable, a saber, su empeño por rastrear del motivo y la intención en la valía del agente —lo cual va de la mano del anhelo del mismo Nagel por salvar el principio de control. En todo caso, el reconocimiento del fenómeno pone en cuestión la concepción heredada de la moralidad, y en especial sus elementos más *kantianos*, y *estoicos* (por bien que ello no suponga, en sí mismo, una refutación de la letra de Kant).

Una muestra más concreta de las divergencias entre ambos es el rechazo de Nagel de los dos casos principales esgrimidos por Williams. En primer lugar, niega que el caso de Gauguin funcione. Para Nagel, Williams no logra explicar por qué este tipo de actitudes retrospectivas pueden ser llamadas *morales*. Si Gauguin no puede justificarse ante los demás y aún siente esta clase de sentimientos, eso demuestra que no es preciso que sean “sentimientos morales”. Por otro lado, Nagel no acepta tampoco el ejemplo del camionero como un caso de suerte moral. En su

ella pudiéramos verificar en provecho o gracia de alguna inclinación y, si se quiere, de la suma de todas las inclinaciones. Aun cuando, por particulares enconos del azar o por la mezquindad de una naturaleza madrastra, le faltase por completo a esa voluntad la facultad de sacar adelante su propósito; si, a pesar de sus mayores esfuerzos, no pudiera llevar a cabo nada y sólo quedase la buena voluntad —no desde luego como un mero deseo, sino como el acopio de todos los medios que están en nuestro poder—, sería esa buena voluntad como una joya brillante por sí misma, como algo que en sí mismo posee su pleno valor. *La utilidad o la esterilidad no pueden ni añadir ni quitar nada a ese valor.*” (Sección I, Ak. IV: 411-ss. Traducción de García Morente; mis cursivas.)

⁴² Nagel (1979), pp. 24-5.

opinión, es necesario un elemento más, que está ausente, para que se pueda hablar de suerte moral: la *negligencia* del agente. Si el camionero fuera culpable de la más mínima negligencia y ésta contribuyese a la muerte del niño —como, por ejemplo, no haber revisado recientemente los frenos del camión, cuando tenía la obligación de hacerlo—, entonces el camionero no sólo tendría que sentirse enormemente mal, sino que debería culparse de la muerte.⁴³ La fuente de divergencia se encuentra en el hecho de que Nagel acepta el requisito de control como irrenunciable, y de aquí el carácter paradójico insalvable de la suerte moral, mientras que Williams considera que “la idea de lo voluntario [*voluntary*] es esencialmente superficial.”⁴⁴

Por otro lado, cabe destacar también la manera diferente de plantear el problema por parte de uno y otro. Nagel establece una formulación muy clara y directa, en la que vemos en seguida cuál es el problema exacto y la tensión entre sus dos elementos principales. Mientras que el planteamiento de Williams es mucho más oscuro. Como él mismo reconoce en su “Postscript”, su artículo ha sido mal comprendido a causa, básicamente, de no haber distinguido suficientemente las tres cuestiones planteadas; que son éstas:

1. La limitación de la importancia de la moralidad (en el sentido estrecho) en contraste con el sentido más amplio de lo ético.
2. La limitación de la importancia, para un agente y para la visión de ciertos agentes, de lo ético incluso en el sentido más amplio, en contraste con el resto del ámbito práctico.
3. La justificación retrospectiva, que va más allá de la racionalidad práctica.⁴⁵

⁴³ Nagel (1979), p. 28, n3; p. 29.

⁴⁴ Williams (1993), p. 253.

⁴⁵ Véase Williams (1993), pp. 255-6.

Así, la confusión en su exposición deriva en parte de su ambición en el tratamiento del problema. Como veremos, lo que parece estar en juego es la supremacía de lo moral contra lo ético y lo práctico, por un lado, y la significatividad de la justificación retrospectiva, por otro.

Hay todavía otra diferencia importante entre la forma de abordar la cuestión de ambos: la perspectiva en el juicio moral. Nagel se interesa por el juicio moral que emitimos acerca de otras personas. Mientras que el interés principal de Williams reside en la influencia que la suerte puede tener en la evaluación reflexiva por parte del agente de sus propias acciones; esto es, en la habilidad del agente para justificar sus propias decisiones y acciones, ante sí y ante los demás. Así, si por un lado Nagel se ocupa de cuestiones concernientes a la responsabilidad moral desde un punto de vista “objetivo” o de tercera persona, esto es, el punto de vista de un espectador que juzga a otro agente desde fuera; por el otro, Williams se centra en la perspectiva de la primera persona, en los juicios hechos por el agente mismo que reflexivamente desea evaluar sus acciones. No en vano, Nagel, a diferencia de Williams, parte del supuesto de que la justificación moral es siempre la justificación de uno mismo hacia los demás. Para Nagel la justificación moral es “objetiva e intemporal”; para Williams, tiene que ver con los sujetos y las circunstancias implicados.

Un último contraste que quiero destacar entre ambos apunta a su diferente diagnóstico del problema. Nagel entiende el problema de la suerte moral como concomitante de nuestra concepción de la agencia. En particular, se trata del conflicto entre la visión interna de la agencia, conectada con las actitudes morales, que extendemos desde nosotros mismos a los demás, y la visión externa, implicada en la consideración de las consecuencias, que nos remite al reconocimiento del hecho de que somos parte del mundo. Habría algo en la misma noción de agencia que es in-

compatible con la consideración de las acciones como acontecimientos, o de las personas como cosas, pero al mismo tiempo nos vemos abocados a reconocer este hecho, y de aquí el carácter paradójico.⁴⁶ Sin embargo, Williams responde a Nagel afirmando su duda respecto a que las dificultades se generen sólo a partir del problema filosófico de la naturaleza misma de la acción.⁴⁷ Lo que demandamos a la idea de lo voluntario tiene fuentes éticas; necesitamos marcar unos límites en la voluntariedad para no censurar a las personas injustamente. En efecto, no responderemos correctamente a la cuestión atendiendo sólo al problema metafísico de qué es una acción por oposición a un evento o de cómo un agente puede producir cambios en el mundo, sino que es imprescindible considerar también los aspectos epistémicos y prácticos implicados.

2.4. La articulación del debate: ¿prácticas o control?

Para finalizar este capítulo, recogeré ciertas conclusiones de la exposición anterior y presentaré el mapa general del debate posterior sobre la suerte moral. Ello nos ayudará a comprender las diversas estrategias de respuesta a la cuestión, así como a sugerir cómo podría orientarse más adecuadamente la discusión.

Tras lo visto, hay que reconocer que estamos ante una cuestión compleja, con incontables repercusiones. Los artículos de Nagel y Williams —las dos fuentes principales de la discusión contemporánea— presentan la cuestión desde puntos de vista diferentes y poniendo el énfasis en elementos distintos, por bien que con un trasfondo claramente compar-

⁴⁶ Véase Nagel (1979), pp. 37-8.

⁴⁷ Véase Williams (1993), p. 253.

tido. Ante esto, cabe señalar que el debate se ha centrado más bien en el planteamiento del primero: en concreto, la bibliografía se articula en torno al conflicto entre el principio de control y ciertas prácticas de juicio ordinarias. Por esta razón, mi discusión subsiguiente de la cuestión se centrará principalmente en el planteamiento de Nagel, lo cual no significa que el planteamiento específico de Williams no haya recibido una gran atención, o que vaya a ser dejado de lado aquí. Los temas del sentimiento moral y del carácter retrospectivo de la justificación, en estrecha conexión con la evaluación moral, así como la cuestión de la supremacía o no del valor moral sobre el resto de valores, o el juicio moral de primera persona o autoatribución de responsabilidad moral, serán de lo más relevantes en la discusión que seguirá.⁴⁸

En resumidas cuentas, el debate en torno a la cuestión de la suerte moral nos sitúa en un escenario en el que se plantea un conflicto entre la intuición general, en principio altamente plausible, de que la condición de control es un requisito para la atribución de la responsabilidad moral, por un lado; y la práctica ordinaria de juzgar moralmente a las personas por factores que están más allá de su control, por otro. Para tratar de salvar esta situación conflictiva, aparecen como disponibles dos vías principales, más o menos inmediatas: a grandes rasgos, o bien nos decantamos por el lado del Principio de Control e intentamos aplicarlo consistentemente, o bien salvamos nuestros juicios morales ordinarios rechazando el Principio de Control.

⁴⁸ Cabe señalar que el interés por el punto de vista del juicio moral de tercera persona, externo al propio sujeto evaluado, ha prevalecido en el debate, por encima del punto de vista de la autoevaluación o juicio de primera persona. Sin embargo, ambas perspectivas son igualmente relevantes tanto para nuestra autocomprensión como para la comprensión de las relaciones interpersonales en que estamos inmersos.

Cabe reconocer que el mismo Nagel ya propuso ambas estrategias, pero las juzgó igualmente irrealizables. Pues por un lado, si intentamos aplicar consistentemente el Principio de Control, erosionaríamos, e incluso haríamos ilegítimos, la mayoría de nuestros juicios morales ordinarios. En particular, las “cosas por las que las personas son juzgadas moralmente están determinadas por cosas que están más allá de su control de más maneras de las que somos conscientes a primera vista.”⁴⁹ De hecho, parece que muy pocos de nuestros juicios morales prerreflexivos quedarían intactos, con lo que el sacrificio sería ciertamente demasiado alto. Alternativamente, podríamos concluir que el Principio de Control es erróneo, y así salvar nuestros juicios morales ordinarios. No obstante, esta salida tampoco nos es asequible dada la importancia que el control tiene para nosotros:

La condición de control no indica en ella misma una mera generalización de ciertos casos claros. (...) La erosión del juicio moral emerge no como la absurda consecuencia de una teoría simplificada, sino como una consecuencia natural de la idea ordinaria de evaluación moral, cuando es aplicada desde una caracterización más completa y precisa de los hechos.⁵⁰

A mi modo de ver, hay aquí una verdadera tensión entre el principio de control y las prácticas ordinarias de juicio, pero creo que la paradoja puede superarse investigando la fuerza de nuestras intuiciones y, en especial, la naturaleza de su funcionamiento; es decir, si logramos entender cómo el modo de presentación de los casos hace que nuestras intuiciones sean unas u otras, lo cual está estrechamente relacionado con la contraposición entre principios abstractos y situaciones concretas y con la

⁴⁹ Nagel (1979), p. 27.

⁵⁰ Nagel (1979), p. 27.

posible incoherencia dentro del conjunto de nuestras intuiciones. Pero dejemos esto para más adelante. En todo caso, creo que no vale con meramente renunciar o desechar la idea de control, que no necesariamente es un reflejo estricto de lo voluntario, sino que hay que tomársela muy en serio si queremos que el debate sea fructífero —y ello aun a pesar de que finalmente pueda acabar sufriendo una importante deflación.

Tenemos, pues, dos grandes vías de posible solución o explicación del problema o fenómeno, las cuales han constituido el marco del debate. Por un lado, una mayoría de autores ha negado que exista tal cosa como la suerte moral, más allá de las meras apariencias. Los presuntos casos de suerte moral no serían tal cosa, pues en ellos la suerte no marcaría ninguna diferencia moral, sino de otro tipo, que no afecta a la consideración moral de una persona. Sin embargo, otros han aceptado su existencia mediante el rechazo o restricción del principio de control y han denunciado la que consideran una concepción equivocada de la moralidad. Básicamente, se puede decir que unos afirman que interpretamos erróneamente nuestras prácticas y otros que es la intuición la que está equivocada. Pero tanto unos como otros tienen que afrontar diferentes dificultades para hacer valer sus posiciones. En concreto, quienes niegan la existencia de la suerte moral deben explicar por qué parece haber tal cosa, y hacer un retrato plausible y coherente de cómo evitar que la suerte se inmiscuya en nuestras evaluaciones morales. Mientras que aquellos que aceptan, en general, la existencia de la suerte moral deben explicar, por su parte, cómo podemos revisar nuestros juicios y prácticas de un modo coherente o mostrar que no estamos comprometidos, en primer término, con el principio de control.⁵¹

⁵¹ Cf. Nelkin (2004) y Latus (2001).

Pero este escenario aparentemente dual es en realidad mucho más complejo, ya que en cada bando hay estrategias muy diversas, además de las propias diferencias internas. Cabe, en primer lugar, distinguir la postura tomada frente a la existencia o no de la suerte moral, de la misma evaluación que se haga de su presunta naturaleza paradójica. Una cuestión es la aceptación o negación de la existencia de la suerte moral como un fenómeno real y otra reconocer o no que se trate de una paradoja. Uno puede, de hecho, reconocer el fenómeno de la suerte moral, pero negar que se trate de una verdadera paradoja, e incluso de un verdadero problema.

Por otro lado, un elemento que añade especial complejidad a la cuestión es la presunta existencia de tipos diversos de suerte moral; lo que hace que pueda decirse que no estamos tanto ante *un* problema, sino ante *distintos* problemas. La relativa independencia entre las distintas clases de suerte moral dificulta el posicionamiento y la confección de los argumentos. Especialmente, los críticos de la suerte moral deben mostrar que cada uno de estos tipos de suerte no tiene una incidencia real sobre nuestros juicios morales, más allá de las apariencias. Esto da origen a que algunos de los críticos se centren principalmente, o incluso de manera exclusiva, en sólo uno o dos tipos de suerte moral, que consideran más cuestionables y a los que sus argumentos van especialmente dirigidos, intentando aplicarlos posteriormente a otros tipos, o simplemente no pronunciándose sobre el resto. Cabe decir que las estrategias más ambiciosas son aquellas que intentan rechazar o defender la suerte moral en todos sus tipos y por medio en una línea argumentativa común. Así, el rechazo de la suerte moral resultante, y circunstancial, tiene que ser completada con el descarte de la suerte formativa y constitutiva. O, en el otro bando, la defensa de la suerte antecedente debe completarse con la defensa de la suerte resultante. Pero no está nada claro que pueda seguirse el mismo tipo de estrategia

para todas ellas. Por ello, también se han propuesto estrategias mixtas, esto es, estrategias que combinan argumentos diversos dirigidos a tipos particulares. O, alternativamente, la defensa de unos tipos de suerte moral y el rechazo de otros. (A modo de estipulación, en lo que sigue llamaré estrategias *mixtas* a aquellas que combinan diferentes argumentos para la defensa o rechazo de los tipos diversos de suerte moral, y estrategias *híbridas* a las posiciones que combinan la defensa de unos tipos con el rechazo de otros.)

Para finalizar este capítulo, quiero adelantar cuál será, a grandes rasgos, mi lugar en el debate. En el resto de esta investigación defenderé que la suerte moral es un fenómeno real en todos sus tipos, afirmando las prácticas sobre el principio. Mi estrategia argumentativa general consistirá en defender, en primer lugar (capítulo 3), que el caso global contra la suerte moral falla, y no puede más que fallar; y en hacer valer, a continuación (capítulos 4, 5, 6 y 7), cada uno de los diferentes tipos de suerte moral aportando argumentos particularizados.

Dónde estamos y adónde vamos

En este capítulo he presentado en detalle el fenómeno de la suerte moral, examinando y comparando las formulaciones de Nagel y Williams. También he discutido los diferentes tipos nagelianos de suerte moral, y he rechazado la idea (más o menos implícita) de que estamos ante modos claramente distintos e independientes en que la suerte puede influir en nuestros juicios morales de los demás. Debe quedar claro que, en principio, hay diferentes maneras en que esto puede ocurrir, y que diferentes descripciones de la misma acción pueden convertirla en ejemplo de tipos

diferentes de suerte moral —las descripciones dependerán de la perspectiva y el propósito que tengamos.

En los próximos capítulos presentaré y discutiré la posición de aquellos que niegan la existencia de la suerte moral, considerando que se trata de un fenómeno que, tras ser adecuadamente comprendido, queda desenmascarado (como ilusorio). La parte II constará de un primer capítulo en el que me ocupo del Caso Global contra la Suerte Moral, y otros cuatro capítulos dedicados a cada uno de los tipos de suerte moral distinguidos. Mi estrategia combinará una reducción al absurdo del Caso como solución global y una defensa positiva individualizada de cada uno de los tipos de suerte moral (constitutiva, formativa, circunstancial y resultante).

Apéndice a la Parte I

La suerte moral y el debate sobre el libre albedrío y la responsabilidad moral

Una cuestión importante que quiero intentar responder antes de seguir adelante en el tema de esta investigación es la de ¿cuál es la relación o conexión entre el tema de la suerte moral y el debate en torno al libre albedrío? Dedicaré este apéndice a intentar clarificar la relación entre el tema que nos ocupa y este problema clásico.

Podría decirse que tanto por la sociología de la disciplina (por aquellos que se dedican a uno y otro tema, por la estrecha vinculación en la bibliografía, etc.) como por las nociones que son discutidas en ambos debates (interés por la responsabilidad moral, las atribuciones correctas de ésta, el control del agente sobre sus propias acciones y carácter, etc.) parece que su vinculación es evidente. No obstante, caracterizar con exactitud esta relación es bastante complicado. En concreto, la cuestión sería la de si estamos ante dos temas completamente diversos —que parece que no; si se trata del mismo tema con diferentes nombres —que parece que tampoco; o si estamos ante temas que sólo parcialmente se solapan — opción que parece la más prometedora para caracterizar correctamente la relación real entre ambos temas.

En principio, como vimos al tratar de los diferentes tipos de suerte moral, parece que es la suerte causal —la suerte en cómo las circunstancias antecedentes nos determinan— la que está más estrechamente relacionada con el tema de la libertad y la necesidad. De hecho, Nagel y otros la han asimilado explícitamente a este debate. Y es tan así que es el tipo

menos tratado en la bibliografía específica de la suerte moral, precisamente porque ya es el centro en el debate sobre el libre albedrío. En este sentido, podríamos discriminar ambos debates diciendo que si bien la cuestión de si somos últimamente libres, o poseemos libre albedrío, es el foco de un debate; en el área de la suerte moral éste punto no constituye el foco inmediato de interés. Sin embargo, la cuestión es más compleja. Pues, por ejemplo, Andrew Latus ha afirmado que la categoría de la suerte causal es incluso redundante dentro del propio debate de la suerte moral, pues su área queda completamente cubierta por la combinación de la suerte circunstancial y la suerte constitutiva.¹ Si esto es así, parece que ambos tipos estarán asimismo estrechamente relacionados con el problema del libre albedrío o la libertad última. Obviamente, la cuestión del poder de la persona en la configuración de su propia constitución es fundamental para saber si ésta es o no libre, así como las circunstancias en las que se ve forzada a actuar. No obstante, y aun siendo esto último cierto, creo que se puede resistir esta nueva identificación de ambos debates.

Particularmente, me siento tentado a afirmar que, en la discusión sobre la suerte moral, el propósito es elucidar el concepto de responsabilidad moral a partir del análisis de su uso en las atribuciones que corrientemente realizamos, asumiendo que no es un concepto completamente erróneo. En otras palabras, se trata de dilucidar problemas internos a nuestras prácticas e instituciones de responsabilidad moral, sin necesidad de entrar en la cuestión metafísica, definitoria del problema clásico del libre albedrío y la posibilidad de la responsabilidad moral. En particular, la discusión parte de la reflexión sobre prácticas particulares de juicio moral, que al ser indagadas chocan con los mismos principios que parecen regirlas, lo cual nos sitúa ante una perplejidad, que habrá que des-

¹ Latus (2001).

hacer. En este sentido, no es fundamental para nuestro debate llegar a discutir sobre el fundamento metafísico último de estas prácticas —aunque, sin duda, la discusión va progresivamente acercándose a este punto. El contraste puede enunciarse así: el tema de la suerte moral es una investigación en *psicología moral*,² mientras que el problema clásico del libre albedrío y la responsabilidad moral es una cuestión de cariz metafísico.³ Aunque esto no impide que haya obvias e importantes conexiones entre ambos tipos de investigación, como veremos.

Por otro lado, el tema es usualmente motivado a través de los ejemplos que aparecen como más desconcertantes, que son los de casos de suerte resultante; un tipo de suerte moral completamente independiente del debate sobre el libre albedrío, que tiene mucho más en común con el debate acerca de dónde situar el *locus* del juicio moral, esto es, la discusión de si lo que primariamente hay que juzgar es la intención o voluntad del agente, su virtuosidad o, más bien, las consecuencias de sus acciones (me refiero a las conocidas disputas entre deontologistas, consecuencialistas y defensores de la ética de las virtudes; esto es, los representantes de las tres principales teorías éticas normativas).⁴ De los casos de suerte resultante se pasa normalmente a discutir ejemplos de suerte circunstancial y de ahí a casos de suerte antecedente y constitutiva. Sólo progresivamente uno va acercándose a la cuestiones de libertad o responsabilidad última.

² Por “psicología moral” [*moral psychology*] se entiende aquí un área de estudio filosófico que se encuentra en la intersección entre la filosofía moral, la filosofía de la mente y la filosofía de la acción. Por supuesto, el estudio del libre albedrío y la responsabilidad moral también hace uso de materiales procedentes de estas áreas. Si se quiere, la diferencia es más bien de énfasis.

³ Agradezco a Carlos Moya la sugerencia de categorizar la distinción en términos del tipo de empresa filosófica que guía cada debate: metafísica o de psicología moral.

⁴ Si uno se ocupa sólo o principalmente en este tipo de suerte moral, no es extraño que afirme, como es el caso de Richards (1986, p. 172), que estamos ante cuestiones completamente diversas.

Sin embargo, la discusión puede ser muy candente aún sin que se refiera en absoluto a estas cuestiones.

Como avancé en el capítulo primero, el problema de la suerte moral discute y desafía cualquier noción o tipo de control, incluso el control menos exigente, como es el sentido cotidiano de control. Y esta es una lección importante que podemos aprender del tema de la suerte moral: que hay multitud de aspectos controvertidos acerca de la responsabilidad moral y sus atribuciones, más allá del tema metafísico de la responsabilidad última. Además, la suerte moral presentaría un reto *gradual* a las atribuciones de responsabilidad moral, mientras que el debate tradicional y contemporáneo acerca del libre albedrío se desarrolla casi exclusivamente en términos de condiciones necesarias y suficientes.

Si esto es así, parece que no podemos pedirle a ninguna teoría sobre la libertad y la responsabilidad moral que para ser aceptable resuelva, a su vez, la cuestión de la suerte moral. En otras palabras, parece que tanto las posiciones compatibilistas como incompatibilistas no comportan, en sí mismas, la negación de todo tipo de suerte moral.⁵ Además, uno puede ser tanto compatibilista como incompatibilista y defender o rechazar la existencia de la suerte moral resultante. Aunque esto debe ser matizado.

Por un lado, la meta del compatibilismo es tratar de distinguir entre diferentes tipos de factores sobre los que uno no tiene control. Si las acciones de una persona están causadas por factores que ésta no controla

⁵ El compatibilismo es la tesis de que la existencia del libre albedrío es compatible con el determinismo; y dado que el libre albedrío se suele tomar como condición necesaria de la responsabilidad moral, el compatibilismo acerca de la responsabilidad moral afirma que ésta y el determinismo son compatibles. (El determinismo es la tesis de que *todo* lo que sucede está determinado por las condiciones antecedentes más las leyes de la naturaleza.) Por otro lado, el incompatibilismo es la postura que defiende que si el determinismo es verdadero, entonces no poseemos libre albedrío, y no somos verdaderamente responsables.

y que le impiden tener o ejercitar ciertas capacidades, entonces no es responsable. Sin embargo, si las acciones de una persona son causadas por factores que ésta no controla, pero que le permiten tener o ejercer las capacidades relevantes, entonces puede tener control sobre sus acciones en el sentido relevante, y así ser responsable de sus propias acciones.

Por otro lado, también la mayor parte de los libertaristas, en tanto que libertaristas, pueden no poner objeción, al menos en principio, a la existencia de ciertos tipos de suerte moral.⁶ De hecho, hay un tipo de suerte (que tradicionalmente se ha achacado a los libertaristas como una objeción acerca del control) que es la que hace que una persona escoja actuar o ser de una manera y no de otra, sin que esta elección venga determinada por las causas antecedentes. Para que una acción sea libre, en esta concepción, no se trata sólo de que no esté determinada (que sea fruto, digamos, de leyes probabilísticas), sino que además ha de surgir de los esfuerzos de la voluntad e intención del agente. Para los defensores del llamado *argumento de Mind*, esta indeterminación sólo añade un factor de aleatoriedad que en realidad viene a reducir el control racional del agente sobre su acción. (Aunque en esta discusión el término suerte tiene un significado mucho más específico que el que tiene en la expresión “suerte moral”.)⁷

No obstante, parece natural pensar que, en principio, la existencia de la suerte moral beneficiaría más (o sería menos adversa) a la causa compatibilista. Y algún destacado compatibilista, como es el caso de Da-

⁶ El libertarismo es la forma de incompatibilismo que afirma que, de hecho, existe el libre albedrío —y así la responsabilidad moral— y, por lo tanto, el determinismo es falso.

⁷ Al hablar de libertaristas en este y en el siguiente párrafo, tengo en mente principalmente a Robert Kane (en especial, Kane 1999) y otros libertaristas de su clase; véase, por ejemplo, Moya (2006). Para los principales representantes del compatibilismo, véase Dennett (1984), Fischer y Ravizza (1998) y Wolf (1990).

niel Dennett, ha reconocido con entusiasmo la existencia de la suerte moral.⁸ También lo ha hecho John M. Fischer, aunque rechazando la suerte resultante.⁹ No obstante, en general, los compatibilistas suelen guardar silencio en relación a la suerte moral resultante y circunstancial. Sin embargo, como afirma Dana Nelkin la suerte moral podría suponer para ellos un recurso desatendido, pues si resulta que la suerte —la falta de control— producida por el determinismo no es más que una fuente de suerte entre otras, resultará que el determinismo no es el único obstáculo para el libre albedrío y la responsabilidad moral, por lo menos en relación al control.¹⁰ Pero, por otro lado, también podría argüirse que la existencia de la suerte moral relativiza la objeción clásica al libertarismo a que hacía referencia antes, pues la suerte en la elección final no sería más que un factor más de suerte entre los diferentes tipos de suerte que intervienen en el origen de nuestras acciones. En este sentido, el libertarista podría replicar, al argumento de Mind, que un mundo determinista tampoco estaría libre de suerte.¹¹

Quizá, sólo el libertarismo del agente-como-causa pueda ser considerado, en principio, y en sí mismo, una teoría con aspiraciones a eliminar toda clase de suerte moral —si bien no se ha ocupado directamente del problema de la suerte moral. La idea básica es que los agentes mismos causan sus propias acciones o al menos la formación de sus intenciones, sin ser a su vez causados a hacerlo. Así, el agente mismo, en tanto que sustancia, ejerciendo su poder causal es causa indeterminada de sus inten-

⁸ Véase Dennett (1984), cap. 3.

⁹ Tanto Fischer (2006) como, por ejemplo, Nelkin (comunicación personal) defienden una postura híbrida, que acepta unos tipos de suerte moral pero rechaza otros (en particular, la suerte resultante).

¹⁰ Nelkin (2004).

¹¹ Sobre la existencia de la suerte en un mundo determinista, véase más abajo. En Rosell (2007) he tratado de rechazar el argumento de que la existencia de la suerte moral podría favorecer la causa del *libertarismo del origen*.

ciones. La diferencia central entre la causación entre sucesos, o causación convencional, y la causación agencial residiría en los *relata*. En la primera ambos *relata* son sucesos, mientras que en la última el primer *relatum* es un agente y el segundo un suceso. Además, la acción sería libre, en el sentido relevante para la responsabilidad moral, sólo si no se da el determinismo.¹² Lo que parece se quiere evitar es que nuestras acciones dependan en alguna medida de factores que están más allá de nuestro control. Sin embargo, este tipo de teorías ha recibido importantes objeciones en tanto que teorías sobre el libre albedrío, y además parecen ser consistentes con que las acciones e incluso intenciones del agente dependan en parte de factores que están más allá de su control (tal como las razones disponibles en el momento de la decisión o acción).¹³

En cualquier caso, sería pedir demasiado a una teoría libertarista que, además, de resolver el problema del libre albedrío tuviese que negar la posibilidad de la suerte moral, en especial en sus tipos circunstancial y resultante. Se puede decir que, aunque el libertarismo que defiende la causación en el agente fuera verdadero, aun así, nuestra discusión permanecería indecida. Es decir, quizá estuviese decidida para la suerte causal y constitutiva, pero no para la suerte circunstancial y resultante. Y, así, incluso demostrándose existencia del libre albedrío humano, el problema de la suerte moral seguiría existiendo. No obstante, sí que hay al menos un caso de libertarismo que es a su vez un intento de eliminación de todos los de casos de suerte moral. Se trata de Michael Zimmerman, que trata de establecer una teoría completamente contrafáctica de la adscripción de

¹² Sus principales representantes fueron Thomas Reid, Roderick Chisholm y Charles Taylor. Para más recientes y sofisticadas articulaciones, véase O'Connor (2000) y Clarke (1993). Véase también Fischer (1999) pp. 106-108, para una interesante exposición y comentario de estas posiciones.

¹³ Cf. Nelkin (2004).

responsabilidad moral, que la haga por principio completamente inmune a la suerte.¹⁴ (Veremos en qué consiste y dónde falla esta teoría, en el capítulo siguiente.)

En el otro extremo, si resultase que la responsabilidad moral es imposible o ilusoria —como han defendido de manera destacada Galen Strawson, Derk Pereboom y Saul Smilansky— parece que la cuestión de la suerte moral perdería su sentido, o gran parte de su interés, dado que ello (probablemente) acabaría con la moralidad misma.¹⁵ No obstante, la cuestión podría quizás plantearse nuevamente al nivel de los sustitutos pragmáticos que viniesen a sustituir la noción substantiva de responsabilidad moral arruinada. Inversamente, si resultara que la investigación en psicología moral muestra que nuestro concepto de responsabilidad moral es incoherente, esto tendría, sin duda, repercusiones para la investigación metafísica.

En todo caso, cabe distinguir este escenario del caso en el que meramente resultara que el mundo es determinista, pues en un mundo determinista seguirían existiendo los casos de suerte en general, y quizá de suerte moral, en particular. De hecho, en entornos deterministas cerrados (cuya existencia no tiene por qué ser negada por nadie) también se dan casos de suerte. Imagínese un caso en el que una persona *P* compra un número de lotería cuando el sorteo que se ha realizado y hay ya un ganador pero nadie sabe de quién se trata. Que *P* resulte la agraciada o no será igualmente cosa de suerte *para P*, a pesar de que esté ya determinado si *P* es o no la agraciada. La suerte, en nuestro debate, no significa más que

¹⁴ Zimmerman (2002).

¹⁵ Véase Pereboom (2001), Strawson (1994) y Smilansky (1999). Aunque, por ejemplo, Honderich (1988) ha defendido que la moralidad subsistiría aún siendo ciertos los postulados de los *incompatibilistas duros*.

falta de control, y el control puede faltar tanto en un mundo determinista como en uno indeterminista.¹⁶

Una conclusión a la que cabe llegar es que el problema de la suerte moral es independiente, por lo menos parcialmente, del problema del libre albedrío, y resolver o dar respuesta a uno no conlleva resolver o dar respuesta al otro; bien sea porque uno y otro no coinciden, al menos completamente, en su objeto, bien sea porque son cuestiones de diferente orden que deben ser estudiadas de diferente manera (desde la psicología moral o desde la metafísica). En todo caso, la cuestión de la suerte moral permanecerá indecisa por los resultados del problema del libre albedrío y la responsabilidad moral ligada a la existencia de éste.¹⁷

En lo que sigue, y siendo fiel a estas conclusiones, trataré de evitar el problema metafísico del escepticismo sobre la responsabilidad moral (si bien diré algo más sobre él en el capítulo 8).

¹⁶ Vemos, pues, cómo la noción de suerte presente en la cuestión de la suerte moral tiene que ver con cualquier tipo de falta de control, mientras que su papel en el *argumento de Mind* es más particular y depende de la existencia de un mundo indeterminista. En el capítulo 4 discutiré la misma noción de suerte relevante para nuestro tema.

¹⁷ Con la posible excepción, de nuevo, de la impugnación de la misma noción de responsabilidad moral.

Parte II

3. El caso contra la suerte moral

Articulación y respuesta

- 3.1. Reconstrucción del Caso Global contra la Suerte Moral
 - 3.1.1. Fundamentos: el Principio de Control y el Argumento Epistémico
 - 3.1.2. Escenarios contrafácticos
 - 3.1.3. La Estrategia Moderada y la Estrategia Radical
- 3.2. Réplica al Caso Global contra la Suerte Moral
 - 3.2.1. Tipos de evaluación moral y la insuficiencia de EM
 - 3.2.2. ER: historias posibles y merecimiento último. Dudas escépticas y revisionismo
 - 3.2.3. ER: identidad e incoherencia
- 3.3. Recapitulación
 - 3.3.1. La estructura del argumento
 - 3.3.2. Contraste con el argumento del *todo-o-nada*
 - 3.3.3. La distinta significación moral de los diferentes tipos de suerte moral

...si el sol no hubiera estado ardiendo sobre Meursault en *L'Étranger* de Camus o si su estado de ánimo hubiera sido otro, o si el árabe que estaba ese día en la playa se hubiese quedado en casa, Meursault no habría matado al árabe.

Lynne R. Baker

Si hubiéramos subido al coche, nuestra vida habría sido muy diferente.

Dennis Lehane, *Mystic River*, II-13.

En esta segunda parte emprendo la exposición y rechazo del caso contra la suerte moral. Los objetivos son, pues, dos: por un lado, reconstruir correctamente el caso; y, por el otro, tratar de mostrar dónde y por qué falla. Para ello, empezaré ocupándome de él en tanto que *caso general contra la suerte moral en todos sus tipos*. En su reconstrucción, distinguiré dos estrategias principales, que pasaré a atacar individualmente —una de ellas

resultará insuficiente y la otra completamente impráctica, y finalmente incoherente. En todo caso, mi réplica no será completa hasta que consiga responder individualmente a la negación de la suerte moral en cada uno de sus tipos principales, cosa que trataré de hacer en los cuatro capítulos restantes de esta parte (4 al 7).

3.1. Reconstrucción del caso contra la suerte moral

En este capítulo, empiezo reconstruyendo lo que llamaré el Caso Global contra la Suerte Moral, a partir de los diferentes argumentos presentados en contra de la existencia de dicho fenómeno. De hecho, creo que podemos hablar de un argumento fundamental, desplegado de forma más o menos independiente por diversos filósofos, a partir del cual caben diferentes variaciones.

Defenderé que es imposible que un argumento de este tipo pueda finalmente prevalecer. Mi razón es que éste habría de depender de una noción tan dudosa como la de *merecimiento verdadero*, en un sentido especialmente fuerte, que muchos atribuyen a Kant, de *función estricta del (o estrictamente proporcional al) control del agente*; noción que llevada a sus últimas consecuencias resultará ser incoherente, además de totalmente impráctica.

3.1.1. Fundamentos: el Principio de Control y el Argumento Epistémico

Como avancé, los negadores de la suerte moral afirman que el supuesto fenómeno de la suerte moral, tras la debida reflexión, resulta ser una ilusión. Un arma fundamental para alcanzar esta conclusión es el lla-

mado *Argumento Epistémico*.¹ Según éste, los supuestos casos de suerte moral no muestran que la suerte pueda realmente afectar al juicio moral que el agente merece, a su estatus moral, sino que sólo lo hacen con respecto a nuestro conocimiento de lo que éste merece, pues no en balde no somos seres omniscientes y nuestro conocimiento está limitado por la evidencia a que nos es accesible.

Partiendo del compromiso con el Principio de Control, el argumento trata de explicar el funcionamiento de los supuestos casos de suerte moral sobre la base del razonamiento anterior, que podemos reformular del siguiente modo:

(AE) Lo que la suerte realmente hace no es interferir en el estatus moral de alguien, sino en el conocimiento que tenemos de éste, dado que no somos seres omniscientes y nuestro conocimiento depende de la evidencia disponible.²

Veamos con más detalle cómo funciona este argumento considerando un supuesto caso particular de suerte moral en las consecuencias, como es el de los conductores temerarios. Aquí uno de los conductores acaba atropellando a una persona y el otro no. Llamémosles, respectivamente, José y Josué. Parece que cotidianamente tendemos a juzgar más duramente —lo que, *prima facie*, indicaría que lo consideramos más culpable— a José, el

¹ El nombre se debe a Latus (2000). Para este tipo de argumento, véase: Richards (1986), Thomson (1989), Rescher (1990), Jensen (1993), Rosebury (1995), Latus (2000) y Enoch y Marmor (2007); también Zimmerman (1987) y (2002), aunque su posición es singular, como veremos.

² Esta tesis sintetiza lo que tienen en común los siguientes pasajes: “El culpable puede así ser afortunado o desafortunado en la claridad de su merecimiento.” (Richards 1986, p. 169.) “La suerte implicada no tiene que ver con nuestra condición moral sino sólo con nuestra imagen; está conectada no a lo que somos sino a cómo la gente (incluso nosotros mismos) nos ve.” (Rescher 1990, pp. 154-5.) “Si el daño real se da, el agente y otros que consideran su acto tendrán un dolorosa consciencia de este daño.” (Jensen 1993, p. 136.) “El daño real tiene sólo la función de hacer vívido cómo de mala fue la acción debido al peligro creado.” (Bennett 1995, pp. 59-60.)

conductor que atropella y causa un daño a otras personas, que a Josué, el conductor que no atropella a nadie. Esto supondría que la suerte marca una distinción moralmente relevante y, por lo tanto, la presencia de la suerte moral. De hecho, en el primer caso tenemos un atropello, y con ello una o más personas que sufren un daño importante, mientras que en el segundo caso no hay atropello y nadie sale mal parado. Sin embargo, el negador de la suerte moral puede reconocer que, ciertamente, la suerte interviene en el caso, pues uno de ellos atropella a alguien y el otro no. José causa un daño a otros y Josué, no. Ciertamente, hay un sentido en el que la suerte interviene en nuestro juicio moral. Hay un sentido por el que José puede ser menos afortunado que Josué en relación al juicio que parece merecer. Pero ello no implica que la suerte marque una diferencia moral verdadera, esto es, que afecte a *lo que la persona verdaderamente o últimamente merece*. En realidad, continúa el argumento, ambos conductores son igualmente culpables, si la temeridad de ambos es la misma, con independencia de las consecuencias de sus acciones.

En condiciones normales sólo sabemos que Josué no ha causado ningún daño, lo cual no nos hace sospechar de su falta de rectitud al volante. Seguramente, Josué suele conducir atendiendo a la carretera, yendo a una velocidad segura, etc. O quizá no. Pero no hay nada que nos haga desconfiar o sospechar, o incluso preguntarnos, si es o no así. Sin embargo, José ha atropellado a alguien, y esto nos hace pensar que posiblemente su atención no era la adecuada, o que su velocidad era excesiva, etc.; en definitiva, que fue temerario, y que no sopesó correctamente el riesgo que su conducción temeraria comportaba a otros usuarios de aquella vía pública. Si esto es así, el negador de la suerte moral puede conceder sin excesivos problemas que la suerte puede influir incluso en el juicio moral acerca de los demás que ordinariamente emitimos, pero ello no repercute

en su verdadero estatus moral. Esta es, en especial, la posición de Norvin Richards, quien defiende explícitamente que cuando no existe un daño real la cuestión permanece incierta, y que es precisamente cuando la suerte abandona a alguien cuando tenemos la base suficiente como para tratarlo de la manera en la que hubiese merecido ser tratado desde el principio.³ En todo caso, lo que cabe retener es que, según los defensores de AE, una persona podrá ser afortunada o desafortunada en la *claridad* de lo que merece, pero ello no significa que la suerte pueda marcar una distinción moral, es decir, afectar a *lo que merece*, a su merecimiento (verdadero).⁴

Con esto, podríamos ya desenmascarar el error de las caracterizaciones típicas de los supuestos casos de suerte moral, y mantener, a la vez, la intuición o Principio de Control y nuestras prácticas de juicio ordinarias. Aunque ello depende del compromiso adicional —que va más allá del núcleo estricto de AE— con una distinción fundamental entre *los juicios que estamos justificados a proferir*, dada la limitación de nuestras facultades cognitivas y la evidencia accesible, y aquello que la persona

³ Richards (1986), p. 171. Esta última idea me parece en sí misma discutible, pero no es mi propósito discutirla aquí. La meta de este capítulo es más bien evaluar si AE puede extenderse adecuadamente a todas las formas de suerte moral o no. Para su discusión con respecto a tipos particulares de suerte moral habrá que esperar a los próximos capítulos de esta parte II.

⁴ Utilizo la palabra ‘merecimiento’ como traducción, más o menos estipulativa, de *desert*, que en la bibliografía especializada (en inglés) se distingue del término *merit*, para el cual es más fácil reservar ‘mérito’. La diferencia principal vendría a ser que *desert* se usa normalmente para hablar sobre las *acciones* de alguien, y *merit* hace más bien referencia a sus capacidades y habilidades, en el sentido de *hechos* sobre alguien. En el primer caso se trata de algo por lo que el sujeto es (más genuinamente) responsable, mientras que no lo es (tan claramente) por lo segundo. Merecimiento y mérito evocarían respuestas diferentes: te admiro por tus méritos, te elogio en la medida de tu merecimiento. Cf. Lucas (1993), pp. 124-7. Sin embargo, esta distinción, que es sólo ambiguamente reconocida en el discurso ordinario, admite diferentes grados de rigidez; lo que la hace inestable a nivel teórico. De hecho, la misma falta de nitidez de esta distinción será parte de la discusión.

realmente merece, su *merecimiento real*.⁵ Llamaré a este compromiso adicional la *Tesis Conservadora* (por antirrevisionista).

Así pues, nuestro trato diferencial de José y Josué dependería, no del hecho de que realmente merezcan un trato diferente; sino de que, en primer lugar, su conducta no nos muestra con la misma claridad que lo que merecen sea lo mismo y, en segundo lugar, nuestro trato de ellos debe reflejar nuestra comprensión (epistémicamente limitada) de lo que merecen, y por lo tanto el modo en que tenemos el deber de tratarlos.⁶ En otras palabras, por un lado, no siempre está claro cuáles son las intenciones de una persona, su compromiso con un curso de acción, su verdadera voluntad, etc. Y, por otro, en tanto que humanos, no tenemos acceso a una perspectiva omnisciente que nos permita conocer exactamente el merecimiento de uno y otro. Así, el resultado de sus respectivas acciones parece ser *el mejor indicador* que tenemos para saber cómo hemos de tratarlos.

Sin embargo, también puede apelarse a una estrategia ligeramente distinta, que en vez de incidir en la falibilidad del juicio humano, lo hace en los diferentes propósitos que cotidianamente guían nuestras prácticas de evaluación moral. A saber, cotidianamente censuramos a aquellos que causan sucesos negativos con la pretensión de cambiar su conducta, con independencia de si realmente merecen ser censurados. Sería esto cuestión de la censura explícita de una persona; cosa distinta de su culpabilidad real. En otras palabras, una cosa serían los *actos de censura*, consistentes en someter a alguien a censura explícita, típica de los reproches, reprimendas, etc.; y otra los *juicios de culpabilidad moral*, en los que la intención primaria del hablante es dar un veredicto impersonal aplicable a

⁵ En particular, Richards (1986), Rosebury (1995) y Thomson (1989) han defendido esta posición.

⁶ Véase Richards (1986), pp. 171-2.

todos aquellos cuyas acciones son iguales en los aspectos relevantes, con el propósito de juzgarlo como merecedor de censura moral y no de meramente influir en su conducta. Ambos constituirían además tipos distintos de aseveración.⁷

Si, de nuevo, aplicamos esta distinción al caso de los conductores temerarios, cabría decir que sólo el conductor que atropella a un peatón es abiertamente censurable —no así el conductor que no atropella a nadie—, aunque ambos sean igualmente culpables. José es censurable por producir un daño moral, dado que causó el atropello (sin tener una excusa adecuada). No así Josué, que no atropelló a nadie. Sin embargo, esto no marca una distinción en su *calidad como persona*; ambos son igualmente buenas o malas personas (y *deberían* producirnos la misma indignación moral⁸). Con ello, la suerte (en las consecuencias, en este caso) sólo puede marcar una distinción moral en la censura abierta, pero no en la culpabilidad y calidad como persona de los agentes, que es precisamente la esfera evaluativa fundamental que permanece inmune a la suerte.⁹

Alternativamente, otro movimiento, mucho más vasto —pero que muestra de modo mucho más intuitivo la meta de la estrategia que esta-

⁷ Véase Thomson (1989) pp. 200-3, y Jensen (1985) pp. 132-4, para defensas de esta posición. En la misma línea, Zimmerman (1987) pp. 218-9, distingue entre elogio y censura *activos e inactivos*. Cabe remarcar que para algunos la noción de censura aparece como más urgente que la de elogio, lo que establecería una plausible *asimetría* entre ambas. Parece que la censura es un fenómeno más complejo que el elogio, que merece una atención más detenida; por lo menos, en tanto que la censura impropia parece producir mayor injusticia que el elogio impropio. Véase Wolf (1990) para una defensa destacada de la asimetría entre elogio y censura, aunque por razones diferentes.

⁸ Véase Thomson (1989).

⁹ Una distinción del mismo tipo es la de Zimmerman entre *alcance* y *grado* de la responsabilidad. Según ésta, José y Josué deben ser evaluados moralmente exactamente del mismo modo; aunque uno sea responsable *de más cosas* que el otro, ambos son responsables *en el mismo grado*. “El grado de la responsabilidad lo es todo; el alcance no cuenta para nada en cuanto a la evaluación moral de los agentes”. Zimmerman (2002) p. 568; véase también Zimmerman (1987). Discutiré con más detenimiento esta distinción en los capítulos 6 y 7.

mos considerando—, es denunciar que en los casos ejemplares de suerte moral (como el del asesino y el asesino frustrado, o los conductores negligentes o temerarios) a menudo se confunde suerte *moral* con suerte *legal*. La idea es que si bien podemos tener buenas razones a nivel legal para tratar a uno y a otro diferentemente —esto es, castigar más severamente a José que a Josué—, es un error inferir de aquí que la ley refleje directamente la evaluación moral correcta de tales casos. Ciertamente, el derecho constituye un sistema que gobierna la atribución de la responsabilidad y la respuesta punitiva, cuyo principal objetivo es el bienestar social y, en particular, impedir la comisión de actos antisociales. Para ello establece un conjunto de convenciones, incluyendo la pena legal, que priva al criminal y trata de disuadir. De este modo, los juicios legales no pueden ser regulados por una concepción de la responsabilidad moral independiente de la suerte, dado su intrínseco carácter convencional.¹⁰ No obstante, de esto sólo se sigue que cualquier caso de suerte legal no constituye, *eo ipso*, un caso de suerte moral. Pero no que por ser casos de suerte legal ya no puedan serlo de suerte moral. Aparte de las razones legales que justifican castigar más severamente el crimen efectivo que el mero intento, pueden también existir razones morales. De hecho, no son pocos los juristas y teóricos del derecho que consideran que el derecho ha de mantener una estrecha correspondencia con la moralidad; ni es inusual que se produzca un fuerte rechazo popular cuando los tribunales dictan sentencias que claramente chocan con la moralidad. Si, por lo menos, la legalidad rastrea la moralidad, habrá casos de suerte legal y moral (si los

¹⁰ Para una desarrollada defensa de este argumento, véase Rosebury (1995), pp. 521-4. Ya Nagel notaba que una objeción probable a su caso del asesino y el asesino frustrado sería que hacer más responsable a uno que a otro “se asemeja a la exigibilidad estricta (*strict liability*), que puede tener sus usos legales pero parece irracional como posición moral” (1979, p. 31).

hay) a la vez. Así, que un caso determinado cuente como caso de suerte legal no impide que pueda ser, a su vez, un caso de suerte moral, aunque tampoco implica que lo sea.¹¹

Por otro lado, hay que distinguir entre el propio AE y la *Tesis Conservadora*. El núcleo de AE es esencialmente negativo, dice lo que no es; mientras que la Tesis Conservadora (que, recordemos, justifica el trato diferente de uno y otro agente en virtud de las claras diferencias en las consecuencias de la conducta —así como en la diferente conducta, como veremos—, aunque sólo sea en relación a la censura abierta, o al mero juicio legal) constituye un determinado desarrollo positivo, entre los posibles, que va más allá del núcleo de AE. El defensor de AE no tiene, estrictamente, por qué comprometerse con ninguna explicación positiva de nuestras prácticas, aunque ofrecer una buena explicación favorece enormemente su posición. De hecho, las mismas explicaciones positivas pueden chocar entre sí —lo que permite diferentes interpretaciones del fenómeno incluso entre aquellos que lo rechazan sobre la base de AE.

Dejando a un lado, de momento, la Tesis Conservadora, tenemos que AE es el fundamento de la respuesta de los adversarios de la suerte moral a la existencia de este supuesto fenómeno. Hasta ahora, sólo hemos considerado su respuesta a los casos de suerte resultante. Pero para conseguir un Caso Global contra la Suerte Moral es necesario que éste pueda ser convenientemente aplicado a todos los tipos de suerte moral. Sin embargo, no está claro de qué modo este tipo de explicación —que tiene su

¹¹ La misma posibilidad de que la suerte (principalmente *resultante*) deba marcar una diferencia en la atribución de responsabilidad legal —especialmente en derecho penal, pero también en derecho civil— es objeto de controversia. La bibliografía al respecto es rica y extensa; véase, en particular, Hart (1968), Davis (1986), Fischer y Ennis (1986), Sverdlik (1988), Lewis (1989), Kessler (1994), Moore (1994), Feinberg (1995) o Levy (2005). Para más referencias, consúltese la bibliografía en éste último.

banco de pruebas principal en la suerte en las consecuencias— puede extenderse sin más a la suerte circunstancial y antecedente. La solución quizá venga de la mano de la sofisticación de las distinciones anteriores, aunque también puede suceder que su adecuada extensión resulte difícil o imposible.

No niego que las distinciones anteriores, y otras que se propondrán, no hagan parcialmente justicia a nuestras prácticas cotidianas. Además, algunas aprovechan oposiciones clásicas, como el elogio y la censura fruto del merecimiento o como medios para conseguir consecuencias deseadas—influendo en la conducta del agente, para reeducarlo, disuadirlo, etc. Sin embargo, como veremos, se requiere algo más para establecer el rechazo de la suerte moral.

3.1.2. Escenarios contrafácticos

La idea que permite generalizar la posición anterior a todos los tipos de suerte moral viene de la mano de la formulación contrafáctica del Principio de Control: el merecimiento moral de una persona —lo que ésta verdaderamente merece— no es el resultado de lo que esta persona haya hecho, sino de *lo que habría hecho si hubiera tenido la ocasión*.¹²

Considérese el caso más clásico de suerte circunstancial: el de los dos simpatizantes nazis. Rudolf y Adolf son adeptos, en la misma medida, a la política que el nazismo está empezando a desplegar en Alemania y Austria, pero Adolf ha tenido que emigrar a Argentina por cuestiones de negocios y no puede participar activamente en ella. Rudolf, por el contrario, permanece en Alemania y tiene la opción de enrolarse en las SS y

¹² Posiblemente, el libertarista sospechará que esta manera de poner las cosas presupone el determinismo. Para no prejuzgar esta cuestión, la expresión anterior, así como otras del mismo tipo que surjan más adelante, deberían modificarse con el adverbio ‘probablemente’: lo que *probablemente* habría hecho si hubiera tenido la ocasión.

acabar comandando un campo de exterminio. Según el defensor de la suerte moral circunstancial, las diferentes circunstancias de ambos les abren y cierran diferentes oportunidades de acción que pueden acabar marcando una diferencia moral en sus respectivos merecimientos. Parece claro que el juicio que Rudolf nos merece (miembro de las SS y comandante de un campo de exterminio) será mucho más duro que el juicio de Adolf (mero simpatizante en la lejanía). Sin embargo, el adversario de la suerte moral afirmará que lo realmente relevante no es cómo cada uno de ellos finalmente actúa, sino que hay algo en Adolf en virtud de lo cual, si hubiese estado en las mismas circunstancias que Rudolf, habría actuado libremente del mismo modo, esto es, se habría integrado en el nazismo en la misma medida y habría cometido los mismos horrendos crímenes. Siendo esto así, ambos serán igualmente responsables.

Si bien, por otro lado, la Tesis Conservadora nos permite tratar más duramente a Rudolf que a Adolf. Dados los hechos, la conducta criminal real del primero constituye una base suficiente para juzgarlo y tratarlo desde el punto de vista legal como culpable; mientras que, con respecto al segundo, no poseemos la misma evidencia para considerarlo igualmente culpable.

Cabe señalar que esta estrategia renuncia ya a la acción misma del agente como el *locus* de la responsabilidad y la sitúa en un elemento previo, en particular en aquello que hace que el agente *habría actuado* de una determinada manera *si hubiese tenido la ocasión*. Para Richards no hay duda de que el locus de la evaluación lo constituyen las disposiciones estables o carácter del agente, con independencia de su realización (puntual) en casos particulares. Esta es su versión de AE, referida a la suerte circunstancial:

El punto central de mi argumento será que si el agente potencial es tan parecido al real [*actual*] como estamos imaginando, entonces habrá alguna cosa más en su conducta que demandará la misma respuesta. Si esto es así, su suerte para llevar a término un hecho *particular* no afectará al trato que merece. Merecerá la misma clase de estímulo o recondicionamiento, pero lo merecerá por una diferente realización de un carácter igualmente virtuoso o vicioso. Su suerte no afectará, así, a lo que merece, sino al momento en el que lo merece y, de nuevo, a la claridad con la que puede verse que lo merece.¹³

Hay un sentido en el que esto se sigue de casos cotidianos en los que juzgamos a alguien en virtud de que parece razonable pensar que si tuviese o hubiese tenido la oportunidad de hacer *x*, lo habría hecho. Piénsese en ese compañero de trabajo del que, por sus palabras y actitudes, sospechas que si se diese la circunstancia te dejaría vendido ante tu jefe. O ese empleado que te trata tan bien, pero que si dejase el trabajo probablemente ni te saludaría. Si estos casos cotidianos pueden generalizarse, el resultado sería el vindicado por Richards, de que vinculamos nuestro juicio de alguien no con lo que ha hecho, sino con aquello *que es plausible pensar que habría hecho si hubiese tenido la ocasión*.¹⁴

Sin embargo, hemos de distinguir de nuevo la mera extensión contrafáctica de AE, el núcleo de la propuesta, de ulteriores compromisos acerca del locus de la atribución de responsabilidad moral. Vemos que Richards se compromete con la idea de que el merecimiento verdadero de un sujeto es función de su *carácter*. Pero este compromiso podría ser un impedimento a la hora de extenderlo a los tipos antecedentes de suerte moral, en los que precisamente está en juego el propio control del agente sobre el carácter que de hecho tiene. Otro impedimento en la misma di-

¹³ Richards (1986), p. 174.

¹⁴ Richards (1986) pp. 173-4.

rección (que también constituye un desarrollo particular de la posición y no su núcleo) es la matización de “lo que es *plausible* pensar que hubiera hecho...”, esto es, el compromiso con el modificador “plausible”. Claramente, conforme retrocedamos más atrás en la consideración de los factores de suerte antecedente más nos separaremos de las condiciones reales y actuales, con lo que este “plausible” se volverá progresivamente menos razonable o inteligible.

En todo caso, el principio puede extenderse (dejando de lado el “plausible”) hasta abarcar la suerte antecedente, del siguiente modo: un agente es moralmente responsable en la medida en que éste habría realizado libremente la misma acción que su hipotético contraparte, si hubiese tenido la misma constitución y el mismo desarrollo y experiencias formativas —o la combinación de constitución más experiencia hubiese resultado globalmente en la misma clase de persona.

Pero podemos también prescindir de la idea de que el locus de la responsabilidad moral es el carácter o las disposiciones que (en algún sentido) el agente *de hecho* posee. La nueva estrategia a seguir queda claramente ilustrada en el siguiente esquema de Michael J. Zimmerman, que radicaliza la idea de que el merecimiento último reside en lo que el agente *habría hecho*:

Dada la Condición de Control [PC, en nuestro caso],
Si (i) P tomó la decisión *d* en la que creyó era la situación *s*,
(ii) P* hubiera tomado la decisión *d* si hubiera estado en una situación que creyese que era *s*, y
(iii) que el hecho de que P* estuviera en una situación que creía era *s*, no era algo que estuviera bajo su control [...],

Ergo: sea cual sea el crédito o descrédito moral que le cabe a P por tomar la decisión d le cabe también a P*.¹⁵

El resultado es que P y P* son igualmente elogiables o censurables. La cuenta moral de ambos, usando la imagen del mismo Zimmerman, se verá igualmente afectada, para bien o para mal, pues *hay algo en ambos que les hace igualmente responsables*.

Pero ese *algo* en virtud de lo que ambos son igualmente responsables no está claro ni tan siquiera qué pueda ser. Y, de hecho, la determinación de qué es ese algo va más allá del núcleo del argumento que estamos considerando.¹⁶ En todo caso, se tratará de *algo que el sujeto posee*, aunque sea sólo potencialmente. Y no puede ser ni su voluntad o intenciones actuales ni su carácter tal y como está actualmente formado; sino un mero contrafáctico acerca de, o bien, como podrían ser su voluntad e intenciones, o su carácter, o bien, meramente cómo decidiría en la situación s . Pero este retroceso en los antecedentes de la conducta y el carácter (o configuración actual) del agente en cuestión implicará también hacer nuevas distinciones, como trataré de mostrar, entre *tipos de merecimiento (verdadero)*.

Creo que, llegados a este punto, podemos empezar a sospechar que el núcleo del argumento que estamos considerando es meramente ya la aplicación consistente del Principio de Control. Esto es, el mismo AE se vuelve superfluo (se torna parte del desarrollo más que del núcleo de la

¹⁵ Zimmerman (1987), p. 381. He suprimido la referencia a “control restringido” (en contraposición a “control irrestricto”) por no ser relevante aquí; lo será más adelante (particularmente, en el capítulo 9). Para la distinción, véase Zimmerman (1987), p. 376.

¹⁶ Para Zimmerman el juicio de responsabilidad moral no puede ser función del carácter del agente, pues para ello ya están los *juicios aretaicos*. Estos últimos, que se refieren al *carácter*, están necesariamente influidos por la suerte. Son los primeros, función del *yo* del agente, los que, libres de suerte, establecen el merecimiento verdadero del agente. Esta posición parece ser la más próxima a Kant, para quien es la mera *buena voluntad* la que marca el merecimiento verdadero o valía moral.

estrategia de rechazo de la suerte moral.) De hecho, Zimmerman, des-
haciéndose finalmente de AE, concibe su propia posición como la “con-
clusión lógica” de la aplicación del Principio de Control hasta sus últimas
consecuencias.¹⁷

Por otro lado, podemos ya vislumbrar importantes diferencias de-
ntro del Caso Global, esto es, entre las posiciones particulares de aquellos
que comparten la meta de negar la existencia de la suerte moral. Y, sin
duda, estas divergencias tienen que ser convenientemente caracterizadas.
En la siguiente sección distinguiré las dos estrategias principales.

3.1.3. La Estrategia Moderada y la Estrategia Radical

Tenemos ya datos suficientes para proponer una distinción entre
dos grandes estrategias, bien diferenciadas, dentro del proyecto general de
negar la existencia de la suerte moral. Son las que llamaré Estrategia Mo-
derada y Estrategia Radical. Por supuesto, ambas estrategias parten de la
defensa del Principio de Control, pero lo hacen de modos distintos y ad-
quiriendo unos compromisos positivos también divergentes.

En primer lugar, la Estrategia Moderada (EM) se caracteriza por
partir de la defensa de AE, para hacer valer PC, junto con el compromiso
con tres ideas adicionales. En primer lugar, defiende la Tesis Conservado-
ra, cuya motivación es tratar de responder de un modo no revisionista a la
cuestión de cómo hemos de juzgar o qué hemos de pensar acerca de nues-
tros juicios cotidianos, dado PC y AE. Por otro lado, cuando extiende PC
y AE contrafácticamente, se compromete con la calificación de “plausi-
ble”, del siguiente modo: juzgamos por cómo es *plausible para nosotros*
*pensar*¹⁸ que el agente habría actuado si hubiera tenido la oportunidad.

¹⁷ Ver Zimmerman (2002), p. 559.

¹⁸ Richards (1986), pp. 173-4.

Finalmente, acepta que el *locus* del juicio es el carácter, disposiciones o intenciones que el agente de hecho tiene, origen de sus acciones.

En síntesis, esta posición se basa en la idea de que si bien el Principio de Control es irrenunciable, no por ello nuestros juicios cotidianos han de ser radicalmente revisados —para estos, lo mejor que tenemos es la evidencia disponible. (Dentro de EM se dan también variaciones: parece que en la versión de la distinción entre juicios legales o pragmáticos y juicios morales, o entre actos de censura y juicios de culpabilidad moral, sí que es factible un tipo de juicio referido a la culpabilidad o el merecimiento verdadero. Sin embargo, este sólo sería posible en relación a los casos de suerte resultante. Aplicada, a otros casos de suerte moral, esta opción se volvería revisionista.) En todo caso, la estrategia se aplica en principio sólo a la suerte resultante y a cierto tipo de suerte circunstancial, y calla en relación a los otros tipos de suerte moral.

Pero estos compromisos que estamos considerando no son necesarios para la defensa del Principio de Control, pues cabe la opción de perseguir todas “las implicaciones de la negación de la relevancia de la suerte para la responsabilidad moral” hasta su “conclusión lógica”.¹⁹ La Estrategia Radical (ER) es esencialmente *revisionista*; esto es, si el Principio de Control, irrenunciable, no es correctamente aplicado en nuestros juicios ordinarios —que parece que, de hecho, no lo es— éstos resultan incorrectos y deben revisarse. En particular, se trata de una *teoría del error*: no es que nuestros juicios morales ordinarios no se refieran al merecimiento último de los agentes juzgados (o sólo lo hagan imperfectamente, como defiende EM), sino que pretenden referirse a él, y de hecho se refieren, pero sistemáticamente fallan.

¹⁹ Zimmerman (2002), p. 556. Los principales valedores de esta estrategia son Zimmerman (1987) y (2002) y Greco (1995).

Además, ER está preparada para ir más allá de la mera aplicabilidad a la suerte resultante y circunstancial de EM, y puede extenderse a la suerte antecedente; pues puede aceptar que las intenciones, la voluntad, el carácter y las mismas disposiciones del agente dependen, por lo menos parcialmente, de factores que están más allá del control del agente. Para lo cual recurre a la idea de basar el merecimiento real en las disposiciones que el agente hubiese podido tener dadas *sus historias posibles contrafácticas*, esto es, las disposiciones que hubiese tenido si su historia fáctica hubiese sido otra, de entre el conjunto de historias posibles para su vida. (En la próxima sección explicaré con más detenimiento esta propuesta y sus dificultades.) Con esto se deshace de los otros dos compromisos que limitaban la extensión de EM: que el carácter o intenciones actuales del agente constituyan el locus para su evaluación moral última, así como la restricción de lo “plausible”.

En resumen, ER es una estrategia más ambiciosa y comprehensiva que EM, pero, como veremos, al precio de perder en intuitividad y aplicabilidad. Distinguidas estas dos estrategias, a continuación intentaré responder individualmente cada una de ellas.

3.2. Réplica al Caso contra la Suerte Moral

Mi estrategia general de respuesta al Caso Global contra la Suerte Moral, tal y como lo he caracterizado hasta el momento, constará de dos movimientos. En primer lugar, consideraré la Estrategia Moderada, respecto a la cual intentaré mostrar que es claramente insuficiente como caso general, pues falla irremediablemente en cuanto se intenta aplicar a la suerte antecedente (suerte formativa y constitutiva). Rechazada esta estrategia, pasaré a considerar la Estrategia Radical. Respecto a ella trataré de

mostrar que, si bien su extensión a tipos antecedentes de suerte moral no es lógicamente problemática, sí lo es llevarla hasta sus últimas consecuencias. ER resultará inaplicable y, lo que es más importante, incoherente. (Cabe decir que si ER sólo fuese inaplicable, esto supondría un escollo importante, pero no impediría que pudiese ser la solución idealmente correcta. Sin embargo, si es también incoherente, ya no tiene salvación). Si ni EM ni ER funcionan, y ambas agotan el Caso Global contra la Suerte Moral, entonces éste habrá sido refutado (en tanto que respuesta global).

En la siguiente subsección, que comienza con algunas consideraciones acerca de la importancia de la distinción entre tipos de juicios prácticos y su papel en el caso contra la suerte moral, trataré de rechazar EM.

3.2.1. Tipos de evaluación moral y la insuficiencia de EM

Como puede verse, un elemento central en nuestra discusión es la correcta comprensión de la diversidad de prácticas de valuación y evaluación moral en las que nos vemos inmersos. A grandes rasgos, y limitándonos a la esfera práctica, podemos distinguir juicios prácticos, juicios prudenciales y juicios morales, en sentido propio. E incluso dentro de esta última área de evaluación moral, encontramos una gran variedad de prácticas de juicio. Hay juicios morales que se refieren a cosas, estados de cosas o sucesos, como cuando decimos que x es moralmente bueno o y es moralmente malo. Y hay otros juicios que tienen por objeto acciones (o grupos de acciones) en sí mismas; en este sentido, decimos que la acción A es correcta y acción B es incorrecta. No obstante, la cuestión de la suerte moral no se plantea respecto a juicios morales meramente referidos a estados de cosas, acciones, propiedades o sucesos en sí mismos, ni respecto a valoraciones de los tipos bueno y malo o correcto e incorrecto, si no está implicada en ellos la noción de agencia o agente. La cuestión de la

suerte moral se refiere exclusivamente a los juicios morales referidos a agentes. Así, los juicios anteriores adquieren relevancia para nuestro tema sólo cuando se refieren también a agentes que producen esos estados de cosas o acciones, o poseen tales propiedades, etc., o a esos estados de cosas, acciones, propiedades o sucesos en tanto que producidos por agentes.

Por otro lado, nuestras mismas prácticas de evaluación moral de personas son muy diversas. Cabe remarcar que una correcta comprensión de éstas podría ser determinante para decidir nuestra cuestión. En particular, los adversarios de la suerte moral deben aportar una clasificación adecuada de nuestras prácticas cotidianas, y mostrar que entre los juicios distinguidos hay uno que claramente es el tipo fundamental, que responde exclusivamente al control del agente y, por lo tanto, está libre de la suerte. Aunque, alternativamente, también podrían promover una revisión de nuestras prácticas ordinarias si resulta que éstas no son realmente fieles al Principio de Control.

Una clasificación que se ha propuesto distingue, por lo menos, tres tipos de evaluación moral en virtud de aquello por lo que juzgamos a una persona: los juicios deónticos, los juicios aretáicos y los juicios propiamente de responsabilidad moral. Los primeros, estrechamente ligados a la obligación moral, se referirían a la corrección o incorrección de nuestras acciones, por lo que evaluarían a las personas por sus actos. Los segundos evaluarían el carácter del agente, sus virtudes y vicios morales, su bondad o maldad. Y, finalmente, los juicios de responsabilidad moral tendrían que ver con el elogio o laudabilidad moral y la censura o culpabilidad moral de las personas. Presuntamente, estos últimos evaluarían sus voluntades o decisiones.²⁰

²⁰ La distinción es de Zimmerman (2002). Nelkin (2004) le añade los *juicios axiológicos*, esto es, los juicios sobre estados de cosas que tienen que ver con las acciones de las per-

He de decir que esta clasificación resulta clarificadora e interesante, pero sólo con ella no se gana nada. Al adversario de la suerte moral no le basta con establecer esta distinción, sino que además ha de convencer-nos de que unos de estos tipos es el fundamental en nuestras prácticas de evaluación moral de las personas y que, además, está libre de la suerte. Asimismo, empezar llamando juicios de responsabilidad moral sólo a unos (dándoles con ello preeminencia) y aislarlos del resto parece ser una petición de principio, pues es esto lo que está precisamente en disputa. En realidad, los tres tipos de juicios tienen que ver con un agente, sus rasgos y sus acciones; la diferencia estaría en todo caso en aquello en lo que principalmente incide cada tipo: los juicios deónticos se centran principalmente en la evaluación de acciones de un *agente* (como correctas o incorrectas); los aretaicos, en el carácter o disposiciones (virtudes o vicios) del *agente* y, por último, los “de responsabilidad” se centrarían en el *agente* mismo (juzgándolo censurable o elogiable). Pero, como vimos en 1.1, las atribuciones de responsabilidad pueden deberse a diferentes aspectos del agente, sin que quepa confinarlos lógicamente o conceptualmente a un aspecto particular. De hecho, está lejos de ser autoevidente que los juicios de responsabilidad moral no puedan referirse a aspectos del agente como sus virtudes o vicios, ni mucho menos a acciones de éste. Además, la responsabilidad moral respecto a una determinada cosa es normalmente el resultado de tener una obligación moral para con esa cosa. Si soy culpable por pegarte es porque tengo la obligación moral de no pegarte.²¹

sonas (como buenas o malas). Zimmerman no distingue entre juicios deónticos y axiológicos (englobando ambos tipos bajo el nombre de *deontológicos*). Por último, a los juicios de responsabilidad moral los llama *hipológicos*. Véase también Zimmerman (2006) para una detenida consideración de las diferencias entre estos tipos de evaluación moral —especialmente los deónticos e hipológicos— y su distinta relación con la suerte.

²¹ Compárese la clasificación anterior con la siguiente de Greco (1995, p. 82-3), quien distingue: (1) juicios del tipo ‘La acción *A* es elogiable’ o ‘La acción *B* es censurable’, y

Parece, además, que esta clasificación confunde sin justificación el *locus* del juicio con el tipo de valoración que produce sobre él. Así, por ejemplo, sobre el carácter sólo podríamos emitir juicios de virtuosidad o vicio, o sobre las acciones de corrección o incorrección, pero no de censura o elogio. Por otro lado, es posible que la distinción fundamental entre los tipos de juicios que conforman nuestras prácticas cotidianas no responda tanto a qué aspecto del agente se juzga, sino a la *profundidad del juicio*. Esto es, hay juicios que se refieren a acciones, decisiones, intenciones o rasgos del agente que parecen ser más accidentales, mientras que otros lo hacen a propiedades más fundamentales, que forman parte del núcleo de su yo. Por ejemplo, no es igual de profundo el juzgar a alguien en virtud de una decisión o intención momentánea de la que pronto se arrepiente que por una decisión fruto de una seria deliberación, o una intención duradera que responde a deseos con los que se identifica más fundamentalmente. La idea sería que cuanto más profunda sea la evaluación más tocará el núcleo moral del agente, y más conectada estará con su merecimiento último.

En realidad, la mayoría de propuestas combinan la defensa de un elemento privilegiado en el agente, que constituiría el locus de la responsabilidad, de su merecimiento, que además, convenientemente considerado, otorgaría la máxima profundidad al juicio, reflejando el merecimiento

‘La persona *S* es elogiable por hacer *A*’ o ‘La persona *S*’ es censurable por hacer *B*’; y (2) juicios del tipo ‘*S1* es mejor persona que *S2*’, ‘*S1* es moralmente inferior a *S2*’, ‘*S* es un santo’ o ‘*S* es malvado’. Esta clasificación reagrupa todos los tipos anteriores, estableciendo como básica una distinción interna al tipo que Zimmerman y Nelkin convenían en llamar juicios de responsabilidad moral. Además, para Greco los fundamentales son los segundos, que se refieren a la *valía moral* o merecimiento verdadero del agente; mientras que los primeros, referidos a su *historial moral*, estarían mediados por la suerte —es curioso que son éstos los que Greco llama “de responsabilidad moral”. Creo que esta es una buena muestra de cómo de confusas son nuestras prácticas cotidianas de evaluación moral. Volveré sobre esta cuestión.

último, incondicionado, del agente. Este sería el tipo de evaluación moral fundamental del agente, que debería quedar libre de la suerte.

Retornando a la consideración de la Estrategia Moderada, veíamos que ésta se basaba en distinguir entre el merecimiento verdadero y el juicio que estamos justificados a proferir, dada la evidencia disponible (AE) —o, en otras variaciones, los juicios legales o pragmáticos, o los actos abiertos de censura. En todo caso, su mayor problema no es la distinción misma, sino la imposibilidad de extenderla más allá de la suerte resultante y circunstancial, dado su carácter esencialmente *moderado*. Ir más allá de ciertas consideraciones *plausibles* le haría perder, por definición, su carácter intuitivo, que le da ventaja sobre ER.

Consideremos el caso de las dos personas que comparten el propósito de asesinar a alguien, y que de hecho lo intentan, pero sólo una lo consigue. Llamemos Alonso al que lo consigue y Alfonso al asesino frustrado. Convengamos que ambos estaban igualmente convencidos de sus intenciones y que ambos han hecho todo lo que estaba en su mano para realizarlas, para llegar a matar a Lola el día D. Pero resulta que cuando Alfonso se disponía a disparar a la que iba a ser su víctima, Lola entró en una tienda repleta de gente, perdiendo así su oportunidad y quedando frustrado su plan. Alonso tuvo más suerte (o menos) y pudo acabar con la vida de Lola, tal y como lo había planeado. Pues bien, si aplicamos EM a este caso, tenemos que el merecimiento (culpabilidad) de ambos es el mismo, dado que sus disposiciones y carácter son iguales en lo relevante; si bien la falta de evidencia puede llevarnos a juzgar más negativamente a Alonso, o a censurarlo de una manera más abierta, o a castigarlo legalmente con mayor rigor (homicidio o asesinato contra homicidio frustra-

do). Parece que la apelación a nuestras intuiciones más kantianas surge efecto aquí.²²

Pero extendamos el caso añadiendo una tercera persona o contraparte, Alfredo, que es igual en lo relevante a Alonso y Alfonso, pero que fue encarcelado (por un robo de poca monta) unos días antes del día D. Lola se vuelve a salvar, esta vez porque Alfredo no pudo acudir al lugar donde se llevaría a cabo el asesinato. Pero según EM, Alfredo es igual de culpable que Alonso y Alfonso, pues su intención era ir a cargarse a Lola el día D. Es igual de culpable a pesar de que lo de Alonso es un homicidio, lo de Alfonso un homicidio frustrado y lo de Alfredo, no sé... un propósito de homicidio —aun no había conseguido la pistola, ni había pensado dónde ni cuándo cargársela, etc. Pero podemos todavía añadir un cuarto contraparte, Alberto, para el que, aunque todo apuntaba a que se convertiría en un exitoso sicario, un accidente de coche, huyendo de la policía, le apartó de la que pensaba iba a ser una muy lucrativa carrera. Alberto ya había sido contratado para acabar con la vida de Lola antes de sobrevenirle el accidente, y por lo tanto era parte de sus planes inmediatos disparar a Lola el día D. Según EM —estrictamente, según el Principio de Control—, Alberto es igualmente culpable que Alberto, Alonso y Alfonso. En otras palabras, su *grado*²³ de responsabilidad será el mismo, aunque unos sean responsables por más cosas que otros.

Sin embargo, parece que EM ya no puede hacerse cargo de este tipo de casos, dado que para el veredicto de la igual culpabilidad de todos

²² Con ligeras variaciones, este caso es representativo tanto de la suerte resultante como de casos de suerte circunstancial que se asemejan profundamente a los de suerte resultante. En concreto, tal y como lo he descrito, es un caso de suerte circunstancial.

²³ La noción de *grado* (en oposición a *alcance*) de la responsabilidad se debe a Zimmerman (2002, p. 568-70). El alcance no es relevante, según Zimmerman, para la determinación de la responsabilidad moral; sólo lo es para juicios de tipo *deóntico* (o también *aretaico*) que son ajenos al merecimiento último del agente.

ya no puede apelar meramente al carácter o disposiciones actuales, a la vez que el abandono progresivo de la historia real del agente hace que el compromiso con lo “plausible” se vaya debilitando. Nuestras intuiciones para con el caso se vuelven cada vez menos inmediatas, convirtiéndose en necesario algún tipo de revisionismo respecto a nuestras prácticas de juicio. Por ello, si la extensión de EM a casos progresivamente más antecedentes de suerte moral resulta cada vez más difícil, EM se torna insuficiente como caso global contra la suerte moral en todos sus tipos. Sin embargo, el relevo lo tomará aquí sin problemas ER, cuyo único compromiso es el Principio de Control.

En todo caso, aunque EM fracase como caso global, ello no impide que pueda resultar adecuado como argumento particular contra la suerte resultante y algunas formas de suerte circunstancial, formando parte de estrategias mixtas (que combinan diferentes argumentos para diferentes tipos de suerte moral) o híbridas (que aceptan ciertos tipos de suerte moral y rechazan otros).

En los capítulos siguientes (6 y 7) se evaluarán más detenidamente los argumentos que EM ofrece contra la suerte circunstancial y resultante, respectivamente. Aquí, doy por demostrada la inviabilidad de EM como Caso Global y paso ya a considerar ER. No obstante, reconsideraremos el núcleo esencial de EM más adelante.

3.2.2. ER: merecimiento incondicionado e historias posibles. Dudas escépticas y revisionismo

Como ya he dicho, el rechazo de la suerte moral depende crucialmente de mostrar (i) que hay un tipo de evaluación moral privilegiado (fundamental), y (ii) que éste está libre de la suerte. Este tipo de evaluación moral fundamental reflejaría el merecimiento verdadero incondicio-

nado del agente, que necesita ser caracterizado como independiente de la acción, e incluso independiente de las disposiciones (o voluntad) actuales del agente. Una objeción que puede resultar fatal para esta estrategia es que, de hecho, no existe un único tipo privilegiado de evaluación moral. Sin embargo, no seguiré esta objeción aquí. En cambio, por mor del argumento, asumiré que lo hay; que hay un tipo de evaluación moral que es el privilegiado, y que es función del merecimiento verdadero del agente, y persiguiendo esta idea trataré de concluir que, llevada a sus últimas consecuencias, ésta resulta incoherente.

Mi argumento, en su conjunto, funciona como una reducción al absurdo, del siguiente modo:

1. Hay un tipo de evaluación (moral) que está completamente libre de suerte.
2. El tipo de evaluación libre de suerte es función del *merecimiento verdadero* o *último* de la persona evaluada (condición necesaria).
3. El merecimiento verdadero puede ser condicionado o incondicionado.
4. Un merecimiento verdadero condicionado es insuficiente para nuestro propósito (pues la suerte no quedaría completamente excluida).
5. Pero la idea de un merecimiento verdadero incondicionado es ininteligible.
6. Por lo tanto, no hay una noción de merecimiento verdadero que satisfaga 1.
7. Por lo tanto, no hay un tipo de evaluación moral completamente libre de la suerte.

Creo que nadie discutirá 1 a 4. Quien dude de 4, debe recordar que estamos buscando aquí un tipo de evaluación moral libre de la suerte en todos sus tipos. Además, 4 ha quedado establecido por la discusión anterior de EM. En lo que sigue, me centraré en 5, esto es, en la idea de un *merecimiento verdadero incondicionado*, pues su posibilidad es crucial para el

argumento contra la suerte moral. Mi rechazo de 5 se deberá a dos razones. La primera, más débil, responde a su inaplicabilidad. Reconozco que para algunos teóricos esta primera razón puede constituir una petición de principio; sin embargo, yo la juzgo sustantiva. En todo caso, la que considero definitiva es la segunda razón: su incoherencia final. De aquí, se seguirán trivialmente 6 y 7.

Veíamos que EM fracasaba cuando los escenarios a evaluar se alejaban en exceso del escenario real. Sin embargo, si estamos dispuestos a llevar la prevalencia del Principio de Control hasta sus últimas consecuencias, abandonando el compromiso con lo intuitivo y plausible, o con un elemento real en el agente que constituya el locus del juicio, distinciones como la de *grado* y *alcance* de la responsabilidad moral,²⁴ u otras ulteriores, pueden hacer el trabajo sin problemas. Lo esencial es que el elemento de contingencia introducido por *lo que de hecho* somos responsables quede excluido.

Esta estrategia se apoya irrestrictamente en la idea de que el agente es responsable por aquello que habría hecho si hubiese tenido la ocasión. Esto es, construyendo sobre el caso anterior de los asesinos, imaginemos que existe un tal Alonso*, contraparte de Alonso, pero que un día en su tierna infancia fue raptado, retenido y violado por un pederasta. Este hecho marcaría el resto de su vida; el que iba a ser un chico extrovertido, de carácter, bravucón, aficionado a vagar por los barrios bajos, trapichear, con facilidad para meterse en líos y para conocer a tipos duros, etc., que

²⁴ Recordemos que la distinción entre grado y alcance de la responsabilidad no puede implicar la existencia de dos tipos distintos de juicios de responsabilidad moral, al menos igualmente fundamentales. Greco (1996), por ejemplo, distingue dos clases de juicios de responsabilidad moral —la que evalúa el *expediente moral* (juicio de responsabilidad propiamente dicho) y la que evalúa la *valía moral* o merecimiento último— pero sólo el segundo es el fundamental.

finalmente asesinaría a Lola; se convirtió, desde entonces, en un niño muy introvertido, apagado, incapaz de meterse con nadie, temeroso de andar dos manzanas más allá de su casa.²⁵ Pues bien, para ER, si la diferencia entre ambos se debe exclusiva o principalmente a este hecho; en otras palabras, que no depende del control de Alonso* (y podemos estipular que si este hecho no hubiese ocurrido su desarrollo no hubiese divergido del de Alonso, salvo por otros posibles sucesos también fuera de su control); entonces ambos son igualmente responsables o culpables. En realidad, Alonso* no es responsable *por* el asesinato real de Lola, pues no la ha asesinado, sino responsable *tout court*²⁶ (o contrafácticamente responsable), dado que la hubiese asesinado de no ser por aquel incidente.

Así, ER no tiene más remedio que recurrir a una multiplicación (¿infinita?) de las distinciones. Cuando nos ocupábamos de José y Josué, veíamos que pese a tener historias morales distintas (uno atropelló a alguien y el otro no), para el adversario de la suerte moral ambos tenían el mismo merecimiento verdadero, sobre la base de su misma voluntad (intenciones, carácter, etc.); con lo que teníamos una distinción entre el juicio basado en el mero *historial moral real*, o conducta efectiva (con sus consecuencias efectivas), del agente, y el merecimiento verdadero. Pero para seguir en esta dirección, ER necesita añadir una nueva distinción, entre este tipo de merecimiento, que va ligado a aquellas disposiciones

²⁵ Alonso* está inspirado en el personaje de Dave Boyle, de la novela *Mystic River* de Dennis Lehane (2001) —interpretado por Tim Robbins en la película homónima dirigida por Clint Eastwood en 2005.

²⁶ La noción de responsabilidad *tout court* (o “sin más”) aparece en Zimmerman (2002, p. 569-70). La idea es que se puede ser responsable *tout court* incluso aunque no se sea responsable *de* nada. En el ejemplo de los nazis de Nagel, aunque no exista algo (no ha dirigido la ejecución de judíos, etc.) de lo que Adolf, el mero simpatizante nazi, sea responsable (culpable); de hecho, podemos y debemos considerarlo responsable en el mismo grado que Rudolf, el colaborador nazi. Sin embargo, me parece más sencillo decir simplemente que Adolf sería (supuestamente) contrafácticamente responsable de la ejecución de judíos, etc.

que uno tiene dada su historia fáctica o real —que podemos llamar *merecimiento verdadero fáctico*—, y un tipo de merecimiento aún más fundamental, que dependerá de un conjunto más amplio de disposiciones posibles, que incluyen todas las historias contrafácticas posibles de la vida del agente. Llamemos a esto último *merecimiento verdadero esencial*.²⁷

Esta noción, surgida de la legítima aspiración a separar el núcleo duro moral de la persona —lo que se refiere estrictamente a la persona misma— del conjunto de factores (de origen especialmente externo, pero también interno) que contingentemente lo envuelven; se ha radicalizado, en virtud del anhelo kantiano de que *lo que más fundamentalmente soy*, y lo que exclusivamente debe ser juzgado, debe estar más allá de la suerte.²⁸ El merecimiento verdadero esencial, reflejo de la idea de responsabilidad moral última, sería ya el tipo de evaluación fundamental e incondicionado buscado, necesario para librarnos por completo de la suerte. En contraste con los tipos de merecimiento y responsabilidad anteriores, de naturaleza fáctica o, parcialmente, fáctica; se caracterizaría por ser *perfectamente racional* e “imputar una responsabilidad absoluta y por completo dentro del poder del agente”,²⁹ expresada en juicios “absolutos” e intemporales —en tanto que independientes de todo propósito y con una finalidad última.

Así, volviendo al caso de Alonso y Alonso*, su merecimiento (verdadero esencial) será el mismo, en tanto que éste ha de basarse en las disposiciones que el agente hubiese podido tener, dadas las *historias contrafácticas posibles* de su vida —y esto aunque, de hecho, Alonso* haya acabado convirtiéndose en un ciudadano modélico a raíz de aquel inciden-

²⁷ Para una explícita defensa de este movimiento, véase Greco (1995). En sus términos, esta nueva distinción sería entre la “valía moral fáctica” y la “valía moral esencial”.

²⁸ Cf. Williams (1981), p. 38.

²⁹ Feinberg (1970), p. 344.

te. Lo que cuenta, para el defensor a ultranza del Principio de Control, serán las disposiciones que el agente habría tenido si su historia fáctica — la historia que de hecho ha tenido y que depende de infinitos hechos que escapan a su control— hubiese sido otra, de entre el conjunto de historias posibles para su vida y que no se han dado de hecho. No cabe duda de que este movimiento, característico de ER, la convierte en una estrategia mucho más ambiciosa que EM, pero lo hace a un alto precio: la necesidad de abandonar la idea de que el carácter, intenciones o decisiones reales del agente constituyen el locus para su evaluación moral última, así como la armonía con nuestras intuiciones de sentido común (lo “plausible”).

Sin embargo, también podría suceder que el hecho de que Alonso acabase siendo un asesino, o alguien con la disposición de asesinar a un semejante, venga a depender de hechos anteriores en su vida, ajenos a su control; pongamos que sus padres le abandonaron de muy pequeño, que vivió en un ambiente muy degradado, etc., sin los cuales muy probablemente no hubiese llegado a ese extremo. Si esto es así, el merecimiento último auténtico de Alonso no es el mismo que el de Alonso*, sino que ahora es mucho más positivo. Tendríamos como resultado que un ciudadano modélico, que además fue víctima de tan cruel abuso sexual en su infancia, merecería un desprecio moral mucho mayor que el sanguinario asesino de Lola; lo cual ya empieza a ser realmente difícil de digerir.

No obstante, aceptaré la posibilidad y coherencia lógica de esta propuesta por mor del argumento; y propondré dos tipos de objeciones, que inciden principalmente en su inaplicabilidad. En primer lugar, la atacaré sobre la base de las dudas escépticas que parece suscitar la separación radical del merecimiento verdadero y la historial moral real del agente. En segundo lugar, defenderé la implausibilidad e imposibilidad real de

llevar a cabo el revisionismo radical (irrealista) que se sigue de sus conclusiones, además de sus consecuencias moralmente perniciosas.

Empecemos con la primera objeción. Nos hemos quedado con, al menos, dos tipos principales de evaluación moral: la que evalúa el *historial moral* del agente, y la que evalúa su merecimiento verdadero, función de lo que el agente *habría hecho*, dado el conjunto de sus historias contrafácticas posibles. Pero, ¿qué hay del *décalage* entre las bases para uno y otro tipo de evaluación? Esta posición crea una separación preocupante (quizá, insalvable) entre historia real y merecimiento que, por lo pronto, es claramente indeseable. No obstante, el adversario de la suerte moral puede reconocer que es realmente difícil emitir un juicio sobre el merecimiento verdadero (esencial), pero que esto no implica un escepticismo radical acerca del merecimiento verdadero. Juicios limitados acerca del verdadero merecimiento pueden ser razonables, incluso aunque debamos ser muy cautos al realizarlos.³⁰ Incluso, podría defenderse que esta posición produce la siguiente consecuencia positiva: un razonable escepticismo acerca del merecimiento último minaría nuestra rigurosidad a la hora de censurar a aquellos que se han visto inmersos en situaciones menos afortunadas que nosotros.³¹ Pero precisamente este es un resultado decididamente vindicado por los defensores de la suerte moral: una consecuencia que se sigue del reconocimiento de que la suerte se introduce en cómo juzgamos y somos juzgados moralmente es que no debemos ser especialmente rigurosos con quienes se han visto sometidos a circunstancias menos afortunadas. Y ello sin necesidad de postular una entidad como la de merecimiento verdadero, del tipo que hemos caracterizado.

³⁰ Véase Richards (1986) Greco (1995) y Rosebury (1995).

³¹ Greco (1995), p. 93-4.

Otra réplica disponible para el adversario de la suerte moral sería ésta: el historial moral de una persona es un signo de su merecimiento verdadero; las circunstancias en que una persona de hecho elige y actúa son un subconjunto del rango total de circunstancias en que la persona habría elegido y actuado; de este modo “el historial moral de una persona provee una ventana a la valía moral de la persona”.³² Sin embargo, esta tesis no parece ser del todo accesible a ER (puede que sólo sea accesible al defensor de AE que reconoce en las consecuencias, acciones y disposiciones efectivas un *indicador* del merecimiento). La idea es que una vez se ha renunciado a una concepción, digamos, falible del juicio de responsabilidad moral y se adopta una perspectiva infalibilista u omnisciente, el criterio para establecer que algo cuenta para determinar el merecimiento verdadero es tan exigente, que difícilmente podrá satisfacerse. Quedarse en un escepticismo moderado o afirmar que la historia de alguien nos suministra una ventana a su merecimiento verdadero es, bajo esas condiciones, mero pensamiento desiderativo. Una vez disociamos, por principio, el merecimiento verdadero del historial moral real, la conexión entre ambos se rompe irremediabilmente.

Quizá, llegados a este punto, la respuesta más coherente del defensor de esta posición fuese decir que el juicio de los demás es en último término imposible para nosotros, los humanos, pues hay una gran cantidad de información contrafáctica, necesaria para conocer el merecimiento verdadero, que nos es inaccesible. Podemos llamar a esto escepticismo à la *Dostoyevski*: no debemos juzgar nunca a los demás, sólo Dios, que es omnisciente, puede hacerlo.³³ A lo cuál podríamos replicar: ¿entonces,

³² Greco (1995), p. 93. Cfr. Richards (1986).

³³ Cabe decir que el *escepticismo sobre el juicio moral* es una posición distinta del escepticismo sobre la responsabilidad moral (como el de Spinoza, Derk Pereboom o Galen Strawson). El primero no tiene porqué incluir al segundo. El escéptico sobre el juicio

para qué sirve esta misma idea de merecimiento verdadero? Y parece que la única respuesta que queda ya disponible es, precisamente, la típicamente kantiana: constituye una *idea regulativa*, un ideal que regula nuestra consideración de las personas. Pero este movimiento, además de ser ajeno a los principales defensores de ER, supone reconocer que, finalmente, la viabilidad de la noción de merecimiento último sólo puede ser *ideal* — cuestión que trataré en el siguiente apartado.

Pero hay aún otra objeción significativa para la posición que venimos considerando: la irrealidad de sus prescripciones. El mismo Zimmerman, defendiendo esta posición, llega a la convicción (que, de hecho, se sigue de su idea de responsabilidad *tout court*) de que todos aquellos que, en ciertas circunstancias, hubiesen actuado libremente del mismo modo que actuó, por ejemplo, el colaborador nazi, son tan culpables como él.³⁴ Y lo mismo se aplica al elogio moral o laudabilidad. La conclusión resulta ser que todos nosotros somos censurables (y elogiables) por incontables cosas de las que “ni nos imaginamos”³⁵, dado que tenemos diferentes contrapartes que en situaciones posibles hubiesen actuado incorrectamente, y no está justificado un juicio diferente en virtud de consideraciones fácticas. Pero, así, la mayoría de nuestros juicios ordinarios, si no todos, se volverían ilegítimos. A lo que, finalmente, Zimmerman responde con la prescripción de una revisión general de nuestras prácticas, que

moral puede conceder que las condiciones para la responsabilidad moral son satisfechas en el mundo real, pero aunque de hecho haya acciones culpables, en ningún caso estamos justificados a juzgar que alguien es responsable de x , que ningún juicio de responsabilidad moral particular está garantizado. Para una defensa de una posición de este tipo, véase Rosen (2004).

³⁴ Zimmerman (2002), pp. 569-70.

³⁵ Zimmerman (1987), p. 226.

afectaría no sólo a los juicios morales, sino también a otras prácticas asociadas a éstos, como el castigo, etc.³⁶

Contra esto, cabe el siguiente razonamiento. Una primera consideración es de tipo metateórico: es siempre la teoría revisionista, se refiera a lo que se refiera, la que tiene todo el peso de la prueba. Estos es, una teoría revisionista parte siempre con la desventaja de chocar con nuestras intuiciones preteóricas, más si se refiere a nuestras prácticas cotidianas, y más aún si pretende alterarlas radicalmente. El carácter recalcitrante de nuestras intuiciones y prácticas supone un fuerte obstáculo para toda teoría revisionista, sin embargo esto no es definitivo (además, aunque la modificación fuese altamente difícil, aún así podría estar justificado el mismo intento). Sin embargo, la pervivencia de éstas prácticas es un hecho que juega a su favor. Por lo que cualquier propuesta de revisión radical, tiene no sólo que vencer en el debate, sino además convencer. Debe mostrar, de sobra, que esta revisión es realmente necesaria. Y, además, ha de estar avalada por un plan de viabilidad. (Puede que una revisión radical, si finalmente consiguiese hacerse efectiva, resultase en un desastre práctico, quizá irreversible; piénsese en ciertas ideologías basadas en una ingeniería social revisionista radical.)³⁷

Pero, además, ¿cómo podríamos revisar nuestros juicios morales de hecho en virtud de una multiplicidad de juicios esencialmente contra-

³⁶ Como dije, ER (y la posición de Zimmerman, en particular) parece ser una *teoría del error* de nuestras prácticas de juicio moral, según la cual los juicios de responsabilidad moral efectivamente se refieren al merecimiento verdadero, pero de hecho fallan en su referencia; por lo que son sistemáticamente erróneos. Por el contrario, la teoría conservadora (EM) sostiene (i) que los juicios de responsabilidad ordinarios no se refieren exactamente al merecimiento último, sino a una versión condicionada (mediatizada por la historia real y la evidencia disponible) de éstos; y (ii) que el merecimiento último existe. Este contraste pone de relieve que si bien ER gana en alcance y coherencia lógica, EM es mucho más razonable y aplicable.

³⁷ Me ocuparé más a fondo del revisionismo y sus formas en relación a la suerte moral en la Parte III de esta tesis.

fácticos? Si los múltiples juicios contrafácticos referidos a cada uno de nosotros, los cuales se definen por su carácter condicional —por ser juicios posibles, pero no reales—, han de ser igualmente tomados en consideración, se tornará imposible fijar un juicio determinado como el adecuado para cada sujeto y circunstancia. ¿Cómo revisar el castigo impuesto a alguien en virtud de consideraciones meramente contrafácticas? ¿Cómo podríamos llegar a una resolución acerca de su pena? La única solución parece ser la de que todos somos igualmente censurables y elogiados por todo. Pero esta generalización universal de la responsabilidad moral que acabaría neutralizando el mismo concepto de responsabilidad moral, bien por incremento de la culpabilidad, bien por su mitigación. El resultado contrario se me antoja, a todas luces, inadmisibles. Supondría, en realidad, otra forma de acabar con la idea de responsabilidad moral. Aplicado al castigo, supondría que nadie debe ser nunca castigado, o que todos merecemos serlo en la misma medida. La idea de administrar la responsabilidad moral, de realizar atribuciones particulares de responsabilidad moral, sólo tiene sentido dadas unas distinciones, unos polos positivo y negativo y unas gradaciones que han de ser adecuadamente establecidas y que se distribuyen desigualmente entre las personas.

La conclusión a la que llego, en esta sección, es que este tipo puro de merecimiento propuesto es completamente impráctico. Esto, en sí mismo, constituye a mi modo de ver una desventaja importante para esta posición. No obstante, es (todavía) posible comprender el elogio y la censura como reflejo de un tipo puro de merecimiento. Esto es, mantener que aunque la idea de merecimiento verdadero (esencial) sea prácticamente inerte, puede sin embargo ser la posición idealmente correcta.

No obstante, hay aun una objeción a esta salida que cabe considerar; a saber, que la cuestión de la suerte moral se origina dentro de nues-

tras prácticas de evaluación moral cotidianas, y no en escenarios lógicamente posibles. Por lo que la solución de la cuestión en términos de condiciones ideales, lo que en realidad hace es cambiar de tema. El fenómeno de la suerte moral parece introducir problemas importantes en nuestras prácticas ordinarias, problemas que afectan a unas prácticas que funcionan como guías de nuestras relaciones interpersonales. Si resulta que la solución a la cuestión no se refiere en ningún extremo a las prácticas cotidianas, no parece que sea una solución muy prometedora en relación al presente problema —lo será, en todo caso, a otro. En general, la cuestión es que esta posición parece haber dejado de lado la misma cuestión en discusión. Particularmente, creo que esta objeción es importante. Sin embargo, si tenemos en cuenta que lo que en el problema de la suerte moral se considera es una difícil compatibilidad entre un principio y unas prácticas, incidir sólo en las prácticas y negar la posibilidad misma de una solución en términos del principio puede constituir una petición de principio. Es por ello que no incidiré más en este punto. Mi réplica final no dependerá de esta cuestión.

Lo que hasta aquí se ha establecido es que la propuesta de ER es completamente impráctica, además de altamente problemática. Sin embargo, esta posición podría aún ser la correcta a nivel ideal, o lógico (o conceptual). Pero no podrá ya serlo si la noción de merecimiento verdadero resulta ser completamente imposible de fijar. Es lo que precisamente trataré de mostrar en el apartado siguiente.

3.2.3. ER: identidad e incoherencia

Tenemos, así, que ER ha de renunciar a una defensa práctica, a su aplicabilidad. Sin embargo, dispone de un arma poderosa: puede que sea la posición *idealmente* correcta. De hecho, la justificación final del kan-

tismo suele ser ideal: como avancé, se trataría de un ideal regulativo, que sobrepasa unas prácticas necesariamente contingentes y limitadas. Tampoco es extraño que este ideal de una justicia o merecimiento perfecto está detrás de la mayoría de creencias religiosas, que abrazan la idea de un juicio divino últimamente justo. Pero, contra esto, defenderé que la noción misma de merecimiento verdadero o último es finalmente incoherente, dado que resulta imposible de fijar, no sólo prácticamente, sino también en condiciones ideales.

Vimos que las acciones, en tanto que externas al agente y nunca libres de contingencias (esto es: la realización externa de las acciones), no podían ser el locus de la responsabilidad última o merecimiento verdadero. Sin embargo, el carácter, las intenciones o la voluntad reales tampoco podían ocupar su lugar. Incluso sus disposiciones más estables tampoco resultaban satisfactorias, dado que estas disposiciones surgieron, en parte, por la intervención de la suerte. Y si la suerte reaparece en escena, seguimos sin resolver la cuestión; lo que hacemos es meramente posponer su aparición —recuérdese que el merecimiento verdadero ha de ser incondicionado, es decir, no puede ser cosa de suerte en absoluto. El movimiento que se le abría aquí al adversario de la suerte moral era favorecer una distinción entre merecimiento verdadero fáctico y esencial, con lo cual se abandona la evaluación de las disposiciones reales del agente, su historia real (que incluye lo que el agente habría hecho dadas sus disposiciones reales), para pasar a evaluar lo que el agente habría libremente elegido y hecho en una diversidad de situaciones pertenecientes a una multitud de historias posibles de su vida. Era este último tipo de evaluación el que parecía evitar la suerte.

Sin embargo, estas historias contrafácticas posibles son, para cada agente, innumerables (aunque posiblemente no infinitas). Cada agente, al

largo de su vida, afronta situaciones que le van llevando a la realización de un historial moral determinado, de entre todos los posibles; a una formación y desarrollo particulares de su carácter y disposiciones. Así, estos sucesos, que están fuera de su control —y que no tienen porqué determinarlo, sino que basta con la presencia de situaciones y factores que abren y cierran oportunidades al agente e influyen sobre él— apartan al agente del resto de otras historias posibles para su vida moral. (Si queremos hacer más concreta esta imagen, creo que caben dos opciones: o bien hay una entre todas las historias morales posibles del agente que sería reflejo de su merecimiento verdadero, o bien su merecimiento verdadero es simplemente una afirmación esencialmente contrafáctica. En todo caso, lo que parece no admitir duda es que el merecimiento verdadero nunca puede identificarse con la historia fáctica, dado que en la vida real de un agente es inevitable la influencia de factores que están más allá de su control.)

En definitiva, vemos que el defensor del merecimiento verdadero se ve obligado a retroceder en aquello de lo que éste es función. Se ve obligado a ir progresivamente retrocediendo en los antecedentes de la acción, decisiones y disposiciones del agente. El resultado de este retroceso rampante es una paralela reducción progresiva de la identidad (real, pero también potencial) del agente. Para cada posible elemento de evaluación siempre hay antecedentes que contrafácticamente podrían haber resultado en otra solución real. Así, para que esta historia tenga un final feliz, en último término será necesaria una constitución original, que establezca las que serían las disposiciones o intenciones *verdaderas* del agente, en una historia ideal en la que ningún factor fortuito se cruzase en su vida. Pero sucede que, el desarrollo consistente de ER impide un movimiento como el que he sugerido. Fundar el merecimiento verdadero en la constitución

original del agente supondría fundarlo finalmente en algo que claramente está más allá del control del agente, pues si parece que hay algo que definitivamente está fuera del control del agente es precisamente su constitución original.³⁸

La gran fuerza de ER —que la hacía superior a EM— era su aspiración a llevar hasta sus últimas consecuencias la idea de que el juicio moral debe estar completamente libre de la suerte. O, en otras palabras, que el merecimiento verdadero es una función estricta del control; esto es, *de lo que exclusivamente el agente controla*. Lo que significa que su desarrollo consistente bloquea el movimiento sugerido. La conclusión lógica de ER es la reducción final de la identidad del agente a la nada, a un *mero yo* sin propiedad alguna, ni física ni psicológica. Lo único que podría quedar es un yo nouménico o trascendental, que está más allá de todas sus cualidades.³⁹ Pero, en tanto que no posee cualidades, no hay nada sobre lo que basar su merecimiento. La conclusión es, pues, que la negación de la suerte moral, llevada a sus últimas consecuencias, se torna una posición sin sentido, pues no hay finalmente nada (no hay agente) de lo que el merecimiento verdadero sea función.

³⁸ El freno último lo constituirían las *propiedades esenciales* del agente; esto es, aquellas propiedades que son esenciales para que el agente sea quien es. Además, sin unas propiedades esenciales sobre las cuales descansar, los contrafácticos se tornan falsos. Zimmerman reconoce que, si tiene sentido hablar de las propiedades esenciales de alguien, entonces la suerte no es completamente eliminable. Sin embargo, aceptar esto supone renunciar al desarrollo lógico de ER y adoptar una nueva estrategia (híbrida), que discutiré en el capítulo siguiente.

³⁹ La posición de Kant parecer ser coherente a este respecto. No obstante, su metafísica está lejos de ser aceptable hoy en día. Y, en todo caso —repito—, o bien este yo trascendental no tiene atributos, y entonces no hay nada que pueda ser la base para el merecimiento; o si los tiene, necesariamente escapan al control del agente. De cualquier modo, fundar el merecimiento de una persona en la constitución original (sus meras propiedades esenciales originarias) parece ser una posición sólo deseable para un calvinista extremo, que sólo tiene sentido a partir del compromiso con al existencia de un *alma* que desde el origen contiene el valor moral del individuo. (Puede que este compromiso sea necesario incluso para comprender la teoría de Kant.)

Cabe concluir, pues, que es la noción de *merecimiento verdadero* —tal y como fue caracterizada— la que finalmente resulta ser una ilusión. Puede que no esté de más apostillar la conclusión de mi argumento con esta sentencia de Wittgenstein: “Una nada presta el mismo servicio que un algo sobre el que nada puede decirse.”⁴⁰

De esta manera, el que empezó siendo un fin legítimo, a saber, aspirar a que las atribuciones de responsabilidad moral (en tanto que referidas al estatus moral de la persona) sean *profundas*, reflejen algo que “realmente pertenece a la persona”, aislando los rasgos que más fielmente definen a la persona de los rasgos formativos y ambientales más externos; si se pretende llevar al extremo, acaba convirtiéndose en una aspiración vana e incoherente. Podría decirse que una aspiración en principio intuitivamente legítima, se ha vuelto manifiestamente ininteligible al ser absolutizada o idealizada.

Con ello, concluyo que ER también fracasa como estrategia argumentativa general contra toda clase de suerte moral.

La meta principal de esta sección ha sido desacreditar la apelación a la noción de merecimiento verdadero, en el sentido fuerte de *función estricta del (o estrictamente proporcional al) control del agente*, para desembarazarse del fenómeno de la suerte moral. Cuando la perseguimos hasta sus últimas consecuencias, la idea resulta ser ininteligible. Así, en la medida en que el Caso Global contra la suerte moral necesariamente depende de esta noción, ningún argumento general podrá finalmente funcio-

⁴⁰ Wittgenstein (1988), p. 249; § 304.

nar.⁴¹ Además, una vez nuestras intuiciones generales contra todo tipo de interferencia de la suerte en nuestros juicios morales quedan desacreditadas, empeñarse en negar la existencia de la suerte moral se torna harto difícil.

3.3. Recapitulación y expectativas

3.3.1. La estructura de mi argumento

En este capítulo he expuesto y rechazado del Caso Global contra la Suerte Moral, esto es, el caso contra la suerte moral *en todos sus tipos*. En su reconstrucción, a partir de los diferentes argumentos propuestos contra la existencia de la suerte moral, distinguí dos estrategias fundamentales. Así, consideré en primer lugar, la Estrategia Moderada, respecto a la cual intenté mostrar que es claramente insuficiente como caso general, pues se halla fundamentalmente limitada a la suerte resultante y circunstancial. Rechazada esta estrategia, pasé a considerar la Estrategia Radical. Respecto a ella traté de mostrar que, si bien su extensión a tipos antecedentes de suerte moral no es lógicamente problemática, sí lo es llevarla hasta sus últimas consecuencias. ER resultó inaplicable y, lo que es más importante, incoherente. Cabe decir que si ER sólo fuese inaplicable, esto supondría un escollo importante para su defensa pero no impediría que pudiese ser la solución idealmente correcta. Sin embargo, al ser también incoherente—como he tratado de mostrar— queda definitivamente desacreditada. Así, ninguna de las estrategias consideradas (ni EM, ni ER) funciona.

⁴¹ La reducción al absurdo se completa con un argumento de tipo *trascendental*: la noción de merecimiento verdadero, cuya incoherencia interna he tratado de mostrar, es un presupuesto indispensable del rechazo radical de la suerte moral.

Pero con esto, el Caso Global contra la Suerte Moral no queda aún refutado. Cabría todavía la posibilidad de una estrategia mixta, como es la combinación de ER con una estrategia particular con respecto a la suerte constitutiva —que exploraré en el próximo capítulo. Además, el mismo rechazo del Caso Global contra la Suerte Moral no implica que todos los tipos queden, *ipso facto*, vindicados; sino solamente —y no es poco— que una estrategia general única ya no es posible. Por ello, mi réplica no será completa, hasta que no consiga responder individualmente a la negación de la suerte moral en cada uno de sus tipos principales —que es lo que trataré de hacer en los capítulos restantes de la Parte II.

Antes de pasar a ocuparme de esto, quisiera explicitar las diferencias entre el argumento que hasta el momento he defendido y el que podemos llamar *argumento del todo o nada*, debido a Michael S. Moore, y que trata también de rechazar la impugnación de la suerte moral.⁴²

3.3.2. Contraste con el ‘argumento del todo o nada’

Moore ha presentado un argumento especialmente dirigido al adversario de la suerte moral resultante, que tratar de mostrar que, retrocediendo en las causas de la acción, no hay ningún punto privilegiado en el que podamos trazar una distinción clara entre lo que el agente realmente controla y lo que no. Su conclusión es que o aceptamos en último término que las consecuencias dependen del agente, y cuentan para su merecimiento moral, o nada lo hace. No obstante, si aceptamos que nada depende de él ni cuenta para su merecimiento moral, nos vemos abocados al escepticismo acerca de la responsabilidad moral. Pero dado que esta conclusión es considerada inaceptable, no hay más remedio que admitir que las consecuencias de la acción dependen del agente y son parte de su me-

⁴² Agradezco a Dana Nelkin que me instara a clarificar esta cuestión.

recimiento moral. Veamos con más detenimiento los pasos cruciales de este argumento.⁴³

El blanco principal del argumento es lo que Moore llama la Posición Educada Estándar (o PEE), según la cual *los resultados son irrelevantes para el merecimiento moral*. Imaginemos el caso de un asesino que intencionadamente causa la muerte de su víctima con un arma de fuego. La explicación causal de la acción, retrospectivamente desde la muerte de la víctima hasta sus orígenes, sería la siguiente:

- (7) Muerte de la víctima.
- (6) Movimiento del dedo del asesino sobre el gatillo.
- (5) La voluntad o volición de mover el dedo.
- (4) La intención, plan o elección de matar a la víctima, que se ejecuta por medio de (5).
- (3) El conjunto de creencias y deseos, que pueden formar un silogismo práctico válido, que se ejecuta con (4).
- (2) Los rasgos más generales de carácter, que causan las creencias y deseos motivantes en una ocasión particular.
- (1) La clase de elecciones amplias, formadoras del carácter.

Con esta secuencia a la vista, cabe preguntarse dónde debe situarse el límite entre lo que está bajo el control del agente, y que cuenta para su merecimiento, y lo que no.

Los defensores de PEE pueden situar el foco de la responsabilidad en aquello que el agente *intenta* hacer. Pero, ¿con que se identifica este intentar? Una primera opción es situar la línea entre los factores que ocurren a partir de los movimientos corporales del agente y los factores antecedentes —entre (6) y (7). El motivo es que hay un gran número de sucesos que pueden ocurrir entre el movimiento del dedo por parte del asesino

⁴³ Cabe remarcar que el argumento de Moore —en Moore (1997), cap. 6—, junto con Adams (1985), son dos de los argumentos más destacados en la defensa de la suerte moral, más allá de los argumentos fundacionales de Williams y Nagel.

y la muerte de la víctima, que el agente no controla como controla los primeros. Pero, de igual modo, hay un gran número de sucesos que median entre los estados mentales (la volición) del agente y sus movimientos físicos (el intento), como disfrutar de una mayor o menor agilidad u otras habilidades corporales, que el agente no controla. Por ello, situar el locus de la responsabilidad en este punto parece arbitrario.

Alternativamente, continúa el argumento, podríamos decir que lo que importa es (5), la volición o intento mental. Pero, a su vez, la formación o mantenimiento de esta volición depende de la presencia o no de factores psicológicos, como la concentración o la distracción, la constancia o la impaciencia, la excitación, el autoengaño, etc., que van unidos a la posesión o no de ciertas capacidades por parte del agente. Así, tendremos que seguir retrocediendo si queremos encontrar un fundamento seguro para el control del agente. Quizá el lugar adecuado donde trazar la línea sea (4), la intención misma, más allá de su substanciación en una volición específica. Pero también en este punto encontramos problemas semejantes a los anteriores.

Quizá (4) —las intenciones, planes y elecciones— es un punto especialmente relevante. Uno puede afirmar que (4) posee una significación moral especial, ya que parece ser algo que el agente controla *completamente*, o que al menos controla de una manera distinta a lo que está más allá de (4). Pero, claro, ¿cómo se forma (4)? ¿Cómo se forman nuestras intenciones, planes y elecciones? Para Moore es evidente que su formación depende de una base anterior (3), constituida por un conjunto de creencias, deseos y actitudes, que no está claro en qué sentido puede controlar voluntariamente el agente. Ulteriormente, éste ha de poseer ciertas capacidades y han de darse una serie de oportunidades favorables (2). Y, por último, las elecciones de (1), en tanto que las más básicas, son, para

Moore, las que menos pueden caer bajo el control del agente. Parece que, retrocediendo, retrocediendo, la conclusión a que nos vemos abocados es una especie de fatalismo, pues las elecciones de (1) dependen a su vez de factores antecedentes, quedando finalmente anulada la responsabilidad moral. El último movimiento de Moore es, pues, afirmar que dado que esta conclusión no parece deseable, no tenemos más remedio que *parar en el carácter*; esto es, “quién somos nosotros *es* nuestro carácter, así las fortuidades que determinan quienes somos son irrelevantes. Ciertamente, no podemos controlar estos factores, pero ello se debe a que no hay *nadie* que los controle [...]”.⁴⁴

El paralelismo entre este argumento y el mío es obvio. En particular, ambos tratan de establecer una reducción al absurdo de la posición que atacan. Sin embargo, creo que el argumento de Moore tiene importantes puntos flacos, que no tiene el mío. En primer lugar, el *argumento del todo o nada* depende de un rechazo *ad hoc* del escepticismo moral, que es la conclusión primera del argumento. Esto es, dado que el argumento nos conduce necesariamente al escepticismo moral y esta es una consecuencia indeseable —para aquellos previamente comprometidos con que las personas realmente somos moralmente responsables—, se sigue, es la conclusión final, que somos responsables por todos los elementos anteriormente listados, incluidas las consecuencias de la acción. En este sentido, el argumento no es válido para los que no acepten la existencia de la responsabilidad moral, o que no lo hagan a priori. Moore podría replicar: mi argumento no es para aquellos de vosotros que no estéis convencidos de que de hecho somos moralmente responsables y de que es erróneo todo

⁴⁴ Moore (1997), pp. 244-5.

razonamiento que conduzca a la conclusión contraria.⁴⁵ Pero esto es reducir de manera muy significativa la fuerza del argumento.

Mi argumento, por el contrario, evita la noción de control y ataca primariamente la noción de merecimiento verdadero, en tanto que imprescindible para impedir por completo que la suerte se inmiscuya en el juicio moral que alguien merece. Mi conclusión en relación al merecimiento verdadero no depende de ningún condicional; por lo que, si el argumento es correcto, la conclusión se sigue necesariamente. No obstante, aunque más sólida, mi conclusión es menos ambiciosa, en tanto que no implica que haya formas de suerte moral que no puedan negarse convincentemente. Y aquí estriba la otra gran diferencia: que Moore pretende que con su argumento queda establecida la existencia de todas las formas de suerte moral. Pero esto es algo que no se sigue, pues podría haber razones particulares para no aceptar unos tipos de suerte moral, aunque no pueda negarse la existencia de otros.

Su conclusión parece comportar que somos igualmente responsables, en la misma medida, de todos los elementos considerados. Pero esto parece constituir una *falacia de la pendiente descendiente*. Su razonamiento es que si no somos moralmente responsables de las consecuencias de nuestras acciones, tampoco lo somos ni de nuestras acciones, ni de nuestras intenciones, deseos y planes, ni de nuestro carácter, en último término. Pero esto es lógicamente distinto de decir que si somos responsables de las consecuencias de nuestras acciones, es porque lo somos de nuestras acciones, y si lo somos de éstas es porque lo somos de nuestras intenciones, etc. Por lo que no se sigue su conclusión de que o controla-

⁴⁵ Hay que tener presente que Moore es un filósofo del derecho que enmarca sus razonamientos en una defensa de la relevancia de las consecuencias para el castigo legal, por medio de su relevancia para el merecimiento moral (lo que para él es fundamental, pues es un *moralista legal*).

mos y somos responsables por toda la cadena causal, o no controlamos nada ni somos responsables de nada, con lo que “la responsabilidad — incluso aunque sólo sea en el sentido de culpabilidad— desaparece.”⁴⁶ Se puede incluso alegar que quizá la falta de control sobre factores que impiden la elección no sea lo mismo que la falta de control sobre factores que impiden que la elección se realice en una acción. Y, en todo caso, es falso que, como afirma Moore, “no tenemos *más control* sobre los factores necesarios para decidir matar del que tenemos sobre los factores necesarios para matar.”⁴⁷ Sí que lo tenemos. Tenemos un mayor control sobre nuestras decisiones que sobre las consecuencias de éstas, aunque la suerte siempre pueda inmiscuirse.

De hecho, contra su conclusión se puede alegar lo que Linda Zagzebski ha llamado el *efecto acumulativo* de la suerte moral. La idea es esta:

Los rasgos internos de carácter originan disposiciones en circunstancias específicas, que llevan a la formación de intenciones particulares, que llevan a la realización de acciones, que a su vez producen consecuencias externas. En cada estadio siguiente se añaden nuevos elementos de suerte, de modo que el mayor grado de suerte se encuentra en las consecuencias, el menor en los rasgos de carácter.⁴⁸

Ante este efecto, y tras aceptar la imposibilidad de eliminar completamente la suerte, Zagzebski propone centrar la evaluación en las disposiciones del agente, con la meta de reducir lo más posible el componente de suerte. Particularmente, creo que este efecto acumulativo es real y significativo, y

⁴⁶ Moore (1997), p. 242.

⁴⁷ Por lo menos, para todo aquél que, sea compatibilista o incompatibilista, no niegue las intuiciones compatibilistas acerca de la cantidad o grados de control de los que puede disfrutar un agente.

⁴⁸ Zagzebski (1996), p. 72.

clarificarlo ocupará una parte de mi esfuerzo en los capítulos siguientes. Sin embargo, hay que adelantar que la cosa no es tan sencilla como pretende Zagzebski. Por ejemplo, un kantiano podría replicar que los rasgos de carácter son menos voluntarios que las decisiones, intenciones o acciones de un agente; esto es, que son menos sensibles a nuestra (buena) voluntad inmediata. Volveré, también, sobre esto.

3.3.3. La distinta significación moral de los diversos tipos de suerte moral

Cabe, pues, destacar que argumentos como el de Moore y el aquí presentado son inversamente paralelos a ER. Esta última estrategia consistía en negar la suerte moral por medio de llevar el Principio de Control hasta sus últimas consecuencias; mientras que la alternativa consiste en afirmar la suerte moral por medio de la reducción al absurdo del mismo principio. Sin embargo, mi argumento, a diferencia del de Moore y de ER, es compatible con la existencia de tipos diversos de control —los requeridos por los diferentes tipos de suerte moral (control de las consecuencias, control de las acciones o control de la constitución)— que es, de hecho, lo que hace que sea necesario para el triunfo final de mi defensa de la suerte moral que los razonamientos de los tres capítulos siguientes resulten correctos.⁴⁹

⁴⁹ Ambos argumentos mostrarían que no hay un merecimiento último, porque no hay un punto de control absoluto o un control último; además de que no hay una distinción clara entre la ejecución interna y la externa ni, por lo tanto, un límite que separe (absolutamente) lo que el agente controla y lo que no. En todo caso, si finalmente resultase que mi argumento es, en algún sentido, una extensión del de Moore, o una nueva versión; sólo con que sea más completo y detallado, ya me daría por satisfecho. Pero creo que no es el caso, pues en tanto que combina argumentos independientes para tipos de suerte moral particular, la mía no es ya una estrategia única para todos los tipos, como la de Moore, sino híbrida.

Esto parece que me situaría en una posición más débil. Como ha destacado Dana Nelkin argumentos como el de Moore o Zimmerman (ER), que demuestran o rechazan la suerte moral en todos sus tipos, tienen de su parte, además de su elegancia lógica, el desplazar el peso de la prueba a aquellos que pretenden combinar argumentos independientes para los diferentes tipos de suerte, o quieren defender un(os) tipo(s) de suerte moral y no otro(s). Si rechazamos la estrategia de tipo general, el reto será encontrar un principio desde el que poder trazar la línea entre lo que es significativo para el merecimiento moral del agente, o aquello que el agente realmente controla, y lo que no.⁵⁰ No obstante, este reto descansa en la dudosa pretensión de que la meta es encontrar un único principio que concluya el debate. Además, la idea de que hay diferencias (morales) importantes entre (algunos de) los diferentes tipos de suerte moral, que justificaría la necesidad de respuestas diversas, viene avalada por consideraciones como la anterior de Zagzebski y otras que trataré brevemente de presentar ahora.

Consideremos, para terminar, el siguiente resumen del contraste entre los diferentes asesinos que vimos más arriba (con una ligera variación):

- Alonso planea el asesinato de Lola, lleva a cabo todos los preparativos necesarios, se presenta en el lugar L el día D, y asesina a Lola.
- Alan planea el asesinato de Lola, lleva a cabo todos los preparativos necesarios, se presenta en el lugar L el día D, pero cuando se dispone a disparar a Lola no acierta en el tiro.
- Alfonso es igual en lo relevante a Alonso y Alan, pero cuando se dispone a disparar Lola entra en una tienda, perdiendo su oportunidad y quedando frustrado el plan.
- Alfredo es igual en lo relevante a Alonso y Alfonso, pero que fue encarcelado el día antes de que alguien le propusiera acabar con la vi-

⁵⁰ Nelkin (2004).

da de Lola (cuando, dada su biografía y penuria económica, habría aceptado sin dudarlo).

Alonso* es igual en lo relevante a Alonso y Alan, con la diferencia que en su infancia fue raptado y violado por un pedófilo. Este hecho influyó en su carácter de tal manera que ya nunca sería capaz de meterse con nadie, y menos de asesinar.

Podemos distinguir entre ellos aplicando el anterior esquema de los estadios de una acción de Moore. Alonso y Alan sólo se diferencian para algo que pasó en (7), a saber, el matar finalmente a Lola o no; el contraste entre ellos constituye un ejemplo paradigmático de suerte resultante. La diferencia de estos para con Alfonso es que no llega a realizar (6), no disparar; y de los tres para con Alfredo, es que éste último no llega ni a realizar (5). Cabe remarcar que el contraste de Alfonso con Alonso y Alan, y de Alfredo con los tres, constituye casos de suerte circunstancial más o menos próxima a la realización de la acción. Finalmente, Alonso*, que no llega ni a formarse la intención de matar a Lola (4), supone un caso de suerte antecedente —constitutiva, en general; o formativa, en particular— en relación al resto. Hay que decir que las diferencias tanto respecto a (4), como respecto a (3), (2) y (1) —esto es, las creencias y deseos que posee, el carácter presente y las decisiones formativas del carácter, respectivamente—, darían lugar a diferentes tipos de suerte antecedente.

Alonso es el único que ha asesinado a Lola; si bien las diferencias de éste con respecto a los demás se deben, por hipótesis, a hechos que están más allá de su control. Por lo tanto, si aplicamos consistentemente el Corolario del Principio de Control:

(CPC) Dos (o más) personas no deben ser moralmente evaluadas de manera diferente si las únicas diferencias entre ellas es deben a factores que están más allá de su control.

tendríamos que todos ellos son igualmente culpables, comparten el mismo merecimiento moral. Sin embargo, esta conclusión es, a mi modo de ver, claramente antiintuitiva. Como también lo es el siguiente principio, inverso a CPC:

(CPC-i) Si es cosa de suerte que una persona no controle (1), es igualmente cosa de suerte que no controle (2), y si lo es respecto a (2) lo es igualmente respecto a (3), etc., hasta finalmente respecto a (7).

El carácter de Alonso* es lo más opuesto a un carácter violento (nunca mataría a nadie), si bien es así sustancialmente a causa de su desarrollo infantil y, en particular, del incidente que sufrió. (Alternativamente, para quien dude de este ejemplo, podríamos considerar que fue cuando era ya un post-adolescente, cuando la experiencia de la muerte de su mejor amigo a manos de un matón de la ciudad, le apartó para siempre del mundo del crimen.) La hipótesis afirma que si eso no hubiese pasado, él habría actuado igual que Alonso. Pero, Alonso decidió voluntariamente asesinar a Lola, planeó al crimen, consiguió la pistola, se plantó en el lugar indicado, disparó y la asesinó. En este sentido, el hecho de que un suceso ocurrido en la vida de Alonso* produjese un cambio para mejor en su carácter, junto con su asimilación de este hecho (que sólo es posible si éste acaece), forma parte del merecimiento de Alonso*. Por lo tanto, yo diría que, al menos, el merecimiento de Alonso* no puede ser en ningún caso el mismo que el de Alonso; así como que el del resto tampoco parece ser el mismo, aunque mis intuiciones son menos claras, y pueden chocar con las intuiciones de otros. Es cierto que fue cosa de suerte que Alberto fuese encarcelado el día antes de que le propusieran asesinar a Lola, pero ello le evitó que aceptara, y lo que sigue.

Así, estrategias generales como la de Zimmerman y la de Moore, vienen a decirnos que si uno de ellos es culpable, lo son igualmente el resto; o que si hay algún factor que es cosa de suerte, lo es igualmente el resto de factores. Y quien no acepte este tipo de razonamiento, continúa el argumento, deberá justificar un punto donde trazar la distinción entre la suerte que puede inmiscuirse en el merecimiento y la que no, o la suerte que socava el control y la que no. Mi respuesta a este reto es la siguiente: creo que existen razones moralmente significativas para dar diferente peso, en cuanto al juicio moral, a elementos distintos dentro del poder del agente. Aunque sea imposible, finalmente, impedir que la suerte se inmiscuya en todos ellos, sí que cabe distinguir entre la mayor o menor relevancia (en relación al juicio moral) con que lo hace.

Hay razones para afirmar que —en los casos normales de no coerción, ni patología, adicción, etc.— escoger es central para nuestra *auto-comprensión moral*, como no lo es lo que está más allá de nuestra elección; controlamos nuestras decisiones porque son nuestras decisiones, si bien éstas dependen causalmente de factores que previamente no escogimos. O quizá lo central para nuestra agencia moral sea más bien nuestro carácter, incluyendo nuestros valores, intereses y compromisos duraderos. O lo es nuestra Voluntad, etc. En todo caso, son opciones que aún son transitables. Ciertamente, mi respuesta es, en este punto, ambigua; pero trataré de clarificarla y defenderla en los próximos capítulos.

Dónde estamos y adónde vamos

En este capítulo he defendido que, por diversas razones, el Caso Global contra la suerte moral fracasa. No obstante, permanece aún disponible la opción de negar la suerte moral (en todos sus casos) mediante una estrategia mixta, que combine la viabilidad general de ER, salvo para la suerte moral constitutiva, la cual puede rechazarse en virtud de argumentos diferenciados.

Los argumentos propuestos a este respecto pueden agruparse en dos tipos: (i) los que niegan la coherencia de una noción como la de suerte constitutiva; y (ii) los que tratan de neutralizar la suerte antecedente, en general, por medio de la correcta intervención del agente en la configuración de su propia constitución o carácter posterior. En el primer caso, se esgrime que si la existencia misma de alguien depende de la posesión de unas determinadas propiedades esenciales, siendo imposible su existencia sin éstas, no puede decirse con sentido que éste tenga o haya tenido *suerte* en relación a estas propiedades. En este tipo de objeción está en juego la misma noción de suerte. Para el segundo tipo de argumento, resultará fundamental la idea de que, si bien poseemos un conjunto de rasgos con los que simplemente nos encontramos (lo que podemos llamar el carácter recibido o heredado), sin embargo aquello por lo que juzgamos y somos juzgados es por el *carácter en tanto que formado*, donde la intervención del agente es esencial.

Mi meta en los dos capítulos siguientes será caracterizar correctamente estos argumentos, esto es, cómo diferentes autores han rechazado la suerte moral antecedente (constitutiva y formativa), y tratar de articular mi réplica.

4. La suerte constitutiva

Sobre necesidad y contingencia en nuestros orígenes e identidad

- 4.1. Suerte antecedente y estrategias
- 4.2. ¿Es la noción de suerte constitutiva incoherente?
- 4.3. La noción de suerte
- 4.4. Suerte y fortuna, o la especial relevancia moral del carácter
- 4.5. Posibilidades formativas e indeterminación

...para su existencia [el ser humano], necesita una esencia, esto es, propiedades radicales que constituyen su carácter y que no requieren más que la provocación exterior para manifestarse.

A. Schopenhauer, *Sobre la libertad de la voluntad*, II

Su temperamento ya está cuajado, es inamovible. Primero el cráneo, luego el temperamento: las dos partes más duras del cuerpo.

J. M. Coetze, *Desgracia*, 1

En el capítulo anterior, vimos que era en la esfera de la suerte constitutiva donde finalmente se dilucidaba la cuestión. La estrategia que principalmente consideré fue la de extender el Principio de Control hasta sus últimas consecuencias lógicas, tratando de deshacernos de la suerte moral en todos sus tipos. Mi conclusión fue que esta posición se volvía insostenible en tanto que últimamente conducía al desvanecimiento de la misma identidad del agente. Sin nada sobre lo que basar el merecimiento verdadero (que pudiese hacer verdaderos los mismos juicios contrafácticos), esta noción se tornaba ininteligible; de ello se seguía que el argumento global

contra la suerte moral ya no podía triunfar. Pero aún es posible una estrategia mixta que combine el argumento anterior con una respuesta específica para la suerte moral constitutiva. De este modo, habría un elemento realmente perteneciente al agente, y libre de la suerte, fundamento incontaminado de su agencia responsable y sobre el cual construir su merecimiento. Además, esta respuesta específica podría estar justificada en tanto que la suerte constitutiva parece relevantemente singular, en relación al resto de tipos. Explorar esta opción es de lo que me ocuparé en este y en el próximo capítulo.

4.1. Suerte antecedente y estrategias

A nadie se le escapa que las personas poseemos diferentes inclinaciones o disposiciones naturales, que acarrearán ventajas o desventajas, más o menos significativas. Hay personas que son más inteligentes, más atléticas, más atractivas, etc., que otras, lo cual parece deberse, en parte, a un mayor potencial biológico inicial en relación a estos rasgos. Este potencial inicial establece talentos o *handicaps* naturales, que hacen más o menos fácil —o, incluso, imposible— alcanzar ciertas metas en el desarrollo. A ello, hay que sumar las circunstancias sociales en las que uno se halla, que pueden acrecentar o disminuir estas desigualdades originales. Esta es una idea comúnmente reconocida, que algunos han llamado *lotería natural*.

Precisamente, John Rawls en su *Teoría de la Justicia* concedió un papel teórico importante a la vieja metáfora de la “lotería”, tanto natural como social, para caracterizar la dotación inicial de las personas: el punto de partida de cada individuo en la sociedad es el resultado de una lotería natural —los potenciales biológicos con que cada cual nace— y de una

lotería social —las circunstancias políticas, sociales y económicas en que se nace y crece. Las ventajas o desventajas naturales —los mayores o menores talentos y habilidades naturales— pueden desarrollarse más o menos dependiendo de las circunstancias sociales y de contingencias accidentales. Rawls comparaba el resultado de la lotería natural y social con el resultado de las loterías ordinarias y afirmaba que, como en éstas, el resultado es cuestión de buena o mala *fortuna* o *suerte*: tenemos buena o mala suerte en la lotería natural y social en forma de ventajas o desventajas naturales o sociales, distribuidas de manera moralmente arbitraria.¹

De hecho, cuando evaluamos a alguien por, digamos, su capacidad para realizar determinadas tareas intelectuales, sea en el campo del arte, de la ciencia, etc., no dudamos en reconocer que su mérito está mediado en gran medida por cierto potencial biológico heredado, del que otros no han podido disfrutar, además de por haber hallado los cauces adecuados para desarrollar ese potencial, junto con, por supuesto, su propio esfuerzo y trabajo. O, si pensamos en un gran atleta, al que admiramos por sus habilidades físicas, o por su dominio de las metas particulares

¹ Rawls (1971), pp. 74-5. Como es de prever, la cuestión de la suerte constituye un foco de atención central dentro de toda teoría de la justicia distributiva. En particular, Rawls afirma que “[i]ntuitivamente, la injusticia más obvia del sistema de libertad natural es que permite que las partes que son objeto de la distribución se vean indebidamente influidas por esos factores tan arbitrarios desde el punto de vista moral.” (p. 72.) En esta línea, Richard Arneson ha afirmado que:

La aspiración de la justicia distributiva es compensar a los individuos de los infortunios. Algunas personas son bendecidas por la buena suerte, otras se ven arrastradas por la mala suerte, y es responsabilidad de la sociedad —todos nosotros considerados colectivamente— alterar la distribución de bienes y males surgidos de la confusión de loterías que constituye la vida humana tal y como la conocemos... La justicia distributiva estipula que el afortunado debe transferir al desafortunado algunas o todas sus ganancias que son debidas a la suerte.

Por supuesto, no todos están de acuerdo en esta aspiración igualitarista o, al menos, en qué es lo que debe y no debe redistribuirse. Para contribuciones destacadas a este debate, véase: Cohen (1989), Dworkin (2000) y Roemer (1996), en defensa del *igualitarismo de la suerte*; Nozick (1974), Anderson (1999), Hurley (2001) y Scheffler (2003), presentan importantes críticas.

del deporte que practica, no nos cuesta reconocer que sus cualidades se deben, en gran medida, a una mejor dotación biológica. Aún así seguimos admirando sus extraordinarias destrezas como artistas, científicos o deportistas y considerándoles dignos de elogio. Sin embargo, parece que cuando nos referimos a la cualidad moral de alguien, reconocer que ésta pueda depender, en cualquier medida, de factores heredados es particularmente difícil.² No obstante, parece ser que hay personas que son naturalmente más frías, o que poseen una escasa sensibilidad moral, a las que les resulta más difícil empatizar con el sufrimiento de los demás, o responder a ciertas demandas morales; mientras que otras personas, con una afección moral fuerte, responden sin gran esfuerzo a sus obligaciones morales. Hay también personas con una voluntad más débil que otras, o simplemente con rasgos de carácter que, desde un punto de vista moral, pueden constituir ventajas o desventajas, como ser envidioso o amable, honrado o mezquino, etc.

Entre los posibles elementos de suerte constitutiva, cabe distinguir entre la constitución original más estricta, que meramente incluiría los potenciales biológicos heredados, y quizá el desarrollo más temprano, incluido el embrionario; y una constitución más desarrollada en la que han intervenido factores ambientales, educativos y sociales, que a su vez pueden facilitar o dificultar el desarrollo. En este capítulo me ocuparé principalmente del primer aspecto —del segundo tratará el siguiente capítulo—, si bien entre ellos no hay, ni mucho menos, un límite preciso. En concreto, me ocuparé de la cuestión de la suerte en relación a la constitución estricta, o propiamente dicha (en adelante, simplemente, *suerte cons-*

² Véase Russell (2008), para una interesante comparación entre la moralidad y el arte en relación a la libertad requerida, la responsabilidad/“autoría” y el mérito.

titutiva), esto es, la suerte en la constitución original (biológica) del sujeto. Pero el mismo intento de resolución de esta cuestión nos remitirá inevitablemente al segundo tema, el de la suerte en la formación y desarrollo de la constitución o carácter del sujeto (o *suerte formativa*, en adelante).³ El uso de la expresión suerte constitutiva en este segundo sentido, más amplio, referirá principalmente al carácter o identidad moral del sujeto en un momento dado de su biografía que, además de los rasgos constitutivos estrictos, incluye toda una historia de formación y desarrollo particular en la que han intervenido innumerables factores internos y externos al agente.⁴ Por último, la expresión *suerte antecedente* engloba ambos tipos.

Por otro lado, cabe notar que una cosa es la suerte constitutiva y otra la suerte *moral* constitutiva. Al inicio de la sección I de la *Fundamentación de la metafísica de las costumbres*, Kant afirma, en relación a este tema:

Ni en el mundo, ni, en general, tampoco fuera del mundo, es posible pensar nada que pueda considerarse como bueno sin restricción, a no ser tan sólo una *buenavoluntad*. El entendimiento, la agudeza, el juicio, o como quieran llamarse los *talentos* del espíritu; el coraje, la decisión, la perseverancia en los propósitos, como cualidades del *temperamento*, son sin duda, en muchos respectos, buenos y deseables. Pero también pueden llegar a ser extraordinariamente malos y dañinos si no es buena la voluntad que ha de

³ En realidad, el problema de la suerte constitutiva desborda irremediabilmente el marco de lo meramente *natural* o biológico y afecta también a lo *social*, en tanto que determinante para el desarrollo psicológico. Los límites (diacrónicos) entre suerte constitutiva/suerte formativa/suerte circunstancial son mucho más difusos de lo que podría pensarse, en un primer momento.

⁴ Cabe precisar, en sintonía con la nota anterior, que la suerte formativa también podría considerarse como un tipo de suerte circunstancial, en sentido diacrónico. La suerte moral situacional (si hay tal cosa) tiene que ver con cómo las circunstancias influyen en el juicio moral que merecemos, al afrontar determinadas situaciones, tener ciertas oportunidades, etc. Y estas situaciones y oportunidades son las que también contribuyen a la formación de una identidad moral, al desarrollo del carácter. Es, en todo caso, un tipo de suerte situacional *remota*. Sobre esto, recuérdese lo dicho en § 2.2.2.

hacer uso de estos dones naturales, cuya peculiar constitución se llama por eso *carácter*.⁵

En este párrafo, Kant distingue entre “temperamento” y “carácter”. El primero designaría los rasgos constitutivos, o el material originario, con que el agente se encuentra. Mientras que el segundo se refiere a los rasgos que el agente ha adquirido o desarrollado, a partir del trabajo de la voluntad sobre los primeros. Es por estos últimos, que conforman el carácter, por los que el agente puede y debe responder, y no directamente por su temperamento. Sin embargo, contra esta posición, el defensor de la suerte moral constitutiva afirma que en la medida en que el carácter que el agente se forme o consiga desarrollar, dependa (o esté mediado) por los factores temperamentales previos —que incluyen rasgos básicos de carácter, inclinaciones y *talentos* o capacidades, y que pueden hacer que la respuesta a los requerimientos morales le sea más o menos fácil o difícil—; entonces, aquello por lo que el agente será moralmente evaluado incluirá (o dependerá en parte de) estos factores temperamentales.⁶

La meta de este capítulo y el siguiente consistirá en caracterizar y responder a cómo diferentes autores han rechazado la suerte moral antecedente. Los argumentos pueden agruparse en dos tipos: (i) los que se

⁵ El párrafo continúa así: “Lo mismo sucede con los *dones de la fortuna*. El poder, la riqueza, el honor, la salud misma y el pleno bienestar y el contento con el propio estado, bajo el nombre de *felicidad*, dan coraje, y tras él, a veces arrogancia, si no existe una buena voluntad que rectifique y acomode a un fin universal el influjo de esa felicidad y con él el principio todo de la acción; sin contar con que un espectador razonable e imparcial nunca podrá estar satisfecho al contemplar las ininterrumpidas bienandanzas de un ser que no ostenta el menor rasgo de una voluntad pura y buena, y así parece constituir la buena voluntad la indispensable condición que nos hace dignos de ser felices.” Traducción de García Morente (ligeramente modificada a la luz de determinadas variaciones en las traducciones inglesa y catalana).

⁶ Ver Nagel (1979), pp. 28 y 32-3; también Williams (1981). Aunque desarrolladas independientemente, y en áreas distintas, la noción de suerte constitutiva y la de lotería natural y social de Rawls están, sin duda, estrechamente emparentadas.

dirigen a la noción misma de suerte constitutiva y (ii) los que tratan de neutralizar la suerte antecedente, en general, por medio de la correcta intervención del agente en la configuración de su propia constitución o carácter. Me ocuparé de (i) en este capítulo y de (ii), en el siguiente.

Por un lado, se ha esgrimido, contra la propia noción de suerte constitutiva, que si la existencia misma de alguien depende de la posesión de unas determinadas propiedades esenciales, siendo imposible su existencia sin éstas, no puede decirse con sentido que éste tenga o haya tenido *suerte* en relación a estas propiedades. Además, si el concepto de suerte remite a un resultado azaroso que, en forma de beneficio o perjuicio, recae sobre una entidad previa y fija, resulta que ni hay una entidad previa a la constitución original del agente, ni parece que mi constitución sea en absoluto azarosa para mí. Esta objeción, en tanto que depende de una determinada concepción de qué es la suerte, nos dará la oportunidad de estudiar y clarificar la misma noción de suerte.

Para completar el cuadro, en el siguiente capítulo, me ocuparé de los argumentos que tratan de anular o neutralizar la incidencia de los factores de suerte constitutiva en el juicio moral que uno merezca. La clave será la idea de autoconstitución del propio carácter. No habría ningún problema de suerte constitutiva, si cada persona *se constituye a sí misma*, asimilando o rechazando los elementos previos o heredados y haciéndose cargo de las influencias posteriores. Esta parece ser la posición de Kant en la cita anterior. No obstante, como veremos, esta autoconstitución está sujeta a innumerables factores de suerte formativa que tendrían que ser asimismo neutralizados.

4. 2. ¿Es la idea de suerte constitutiva incoherente?

Algunos pensadores han visto algo extraño en la misma idea de suerte constitutiva, y han intentado mostrar su carácter incoherente. En concreto, se ha argüido que “no se puede decir significativamente que alguien tiene suerte en relación a quien es, sino sólo con respecto a lo que le sucede. La identidad debe preceder a la suerte.”⁷ Esto es, si la misma existencia de P depende de la posesión de unas determinadas propiedades constitutivas (esenciales) *pc*, no puede decirse con sentido que P tiene *suerte* por poseer *pc*, pues la existencia de P depende de que posea *pc*. A favor de esta idea —según la cual, de lo más básico que uno es, no puede decirse con sentido que es fruto de la suerte— caben diferentes argumentos.

En particular, Susan Hurley ha desplegado un argumento que se sustenta en la idea de que la suerte, como lotería, es un tipo de suerte en las consecuencias, lo que tiene implicaciones que no se proyectan lisa y llanamente en la suerte en las causas en general, o en la suerte constitutiva en particular.⁸ Si la idea de suerte remite a un resultado azaroso que recae sobre una identidad previa y constante “externa a esas posibilidades, que podría en cierto sentido reconocerlas y para la que el resultado de la lotería cuenta como bueno o malo”; resulta que no hay nada previo a la misma constitución de un agente. “Sin una identidad que sea constante a través de las alternativas, no tenemos una jugada o lotería sino meramente

⁷ Rescher (1990), p. 155. También Daniel Statman afirma que “la suerte necesariamente presupone la existencia de algún objeto que sea afectado por ella” (1993, p. 12).

⁸ Hurley (2003), cap. 4, § 2. Vimos en el capítulo 2 que Hurley establecía una distinción entre suerte *en las causas* y suerte *en los efectos*, reminiscente de la distinción entre *brute luck* (por completo más allá del control del agente) y *option luck* (suerte originada a partir de las elecciones del agente) de Dworkin; véase la sección 2.2.2 del capítulo segundo. El rechazo de la suerte constitutiva parece ser fundamental para su crítica al igualitarismo de la suerte; ver Hurley (1993) y (2001).

diferentes posibilidades —algunas buenas, otras malas— en que existen diferentes entidades y suceden diferentes cosas.”⁹

Hurley traza un paralelismo entre su posición y la discusión acerca del impacto de las políticas medioambientales y su repercusión en las generaciones futuras.¹⁰ Brevemente, la cuestión es la siguiente: si políticas medioambientales diferentes tienen efectos acumulativos diversos sobre la vida diaria de las personas, resultará que personas diferentes serán engendradas, como resultado de la unión de deferentes gametos, según se lleven a cabo unas políticas concretas u otras. Esto es, supongamos que un gobierno puede poner en marcha la política medioambiental A, la B o la C, siendo significativamente diferentes entre ellas. Supongamos, además, que la política A es, en general, más dañina con el medio ambiente que B o C, aunque puede ser más beneficiosa en otras áreas. Como resultado de los efectos acumulativos sobre la vida diaria de los ciudadanos, hay cierta gente que existirá en el futuro sólo si se pone en marcha la política A, pero no si se implementa la política B o C; como hay cierta gente que sólo existirá si se pone en marcha la política B, pero no la A o la C; y lo mismo con C. Por ello, no puede decirse con sentido que estas personas futuras, que sólo existirán si se pone en marcha A, resultarán medioambientalmente dañadas por A (siendo A medioambientalmente peor que B y que C). Si esta política determinada no se diese, estas personas particulares no existirían. Ciertamente, puede replicarse que, no obstante, sí que tiene sentido decir que la sociedad en su conjunto que resultará dañada (o beneficiada) por la implementación de una política más (o menos) dañina. Sin embargo, lo que está en juego es el daño o beneficio para sujetos particulares, cuya existencia misma depende del resultado. En los casos en

⁹ Hurley (2003), p. 119.

¹⁰ Hurley (2003), p. 118 y 120.

los que las identidades de las personas no son constantes entre los diferentes escenarios posibles, no hay *nadie* que resulte dañado o beneficiado; pues para que hubiese daño o beneficio se requerirían identidades constantes, independientes de los resultados particulares. Por ello, concluye el argumento, hablar de lotería en relación a la constitución que define al individuo mismo no tiene sentido. Hurley insiste en que “la misma idea de lotería va unida a la idea de buena o mala suerte para *alguien*”.

Esta cuestión ha sido tratada por extenso por Derek Parfit como una instancia de su famoso *Problema de la No Identidad*.¹¹ Creo que considerar brevemente esta cuestión puede acabar de aclarar el argumento que trato de exponer. El Problema de la No Identidad, en general, consiste en el hecho de que en algunos casos parece que hacer lo más correcto implicaría que alguien no existiera, mientras que permitir que ese alguien exista supone no hacer lo más correcto; y esto, si es así, nos sitúa ante una situación ciertamente paradójica. Detengámonos en el caso, ideado por Parfit, de una adolescente de catorce años que se queda embarazada y decide tener al niño. Dada su extrema juventud, podemos conceder que, *ceteris paribus*, le dará (al niño) un mal o peor comienzo en la vida. Pero aunque esto tendrá efectos negativos a lo largo de la vida de este niño, podemos también conceder que su vida no dejará de ser, predictiblemente, digna de ser vivida. Sin embargo, si la joven hubiese esperado unos años, predictiblemente también, habría dado a su hijo un mejor comienzo en la vida; pero si hubiese esperado, este niño particular, al que supuestamente da un peor comienzo en la vida, nunca habría existido. Con lo que parece que no podemos intentar persuadirla de que es mejor *para el niño* (para *este* niño particular) que espere a quedarse embarazada con más edad. De hecho, la decisión de la adolescente no fue *peor para su/ese* hijo. ¿Pero,

¹¹ Parfit (1984), cap. 16, § 122.

supone esto que debemos cambiar entonces nuestra valoración de su decisión?

Ciertamente, la cuestión es compleja y desconcertante. Parfit propone la siguiente solución (“incompleta”): si en cualquiera de los resultados posibles viviera el mismo número de gente, sería malo si aquellos que viviesen salieran perdiendo o tuvieran una más baja calidad de vida que aquellos que hubiesen vivido.¹² Aplicado a la cuestión anterior (y dejando a un lado el problema nada baladí del número de personas que vivirán), está claro que la mejor política medioambiental será aquella que permita una vida más beneficiosa a aquellos que de hecho vivan. (Lo que es acorde con la repuesta de sentido común, dado que no podemos controlar el número de gente que vivirá, y no está claro que el que viva más gente sea mejor a que viva menos gente.)

No obstante, es importante no confundir aquí dos cuestiones diferentes; a saber, lo que es bueno o mejor *para alguien* y lo que es bueno o mejor *para el mundo*. Hurley incide en este punto afirmando que, en relación a la aplicación de la idea de lotería a la suerte en las causas, podríamos hablar, con mucho, de una buena o mala suerte *sin más* —suerte para el mundo— en que esto o aquello suceda, pero no buena o mala suerte *para alguien*.¹³ Pero así, el paralelismo anterior no sería correcto. La solución propuesta se mueve en la esfera de lo *mejor para el mundo*, pero esta no es nuestra cuestión; a nosotros nos importa lo *mejor para alguien* —lo mejor para un sujeto determinado *S*. En esta línea, Hurley insiste en

¹² Ciertamente, como remarca el mismo Parfit, esto no constituye una respuesta completa al problema de la no identidad, pues aun queda responder a la verdadera cuestión —la de si resulta que vive diferente número de personas. Por otro lado, y según Parfit, es una consecuencia de esto que si yo fuese el hijo que a sus catorce años tuvo la adolescente, yo *podría aceptar* que sería mejor que el niño que hubiese existido no fuese yo. Ver Parfit (1984), pp. 361ss.

¹³ Hurley (2003), p. 119.

que puede ser cosa de buena suerte (para el mundo) que se unieran los gametos de donde surgió Mozart, pero esto no significa que *Mozart* tuviese suerte por ser Mozart.

Sin embargo, mi intuición es que efectivamente Mozart tuvo suerte de ser un virtuoso de la música, o con más precisión: de tener los rasgos constitutivos que le permitieron ser un virtuoso de la música —lo que podemos llamar sus *innatas dotes musicales*.¹⁴ Pero, para Hurley, decir esto nos comprometería con la idea de una preentidad o prototipo, cuya buena o mala suerte consistiría en tener uno u otro conjunto de propiedades constitutivas o esenciales. La idea de suerte constitutiva, entendida literalmente, requeriría que pudiéramos dar sentido a la idea de un *yo sin constitución*, —lo cual no parece compatible con una metafísica aceptable.¹⁵ Sin duda, esta no es una idea a la que yo desee hacer costado. Recuérdese cómo en el capítulo anterior llegué a la conclusión de que la idea de merecimiento verdadero, llevada hasta sus últimas consecuencias, se tornaba ininteligible a causa del desvanecimiento del mismo agente. La única alternativa disponible, la de considerar que el yo era esencialmente noumenal o, en otras palabras, que la identidad del yo no estaba esencialmente conectada con ninguno de sus rasgos, físicos o mentales, era rechazada precisamente por ininteligible —por lo que no servía para realizar el trabajo requerido: ser el fundamento del merecimiento. Por todo ello, parece curioso que ahora se acuse al defensor de la suerte constitutiva de requerir esta apoyatura.

¹⁴ Estas célebres palabras del propio Wolfgang Amadeus Mozart acerca de sus ideas musicales, suponiendo que son sinceras, parecen apuntar al carácter esencialmente innato de su genialidad: “¿De dónde y cómo me vienen? Ni lo sé, ni tengo nada que ver con ello.”

¹⁵ Hurley (2003), p. 121-2; ver también Hurley (1993), pp. 197-8.

Pero, ¿realmente mantener que Mozart tuvo suerte por sus dotes musicales nos compromete con la idea de un prototipo, de un *proto-Mozart*? Para contestar a esta pregunta tendremos que investigar la misma noción de suerte que está detrás de esta posición.

4. 3. La noción de suerte

Es preciso remarcar que la validez del argumento anterior depende, en primer lugar, de una concepción determinada de la suerte. Pero, ¿cuál es esta concepción de la suerte? Bien, la noción de suerte que hay en juego sigue el modelo del funcionamiento de la lotería o, en general, el de la *suerte como azar*. La siguiente definición de Rescher es una buena caracterización de este modelo:

La suerte... implica tres cosas: 1) alguien que recibe un bien o un mal, 2) un acontecimiento que es benigno o maligno desde la perspectiva de los intereses del individuo afectado y que, más aún, 3) es fortuito (inesperado, azaroso, imprevisible).¹⁶

Según este análisis, un suceso será fortuito *para alguien* cuando, en relación a aquello que tiene que ver con él, el resultado suceda “por accidente”, y este resultado tenga un valor positivo o negativo para la persona.¹⁷

El ingrediente distintivo de este análisis es la inclusión de la probabilidad:

¹⁶ Rescher (1995), p. 35. Véase también Rescher (1990), pp. 145-7. Este modelo hunde sus antiquísimas raíces en la *Física* de Aristóteles (libro 2, 197a-b).

¹⁷ Sin duda, un suceso puede tener un valor positivo y un valor negativo, a la vez, para un sujeto; esto es positivo en un respecto y negativo en otro. Por ejemplo, uno puede ser afortunado de recibir un ascenso, en relación a su carrera profesional, y a la vez desafortunado, en relación a su matrimonio. O digamos que uno tiene un día la mala suerte de perder el avión, pero muy buena si resulta que el avión se estrella. Ver Smilansky (1994), para el estudio de algunos casos de “afortunados infortunios”.

la suerte es una propiedad de los sucesos que varía inversa-mente a la probabilidad del suceso.

Este ingrediente probabilístico podría también formularse en términos modales, de la siguiente manera:

Suerte: Un suceso fortuito es un suceso que se da en el mundo real, pero no se da en los mundos posibles más cercanos al mundo real, en los que las condiciones iniciales relevantes para ese suceso son las mismas que en el mundo real.¹⁸

Si a este elemento probabilístico, añadimos los otros dos factores anteriores —el personal o perspectivo y el valorativo—, podemos avanzar hasta la siguiente caracterización:

*Suerte**: Un suceso *s* es fortuito para un sujeto *S*, cuando *s* acaece en el mundo real, pero no acaece en los mundos posibles más cercanos al mundo real, en los que las condiciones iniciales relevantes para *s* son las mismas que en el mundo real; y *s* tiene valor positivo o negativo en relación a *S*.

Hay que reparar que esta concepción, en tanto que basada en la idea de mundos posibles (esto es, como *toda* concepción en términos de mundos posibles), conlleva un compromiso metafísico con la *rigidez* de las propiedades necesarias de los individuos que deambulan por estos mundos —lo cual es importante para nuestro tema. En particular, la identidad personal vendría definida por la rigidez de los orígenes de cada individuo, en virtud de un principio de este tipo:

¹⁸ Esta formulación es de Duncan Pritchard, en Pritchard (2006), p. 3; ver también Pritchard (2005).

(Principio Z) La identidad de un ser humano, y de otras criaturas con reproducción sexuada, surge de la unión de dos gametos (espermatozoide y óvulo) determinados.¹⁹

Este principio fija las propiedades constitutivas necesarias o esenciales de cada individuo; de modo que si la persona P existe, entonces es necesario que tenga unas determinadas propiedades constitutivas *pc*.²⁰

De este modo, la concepción de la suerte que estamos considerando comporta las dos tesis siguientes:

- (i) la *tesis probabilística* de que todos aquellos sucesos que no puedan ser considerados como azarosos, no pueden constituir casos de suerte, y
- (ii) la *tesis modal* de que nada podrá ser cuestión de suerte (para el individuo) si no hay ningún mundo posible en el que el individuo existe sin ello.

Así, (i) tiene como consecuencia que nada que para el sujeto no sea azaroso, imprevisible o sorpresivo, puede ser cuestión de suerte (para él); y parece que el tener los rasgos constitutivos que uno tiene no es algo que

¹⁹ El nombre de “Principio Z” se debe a Parfit (1984) y es consecuencia de la noción de designador rígido —una expresión que identifica la misma cosa en todos los mundos posibles (en los que esta cosa existe). Sobre esto, véase Kripke (1980). Sin duda, esta concepción de la identidad tiene sus problemas, pero posee una fuerza intuitiva especial, que no se basa tanto en ninguna creencia sobre la importancia de una información genética particular (como causante de los rasgos de una persona), sino en la importancia que tienen los orígenes en nuestra idea de un ser vivo particular. Véase sobre este punto, Williams (1995b), p. 225.

²⁰ Cabe aclarar que hay rasgos constitutivos que pueden no ser esenciales, esto es, rasgos que poseemos desde nuestra constitución pero sin los cuales seguiríamos siendo nosotros mismos (por lo tanto, contingentes, pero en un sentido constitutivos). Haber nacido en Castelló de la Plana es constitutivo para mi identidad, pero no esencial, pues podría haber nacido en otro sitio (siempre respetando el Principio Z). Lo que está aquí en juego son los rasgos que poseemos desde nuestra constitución pero sin los cuales no seríamos quienes somos (los rasgos constitutivos y necesarios). Respecto a las propiedades constitutivas contingentes, no es en principio problemático que podamos hablar de suerte (constitutiva). Cfr. Hurley (2003), pp. 120-1.

pueda considerarse sorprendente para uno mismo.²¹ Pero esto choca frontalmente con ciertas atribuciones cotidianas de suerte, y nos forzaría a dejar de considerar como casos de suerte algunos casos que comúnmente tenemos por tales. Más grave serían las consecuencias de (ii); en particular, no podría decirse que *fue mala suerte para alguien tener los padres que tuvo*.

Ante resultados que chocan con nuestras prácticas e intuiciones cotidianas, hay en general dos alternativas: o nos deshacemos de las intuiciones, pues el análisis demuestra que están equivocadas —de hecho, muchas cosas que nuestro lenguaje ordinario nos permite decir pueden resultar incoherentes, tras la reflexión filosófica— y revisamos nuestras atribuciones de suerte; o nos deshacemos del análisis porque contradice nuestras intuiciones —y son éstas las que, en último término, dan validez o no al análisis propuesto. Para decidirnos por una de estas opciones, tendremos que evaluar primero la corrección del análisis que nos ha llevado a esta situación.

Hay que empezar constatando que la concepción de la suerte como azar se separa de la concepción predominante en el debate. Como vimos, Nagel consideraba que el fenómeno de la suerte tiene que ver con tipos de escenarios en los que “un aspecto significativo de lo que alguien hace depende de factores que están más allá de su control [y] aún así continuamos tratándolo a este respecto como objeto de juicio moral”.²² Por su parte, Williams afirma que “lo que no está en el dominio del yo está fuera de su control, y así está sujeto a la suerte”. Además, afirma que usará una

²¹ Hurley (2003), pp. 126-7, plantea también una objeción a la noción de suerte constitutiva basada en la impredecibilidad de ciertos hándicaps constitutivos.

²² Nagel (1979), p. 25.

“noción generosa” de suerte, “indefinida pero comprensiva”, y añade que el hecho de “que una cosa sea cuestión de suerte no implica que sea incausada”.²³ En la estela de ambos, la gran mayoría de autores que han contribuido al debate han asumido que “suerte” significa meramente *falta de control*.²⁴

La cuestión es qué noción de suerte encaja mejor con nuestros usos cotidianos de “suerte”. Y, de hecho, los defensores de la suerte como azar han considerado que su análisis se adapta mejor a nuestros usos cotidianos, mientras que la noción más general de falta de control parece una concepción *ad hoc*. En efecto, parece que corrientemente no diríamos, por ejemplo, que hemos tenido suerte de que al ir esta mañana a la panadería, el panadero había hecho pan, aunque no depende de nosotros que el panadero hubiera horneado el pan ese día. Parece que, efectivamente, aquí lo que impide que podamos hablar de suerte es la gran probabilidad de que el panadero hornee el pan todos los días en que abre su panadería. Así, hay cosas que no controlamos y de las que no diríamos que son cosa de suerte para nosotros. La falta de control parece que no es suficiente. No obstante, la mera baja probabilidad es igualmente insuficiente. Por ejemplo, en el macabro juego de la ruleta rusa, en el que a pesar de la significativamente menor probabilidad de que al apretar el gatillo del revolver

²³ Williams (1981), p. 20.

²⁴ Estos pasajes dejan constancia de ello: “Empecemos aclarando lo que usualmente queremos decir con el término ‘suerte’. La buena suerte se da cuando algo bueno le sucede a un agente P, y que esto suceda está más allá del control de P. De manera similar, la mala suerte se da cuando algo malo le sucede a un agente P, y que esto suceda está más allá de su control” (Statman 1991, p. 146); “[A]lgo que ocurre como cosa de suerte con respecto a alguien P es algo que ocurre más allá del control de P” (Zimmerman 1987, p. 231); “[D]ecir que algo ocurre como cosa de suerte es precisamente decir que no está bajo mi control” (Greco 1995, p. 83). “‘Es cosa de suerte’ significa aquí: de un modo que está más allá de nuestro control” (Moore 1990, p. 301); “Con ‘suerte’ me refiero a los factores, buenos o malos, más allá del control del agente en cuestión.” (Card 1990, p. 199).

realmente salga disparada una bala (5 a 1), no es extraño decir que se ha tenido suerte de no haber recibido el disparo. Andrew Latus se refiere también a casos en los que uno tiene la habilidad de producir algún suceso valioso, que tienen una baja probabilidad, y en los que el agente consigue producir el suceso, realizando su habilidad; para los que juzga que no parece que uno haya tenido suerte por ello.²⁵ Particularmente, no tengo tan claro que comúnmente no se suela hablar de suerte en relación a estos casos, o que al menos no haya un uso aceptable de suerte aquí. Piénsese en el caso de un baloncestista que anota un triple en el último instante del partido. Por buen anotador que sea, hay un sentido en el que diríamos que ha tenido suerte. Parece que el elemento fundamental aquí para explicar este hecho, sería lo que podemos llamar el carácter contrastivo de la suerte; es decir, que si en casos como el anterior cabe hablar de suerte es porque este resultado contrasta con otros resultados posibles, de un modo que es significativo para el sujeto o aquel que atribuye la suerte —en concreto, anotar el triple en el último instante o no hacerlo.²⁶

Parece, pues, que (i), el carácter probabilístico o azaroso de la suerte podría incluir casos como los anteriores, en los que cotidianamente hablamos de suerte, si vinculamos la probabilidad al contraste particular considerado. En los términos de mundos posibles, el carácter contrastivo de la suerte impondría una restricción en relación a los mundos posibles que cuentan, para el caso, como más cercanos al mundo real. Este resultado apunta, en general, a la idea de que ambos modos de entender la suerte

²⁵ Ver Latus (2003), p. 467.

²⁶ Driver (2006) defiende un análisis contrastivo de la suerte, cuyo factor distintivo es la tesis de que “las atribuciones de suerte (incluso respecto a la misma persona, el mismo conjunto de intereses y las mismas circunstancias, incluso manteniendo todo esto constante) están sujetas a contrastes” (p. 9). Ver Coffman (2007) sobre el elemento de la significatividad.

(como azar o como falta de control) no tienen por qué concebirse como incompatibles.²⁷

Más fundamental, y parece que intratable, resulta (ii). Recordemos que (ii) nos impediría realizar una atribución de suerte como la de que *fue mala suerte para alguien tener los padres que tuvo*. Esto justificaría la tesis de Hurley de que sostener que Mozart tuvo suerte por sus innatas dotes musicales es incoherente, dado que no hay ningún mundo posible en el que existe Mozart y éste no posee esas innatas dotes musicales —a no ser que rechacemos el Principio Z, u otros semejantes, y situemos la rigidez de su identidad en proto-Mozart (cosa que excluimos por principio). Mozart-sin-sus-dotes-musicales no es Mozart, pues ser Mozart incluye necesariamente tener las innatas dotes musicales de Mozart. (Hay que contrastar esto con Mozart-no-siendo-un-compositor-de-éxito, pues la constitución originaria esencial de Mozart no incluye necesariamente *ser un compositor de éxito*; esto es algo que dependerá de circunstancias sociales posteriores.²⁸) Sin embargo, el resultado de (ii) parece contraintuitivo, pues decir que *fue mala suerte para alguien tener los padres que tuvo* es algo que no parece irrazonable decir.

²⁷ Latus (2003) es un intento de avanzar en la dirección de hallar un modelo intermedio. Coffman (2007) ofrece un análisis más fino, construyendo sobre los análisis previos de Rescher, Pritchard y Latus; su esquema es éste:

S tiene suerte con respecto a *E* en *t* si y sólo si (i) *S* es sintiente en *t*, (ii) *E* tiene algún estatus evaluativo objetivo para *S* en *t*, (iii) hubo justo antes de *t* gran probabilidad de que ningún suceso *suficientemente similar y igual en significando* a *E* ocurriera en *t*, y (iv) *E* está más allá del control directo de *S* en *t*.

Por otro lado, Latus (2001, § 3) ha insistido en la idea de que sólo investigando la naturaleza de la suerte podremos alcanzar algún tipo de conclusión final respecto al problema de la suerte moral. Personalmente, no veo muy prometedora esta idea.

²⁸ La diferencia es entre la (im)posibilidad de que *alguien fuese otra persona* y la posibilidad de que *alguien hubiese tenido una vida o historia vital diferente*. En el primer caso tendríamos la historia vital de *un individuo diferente*, que ya no sería yo; y en el segundo, una historia diferente para *el mismo individuo*, esto es, yo mismo. Sobre esto, véase Williams (1973b), p. 45; y (1995b), p. 224.

En el caso de (ii) volvemos al *impasse* anterior, pues de nuevo el análisis choca con nuestras intuiciones, de un modo no decisivo para ninguna de las partes. Si damos un mayor valor a las intuiciones, el resultado apuntar a un límite en el análisis de la suerte en términos de mundos posibles; un límite que sólo se hace evidente cuando se aborda la cuestión misma de los rasgos esenciales de la constitución de un sujeto. La cuestión es que, cuando se trata de la suerte constitutiva (esencial), el recurso al mecanismo de los mundos posibles se vuelve inservible, dado el compromiso metafísico de este mecanismo con el esencialismo. (Este resultado podría ser demasiado fuerte; quiero decir: podría arruinar un análisis de la suerte satisfactorio en otras esferas, o incluso constituir un contraejemplo al método mismo de los mundos posibles.)

A mi modo de ver, el problema no es el esencialismo en sí mismo; esto es, no es incorrecto afirmar que Mozart no pudo existir sin tener las dotes musicales innatas que tuvo (en otras palabras: sus propiedades esenciales, que incluyen sus innatas dotes musicales). Sin embargo, ello no implica que Mozart no tuviese suerte de tener las dotes musicales de tuvo. El error se encuentra, a mi juicio, en presuponer que para atribuir (coherentemente) suerte respecto a cualquier rasgo de una persona es necesario que esa persona pudiera *realmente* no haber tenido ese rasgo.²⁹ Me explico.

Convengamos en la siguiente estipulación entre lo *imaginable* y lo *concebible*. La imaginación (en el uso cualificado que aquí propongo) sería la habilidad para representarse objetos o estados de cosas que no existen aquí y ahora, que no existen en general, o incluso que no pueden existir. Así, uno puede imaginar, y desear, en un sentido, *incluso lo imposible*. Por otro lado, lo concebible depende de la noción de posibilidad,

²⁹ Véase Zimmerman (2002), p. 574; y Latus (2003); para defensas distintas de esta idea.

donde para que algo sea posible —en relación a la identidad personal, que es lo que aquí nos interesa— ha de satisfacer el Principio Z.³⁰ Aplicado a nuestra cuestión, aunque la identidad de una persona, y las posibilidades físicas de su variabilidad, vengan constreñidas por lo concebible, por el contrario la aplicación del predicado “tener (buena o mala) suerte” no conlleva esta restricción, sino que se extiende a todo lo imaginable. De este modo, podemos decir incluso que fue mala suerte para alguien ser de raza negra (durante una época histórica esclavista negra), o ser una mujer (en una sociedad misógina), y ello aunque ser (genéticamente) de raza negra o mujer sean rasgos constitutivos, sin los cuales esas personas no podrían haber existido. Con lo que tampoco es contradictorio creer que estas personas, conscientes de este hecho, pudiesen desear no tener esta cualidad.

Así, si entendemos la noción de suerte en términos de lo imaginable, y no de lo concebible, el paralelismo entre la idea de suerte constitutiva y el Problema de la No Identidad sugerido por Susan Hurley resultará finalmente fallido. (Además, este resultado convertiría, en último término, en insatisfactorio el análisis de la suerte mediante el mecanismo de los mundos posibles. Pero repárese en el hecho de que el análisis funciona bien para todo tipo de suerte excepto para el tipo referido a los rasgos constitutivos esenciales de un agente.³¹)

Sin embargo, es aun posible resistir esta conclusión apelando a la distinción entre *suerte* y *fortuna*. En la sección siguiente, empezaré expo-

³⁰ Reconozco que estas etiquetas son discutibles, pero la distinción es real y significativa. Para una importante recopilación de artículos sobre la noción de *concebibilidad* y su papel en diferentes debates filosóficos, véase Gendler y Hawthorne (2002) —la introducción de los compiladores es especialmente informativa.

³¹ Retomo esta cuestión en las consideraciones finales de la siguiente sección.

niendo en qué consiste esta distinción, para pasar luego a criticarla. A continuación, trataré de explicitar lo que, a mi entender, hay de realmente importante tras esta distinción.

4. 4. Suerte y fortuna, o la especial relevancia moral del carácter

Un modo de hacer plausible un resultado como el de (ii) para los defensores del análisis anterior —plasmado en el esquema *Suerte**— es proponiendo la siguiente distinción entre *suerte* y *fortuna*. Por ejemplo, Julia Driver, tras ofrecer una versión de este análisis, afirma:

El carácter de una persona puede al menos en parte ser debido a sus padres, a pesar de que no haya ningún mundo posible en el que hubiese tenido otros padres. En esta concepción de la suerte, pues, mucho de lo que la gente llama suerte moral es de hecho *fortuna* moral. (...) Una manera en la que “carecemos de control” es por accidente. Otra manera es por simple falta de elección. Ambas pueden coincidir o no.³²

La idea es que una cosa es tener suerte, contando como suerte sólo aquello que satisface la caracterización anterior; y otra tener fortuna, que incluiría cosas como haber nacido con unos determinados talentos y ventajas. De este modo, el análisis anterior queda limitado a la noción (restringida) de suerte, dejando a parte la cuestión de la fortuna. Lo que haría singulares a nuestras disposiciones, rasgos de carácter e inclinaciones heredados es que son cosas que por naturaleza no se eligen, que constituyen a una persona como la persona que es, que conforman “el destino del sujeto”. Abundando en esta línea, Rescher añade:

³² Driver (2006), p. 21.

Su disposición y talento forman parte de aquello que la constituye [a esa persona] en el individuo que es; no es algo que el azar añade a una identidad preexistente. Pero la posesión de ese talento es una cuestión de fortuna y no de buena suerte. No es como si hubiera una versión de nosotros mismos, externa al mundo y previa a la fertilización, que tuvo la suerte de obtener una asignación favorable.³³

No negaré que en esta distinción hay algo significativo, que descansa en las mismas intuiciones que apoyaban la distinción de Kant entre “temperamento” y “carácter” que vimos al principio de este capítulo. Sin embargo, esta distinción difícilmente coincide con la distinción entre propiedades esenciales y contingentes, que es lo que creaba las dificultades que ponían en entredicho el análisis anterior. Esto será más obvio tras leer las siguientes palabras del mismo Rescher:

Las cosas positivas y negativas con que nos topamos en el acontecer normal —incluida nuestra herencia (biológica, médica, social, económica), nuestras aptitudes y talentos, las circunstancias de nuestro tiempo y lugar (la paz o el caos, por ejemplo)— dependen de lo que podría caracterizarse como destino y fortuna. La gente no tiene mala suerte por haber nacido tímida o malhumorada, sólo es desafortunada. Pero los elementos positivos y negativos que nos encontramos por azar y en circunstancias imprevistas —encontrar la cueva del tesoro, salir indemne de un accidente que resulta fatal para la mayoría de los afectados— son cuestión de suerte.³⁴

Además, negar la suerte constitutiva en virtud de la mera distinción entre “suerte” y “fortuna” no parece muy prometedor. ¿Por qué no es igualmente “problemática” la *fortuna*, que la *suerte* (constitutiva)? Si el fenómeno es el mismo, ¿qué importa como lo llamemos? En realidad, la conclusión del análisis anterior no puede ser otra que el hecho de que hay

³³ Rescher (1995), p. 43.

³⁴ Rescher (1995), p. 41.

un tipo de suerte moral —la *fortuna* o suerte constitutiva— que no hay más remedio que aceptar. Lo cual, más que establecer la incoherencia de la noción de suerte constitutiva, le acaba otorgando un respaldo teórico —aunque sea convirtiéndola en una categoría *sui generis* dentro de los tipos de suerte.

No obstante, para convencer a alguien de que debe aceptar la distinción entre *suerte* y *fortuna* se requiere de un argumento —más allá de vagas intuiciones, que además no coinciden con la distinción entre propiedades esenciales y contingentes que el análisis anterior requería. (La misma estipulación de los dos términos es un resultado forzado que violenta nuestro uso del lenguaje, pues comúnmente tanto puede decirse que “S fue afortunado de tener los padres que tuvo” como que “S tuvo suerte de tener los padres que tuvo”.³⁵)

En realidad, lo fundamental aquí, lo que está debajo de esas vagas intuiciones aludidas, es la diversa influencia que la suerte puede tener en la identidad, carácter y vida de una persona. En esta línea, la distinción entre suerte y fortuna, más que una distinción con fundamentos metafísicos, podría interpretarse como un compromiso para con las ideas de agencia y responsabilidad moral. Rescher dice:

Los atributos morales de una persona no le llegan por suerte sino que surgen de ella en cuanto agente moral libre. Responsabilizar a las personas por su carácter moral... forma parte del supuesto moral fundamental que adoptamos al tratar a una persona como persona. Verlo como una adición extra que la suerte puede poner o no

³⁵ Puede que en inglés la distinción parezca más plausible. No obstante, como en castellano, igual puede decirse “*I am fortunate to have the parents I have*” que “*I am lucky to have the parents I have*” —y “*I was fortunate to have won the lottery*” como “*I was lucky to have won the lottery*”.

en nuestro camino equivale a dejar de tratar a las personas como tales.³⁶

Por su parte, Dana Nelkin ha sugerido que la mejor manera de atrapar la idea de que hay algo especial respecto a la relación entre suerte y constitución podría ser mantener que *la suerte moral constitutiva simplemente no es problemática para la moralidad de la manera que lo es la suerte moral resultante*.³⁷ Quizá, para los propósitos de la evaluación moral, no importe cómo llegas a ser quién eres; sino que lo que importa es qué haces con lo que eres. Esta estrategia puede ser vista con simpatía por aquellos que desean poner freno al que puede convertirse en un regreso al infinito del control y la responsabilidad moral.³⁸ No obstante, sucede que lo que uno hace con lo que es también depende de lo que es, y abandonar la preocupación por lo que se es resultará insatisfactorio para aquellos que piensan en juicios de responsabilidad moral profundos. Esta consideración, si bien aparentemente trivial, es fundamental. Puede que lo que uno hace importe más que lo que es, y que importe más lo que se es que cómo se llegó a ello, pero también importa cómo uno llegó a ser quién es en relación a lo que es, y lo que es en relación a lo que hace. Intentaré explicar algo esto.

La distinción entre estratos en la propia agencia en relación al control y a la responsabilidad moral me parece significativa en este sentido: apunta a algo importante en nuestra propia autocomprensión y en la comprensión de los demás; a algo *moralmente* importante. Pues parece que el carácter o identidad moral define a la persona como la persona que es, y no hay más remedio que asumir la responsabilidad por la propia identidad

³⁶ Rescher (1995), p. 169; y Rescher (1990), pp. 155-6.

³⁷ Nelkin (2004).

³⁸ Esta es la motivación confesa de Hurley. Algo parecido sucede en el caso de Rescher.

o carácter. Rescher incide en la idea de que aunque no seamos responsables del carácter que heredamos, sí que lo somos por tenerlo actualmente.³⁹ O, en otras palabras, en nuestra experiencia diaria no nos resulta igualmente problemática la suerte en las causas que la suerte en las consecuencias, por usar la distinción de Hurley. La falta de control respecto a nuestra constitución conduce a cuestiones de diferente tipo a las que comporta la falta de control en nuestras decisiones o acciones, y en sus consecuencias.

No obstante, la mera constatación de este hecho resultará muy poco satisfactoria a aquellos que se preocupen por el control del proceso de formación y desarrollo del propio carácter por parte del agente. Simplemente decir que el carácter es el locus último de la responsabilidad, sin decir nada sobre nuestra adquisición y control del carácter, es claramente insuficiente. Por ejemplo, Richards reconoce que el problema de la determinación última permanece abierto. “[S]i el individuo contribuye en alguna medida a ser la clase de persona que es, esta contribución puede ser la base para que merezca el elogio o censura por lo que es.”⁴⁰ No obstante, según la definición de Nagel, parece más bien que, en realidad, para que el rechazo de la suerte moral antecedente tenga éxito, el carácter de una persona ha de ser (re)creación propia. Esta (re)creación parece tener sus límites. Pues al mismo tiempo que, en ocasiones, desearíamos ser de otra manera, o no tener un rasgo de carácter (que a pesar de nuestra lucha, sigue presente en nosotros), también sospechamos que puede que este

³⁹ E insiste también en basar esta distinción en el argumento o explicación anterior: “No somos moralmente responsables de escoger nuestro mal carácter (el carácter no es algo que se elija), pero somos moralmente responsables —y moralmente reprobables— por tenerlo. Y aquí no interviene la suerte. No puede decirse propiamente que tengo suerte respecto de quién soy, sino sólo respecto de lo que me acontece. La identidad debe preceder a la suerte. [...] La identidad no se asigna por medio de una lotería a individuos igualmente desnudos.” (Rescher 1995, p. 169-170)

⁴⁰ Richards (1986), p. 172

modo del que somos o este rasgo de tenemos forme parte de nuestra misma constitución.

Así, aunque la distinción entre suerte y fortuna resulte artificial para salvar el análisis anterior de la suerte, sí que es reflejo de importantes diferencias en relación al grado de control que el agente puede ejercer sobre el conjunto de todos aquello que conforma su identidad moral. En concreto, lo que hay en juego es la cuestión de las posibilidades efectivas de autoformación. De hecho, esta compleja cuestión de la auto(re)creación será la que centre mi interés en el capítulo siguiente. Pero antes de pasar a ello, hay algunas consideraciones que quiero plantear.

4.5. Posibilidades formativas e indeterminación

En primer lugar, por lo que concierne a la suerte moral, existe un modo de rescatar el análisis anterior de la suerte en términos de mundos posibles. Hay un sentido, que es el relevante para nuestro tema, en el que podemos seguir diciendo correctamente que *fue mala suerte para alguien tener los padres que tuvo*, aun aceptando la idea de que las atribuciones de suerte no pueden violar la constricción de lo *concebible* —de las propiedades esenciales reales del agente. Esto es, cuando en el contexto de la suerte moral se dice, por ejemplo, que fue mala suerte para los hijos de Himmler tener a éste como padre, la cuestión no es en realidad que tuvieron mala suerte de surgir de cigotos formados por gametos de Himmler y su esposa; sino que la mala suerte consiste, en todo caso, en haber nacido en una familia de fanáticos seguidores del nazismo, en la que era muy probable heredar ciertas creencias morales fuertemente reprobables. De esta manera, es posible que los hijos de Himmler siguieran siendo hijos de Himmler (se desarrollasen a partir de cigotos engendrados por los Himm-

ler), sin estar sometidos a semejante influencia moral perniciosa, debido a que la historia vital del mismo Himmler podría haber sido muy distinta — por ejemplo, porque nunca hubiese sido seducido por las ideas nazis. O pensemos en el caso de una pareja de personas que transmiten su extraordinaria sensibilidad moral a sus hijos. En comparación con los hijos de familias menos sensibles, éstos *tienen suerte de tener los padres que tienen*. Pero este hecho no es incompatible con la constricción impuesta por el análisis de la suerte que venimos considerando, si caemos en la cuenta de que estos padres, siendo las mismas personas, podrían haber sido menos (o más) sensibles moralmente, si su historia vital previa hubiera sido otra. Es decir, aun manteniendo fijas las propiedades esenciales de los sujetos en cuestión (satisfaciendo el Principio Z), sus propiedades morales *fenotípicas*, por decirlo así, podrían haber sido muy diferentes; lo que se adecua al análisis anterior de la suerte.⁴¹ Así, aunque esto no resuelve la cuestión de las propiedades *esenciales* heredadas, sí que es suficiente para satisfacer las necesidades de nuestro tema. Con ello, además de rescatar el análisis anterior, puedo reforzar mi réplica a la acusación de que es incoherente hablar de suerte constitutiva.

Una segunda consideración, en la que quiero explayarme, tiene que ver con la maleabilidad de los propios rasgos. La constricción impuesta por los límites anteriores a lo *concebible* respecto a la identidad de una persona, aunque no desempeñe el papel definitivo contra la noción de suerte constitutiva que pretendían sus defensores, sí que resulta relevante en relación a las posibilidades formativas. Pues, por mucho que podamos imaginarnos sin determinados rasgos, o desear no tenerlos, existe una dramática diferencia entre aquellos rasgos que juzgamos negativos y respecto a los que efectivamente podemos resistirnos —esto es, que no es en

⁴¹ Agradezco a Carlos Moya que me sugiriese y animase a perseguir esta línea de réplica.

principio imposible que llegáramos a deshacernos de ellos o, al menos, es concebible que no los hubiéramos desarrollado— y aquellos que, dada nuestra constitución original, es físicamente imposible existir sin ellos — por bien que la indeterminación epistémica oscurezca la distinción, impidiéndonos en muchos casos saber si realmente teníamos alguna posibilidad de modificar cierto rasgo. En concreto, hay un sentido en el que es inútil que una persona intente resistirse a ciertos de sus rasgos físicos, que no puede modificar, como no lo es que se esfuerce modificar ciertas de sus tendencias o actitudes. Imaginemos el siguiente caso:

Ana es una adolescente con una baja autoconfianza somática — menor de lo usual—, que a menudo la conduce a comportarse de un modo inseguro, limitando sus deseos, evitando confrontaciones con otros, incluso las más leves, etc. Hasta cierto punto, se siente alienada o limitada en el autocontrol de su conducta, e incluso en ocasiones se siente avergonzada de sí misma por actos que en realidad no merecen esa vergüenza. Esta disposición es algo que de niña ya mostraba, algo con lo que se encontró, y que la ha llevado a adquirir ciertos hábitos que le han impedido desarrollar ciertas habilidades positivas en el trato social que desearía haber desarrollado. Así, la introversión se ha convertido en un rasgo temperamental central de su personalidad, incrementándose su baja autoconfianza somática.

La relevancia de esta historia para el juicio moral es fácilmente apreciable si consideramos la siguiente continuación del relato:

Un día, Ana, con más de treinta años, desafortunadamente atropella a un peatón y, a causa de los rasgos psíquicos anteriores, en lugar de pararse a auxiliar a la persona herida, acelera y se da a la fuga.

Con esta historia pretendo hacer notar cómo las propiedades, aún psicológicas, que uno desarrolla pueden estar profundamente marcadas por los rasgos temperamentales recibidos. Sin embargo, tiene perfecto sentido que Ana se esfuerce por superar esta situación, si es lo que desea; aunque puede que, a pesar de todos sus esfuerzos, no pueda modificar este rasgo.⁴²

Podemos hablar de la maleabilidad o posibilidades de realización o modificación de las mismas propiedades constitutivas esenciales; es decir, que no es *inconcebible* llegar a resistir algunas de estas propiedades, *en tanto que rasgos efectivos*.⁴³ Los límites de lo coherentemente concebible respecto a las propiedades de una persona pueden variar, aun manteniendo fijada la constricción física (genética) del origen; pues en muchos casos (especialmente en los psicológica y moralmente relevantes) ésta no determina los rasgos fenotípicos, que son los que influyen en nuestra vida, y los que realmente afectan a la cuestión de la suerte moral. La variabilidad cuenta con los cambios genéticos ocurridos antes y después de la concepción.⁴⁴ Puede haber igualmente variación durante el período de gestación, dadas las influencias positivas y negativas (benéficas o dañinas) que el feto puede sufrir; así como a causa de la interacción con el ambiente familiar particular durante los primeros meses de vida, y con

⁴² Por supuesto, el problema de la suerte constitutiva es mucho más amplio, y no se circunscribe al ámbito de la moralidad o responsabilidad moral; sino que tiene que ver con la constitución de que cada cual está dotado, incluyendo características físicas, mentales, etc.

⁴³ En el lenguaje de los mundos posibles: existe una pluralidad de mundos posibles en los que *S* puede concebirse coherentemente como afortunado o desafortunado según posea o carezca de la propiedad *P*, o cuánto disfrute de esta propiedad, manteniendo fijo su origen zigótico.

⁴⁴ Desde el punto de vista origenista, hay un tiempo variable, aunque limitado, de concepción, compatible con el mantenimiento de la identidad numérica del cigoto. Cf. Parfit (1984), pp. 352-3. El Principio Z plantearía un problema con los casos de gemelos monozigóticos, esto es, de dos personas surgidas de la misma célula, pero esto no debe preocuparnos aquí.

un entorno social más amplio con posterioridad, en el que el niño crezca.⁴⁵

Así, el ámbito de lo concebible con respecto a la identidad o carácter subsiguiente del agente, a partir de unos mismos rasgos constitutivos, es más amplio de lo que Hurley y Rescher parecen reconocer. Sin embargo, tiene sus límites: por seguir con dos ejemplos anteriores, ser mujer o negro viene fijado por el origen mismo (es *inconcebible* que alguien mujer o negro no fuese mujer o negro), y puede constituir un elemento de desventaja en ciertas sociedades; lo que nos debe hacer reparar también en el hecho de que el entorno social, educativo e histórico juega un papel no sólo en el hecho de favorecer o dificultar la formación de inclinaciones y rasgos determinados, sino también en la misma definición de cuáles comportan ventajas y cuáles desventajas. Por ejemplo, ser reservado o desinhibido puede resultar ventajoso o desventajoso según nos encontremos en una sociedad victoriana o en una sociedad más liberal.⁴⁶

Además, las diferencias en los rasgos constitutivos pueden posteriormente influir en nuestras acciones de modos muy distintos e inesperados. Por ejemplo, comparemos a los dos asesinos que protagonizan *A sangre fría*, la novela de Truman Capote. Perry Smith, el hombre que em-

⁴⁵ A mi modo de ver, existe cierto contraste entre los rasgos naturales y sociales, pero no una dicotomía sustantiva. Los rasgos naturales constituyen la materia prima sobre la que nuestros rasgos determinados parcialmente se construyen. Pero, los rasgos naturales son normalmente componentes de rasgos socialmente contruidos y, además, el desarrollo de los rasgos naturales está condicionado por factores sociales. En este sentido, una cuestión como la de si un rasgo es natural o social no es muy significativa —más bien, dependerá en parte de cómo se describe el rasgo y en qué aspectos de su historia causal estemos interesados. De este modo, la cuestión tradicional de la relación entre naturaleza y educación (de cuánto depende de los genes y cuánto del entorno), por importante que pueda ser, no afecta directamente a nuestro tema. Véase Flanagan (1991, p. 41-6) para la defensa de una posición similar.

⁴⁶ Ver Rorty y Wong (1993) para una presentación general de los distintos aspectos psicológicos de la identidad y la agencia, así como de sus dimensiones social y cultural, y la defensa de la importancia de su consideración para el estudio filosófico de la identidad personal. Me ocuparé más a fondo de estas dimensiones al tratar de la suerte formativa.

puñó la navaja y apretó el gatillo en los asesinatos por los que él y Hickock fueron ejecutados, es descrito por Capote como un hombre con una muy débil conexión con la realidad, dado a la ensoñación y a la ilusión, y con gran predisposición a la rabia explosiva, factores que al combinarse encienden su instinto asesino. El salvajismo de su conducta se vuelve aun más extraordinario a causa de su interés previo por el confort y la seguridad de sus víctimas. Durante el crimen y los sucesos siguientes, Smith exhibe su lado blando y su capacidad para la simpatía y para conectar con los demás. Estas disposiciones están ausentes en el carácter de Hickock, que Capote describe como un hombre excepcionalmente superficial, siempre intrigante, dominado por emociones triviales cuando no consigue sus objetivos e incapaz de sentimientos más profundos acerca de sí mismo y de los demás. En particular, no siente ninguna compunción por sus actos, por muy malos o dañinos que puedan ser, ni muestra vergüenza, arrepentimiento o remordimiento por ellos. El crimen es su idea, y se hace amigo de Smith porque ve en su capacidad para la violencia algo muy útil para llevar a cabo sus ambiciones criminales. Por otro lado, no es alguien especialmente inclinado a la violencia, al margen de sus bravatas; y, a diferencia de Smith, no tiene problemas en mantener su cabeza fría y su mente lista en la búsqueda de sus fines malvados.⁴⁷ Obviamente, los rasgos de carácter de ambos no se deben exclusivamente a sus constituciones originales; no hay duda de que, además, ha sido necesaria toda una histo-

⁴⁷ Tomo la descripción de los rasgos de ambos de Deigh (1995). Un contraste semejante puede apreciarse entre los dirigentes nazis Rudolf Höss y Adolf Eichmann, ambos asesinos criminales, pero de temperamentos muy diferentes: el primero, un hombre de acción y el segundo, un hombre de ideas. Hanna Arendt afirma que “si le hubieran nombrado comandante de un campo de exterminio, como le ocurrió a su buen amigo Höss, Eichmann se hubiera suicidado porque se consideraba incapaz de matar.” Entiéndase: de matar con sus propias manos, pero no de idear y poner en marcha la Solución Final. No obstante, era improbable que le hubieran encomendado una tarea de aquella clase, pues sus superiores “conocían muy bien los límites de cada individuo”. Véase Arendt (2006), pp. 136-7.

ria causal particular para convertirlos en actuales. Lo único sobre lo que pretendo llamar la atención con este caso es sobre cómo de intrincada puede ser la conexión entre la posesión y desarrollo de ciertas disposiciones, su rol en la economía global del carácter de una persona y las circunstancias en las que uno se ve inmerso. De estas cuestiones trata el siguiente capítulo.

Dónde estamos y adónde vamos

Cabe concluir que resulta innegable que la noción de suerte constitutiva, que no presupone que el agente podría haber tenido realmente una constitución original diferente a la que de hecho tiene, no es ni mucho menos incoherente; por lo que la suerte constitutiva es un hecho. Además, el carácter de una persona, en tanto que definitorio de su identidad moral personal, ocupa un lugar especialmente importante en nuestra consideración moral de los demás, si bien esto no excluye la relevancia de la cuestión de la relación del agente con su propio carácter y su grado de control sobre él. En todo caso, la diferencia entre lo concebible en relación a la identidad de una persona y lo que puede ser cosa de suerte para ella pone en claro que muchos elementos de la propia identidad personal, e incluso de la agencia, están necesariamente más allá del poder del agente.

En el siguiente capítulo indagaré la cuestión del poder del agente para formar y reformar su propia identidad moral y carácter y de los avatares de la historia formativa particular que lleva a cada persona a forjarse, en mayor o menor medida, a sí misma; lo cual podría servir para alcanzar la tan ansiada neutralización —por parte del adversario de la suerte moral— de las desigualdades constitutivas.

5. Suerte formativa

Responsabilidad por el carácter

- 5.1. Autocontrol reflexivo
 - 5.1.1. Reflexividad volitiva
 - 5.1.2. Reflexividad racional
 - 5.1.3. Dos proyectos
- 5.2. Competencia normativa y condiciones sociales
 - 5.2.1. Apremiar lo Verdadero y lo Bueno
 - 5.2.2. Afectividad y relaciones interpersonales en la formación del carácter
- 5.3. Autocreación. Decisiones y acciones autoformativas
 - 5.3.1. ¿Autocreación?
 - 5.3.2. Decisiones y acciones autoformativas
 - 5.3.3. La “decisión” de Raskólnikov
 - 5.3.4. La repercusión futura de las decisiones autoformativas
- 5.4. El rastreo de la responsabilidad y la condición epistémica
- 5.5. Carácter y evaluación. Conclusiones

En tales trances de vacilación y duda, sólo la experiencia y el ejemplo pueden determinar razonablemente el impulso del corazón. Mas la experiencia no se adquiere a voluntad; depende de la situación en que a cada uno le coloque la suerte.

Abate Prevost, *Manon Lescaut*, Advertencia del Autor.

El resultado del capítulo anterior nos dejó prácticamente como estábamos al principio del mismo en cuanto a la posibilidad de deshacernos de la suerte constitutiva. Uno podría todavía aceptar que, en efecto, poseemos rasgos constitutivos diversos, que *nos tocan*, esto es, que no depende de nosotros haberlos recibido, pero que lo que juzgamos no son estos rasgos *brutos*, por así decir, sino cómo los afrontamos, modificamos o aceptamos

como elementos de pleno derecho de nuestro carácter —del carácter que nos formamos. En otras palabras, no habrá ningún problema de suerte moral con respecto a los rasgos constitutivos, si somos nosotros quienes, haciéndonos cargo de ellos, modificándolos o rechazándolos, nos construimos a nosotros mismos, o elegimos nuestros propios rasgos. Así, si tenemos suerte (o fortuna) en nuestra constitución, pero ésta no repercute en el juicio moral que merecemos, en último término no habrá suerte *moral* constitutiva. Hay que recordar que la mera suerte, no es suerte moral; de hecho, los mismos adversarios a la suerte moral reconocen que la suerte está presente en nuestras vidas, incluso que es una de “las características sobresalientes de la condición humana”.¹

Este tipo de respuesta, que aspira a neutralizar la suerte constitutiva estricta mediante el control en la configuración del propio carácter, puede concretarse de muy diversas maneras. La más inmediata consistiría en que el agente disfrutara siempre del poder para recrear su propio carácter, a partir de los diferentes elementos con que se ha encontrado. De hecho, esta parece ser la posición de Kant, para quien cualquier persona puede volverse virtuosa, en cualquier momento de su vida, sin importar ni su constitución, ni su formación y experiencias, con independencia de cualesquiera influencias contrarias a que haya estado sujeto o de las veces que en el pasado haya actuado viciosamente. Para él, lo único que se requiere es el *esfuerzo de la voluntad*. La virtud —esto es, el dejar que la buena voluntad dirija nuestras acciones— es siempre posible, en tanto que resultado de una *revolución* instantánea de la voluntad.²

¹ Rescher (1995), p. 80.

² Para Kant, la virtud consiste en el compromiso firme de actuar por el motivo del propio deber, en ausencia de ninguna inclinación o contra nuestras mismas inclinaciones. Sobre la virtud en Kant y su relación con la suerte moral, véase Athanaloussis (2005), cap. 5.

En este capítulo, exploraré esta cuestión echando mano de las más destacadas teorías filosóficas acerca de la responsabilidad moral. En concreto, se trata de teorías que aspiran a establecer las condiciones suficientes y necesarias para que un agente pueda ser satisfaga lo que en el capítulo primero llamamos AMR, es decir, pueda ser considerado un agente moralmente responsable. Mi propósito último no será evaluar directamente estas teorías; es decir, dictaminar si las condiciones que proponen para que un agente pueda ser considerado como moralmente responsable, son suficientes (o incluso necesarias). Más bien, lo que me interesa es evaluar si, en el caso en que lo fuesen, podrían bloquear el fenómeno de la suerte moral. Puedo avanzar ya que mi posición será la de que ninguna de estas teorías puede hacerlo. Por comodidad, llamaré al agente que cumple con las condiciones que cada teoría propone agente *autónomo*.³

5. 1. Autocontrol reflexivo

5.1.1. Reflexividad volitiva

Un lugar interesante donde comenzar es en las disquisiciones de Harry Frankfurt al respecto. Frankfurt ha afirmado explícitamente que una persona es moralmente responsable por las elecciones y conducta a que le

³ Lo que nos interesa aquí no son los requisitos para atribuciones particulares de responsabilidad moral, sino las condiciones generales para ser un agente *autónomo* o, en otras palabras, candidato adecuado para las atribuciones de responsabilidad moral — “autónomo” significará meramente eso. Pues aún siendo un candidato adecuado para estas atribuciones, pueden existir excusas que en una situación determinada excluyan la atribución de responsabilidad moral; asimismo, el mero no ejercicio de las capacidades necesarias para ser un agente autónomo no impide la atribución.

Por otro lado, mis argumentos pretenden ser igualmente válidos para compatibilistas e incompatibilistas, a la vez que trataré de evitar la cuestión metafísica de la responsabilidad moral última y del libre albedrío. (Mi postura oficial en esta investigación será el agnosticismo.) A mi modo de ver, esto tiene un paralelo claro en epistemología: no toda investigación acerca de las condiciones para el conocimiento deben abordar necesariamente el problema del escepticismo epistémico.

conduce el carácter que actualmente tiene, más allá de su control sobre los factores antecedentes (intra y extrapersonales) que le llevaron a tener este carácter. Estos factores sobre los que el agente ya no tiene control no son importantes para su autonomía actual, en tanto que éste es capaz de asumir la responsabilidad por ellos, de evaluarlos reflexivamente y aprobarlos o desaprobarlos como parte de su yo. Lo fundamental es “si las disposiciones que [el agente] posee [...] son características con las que se identifica y que incorpora en sí por su propia voluntad como constitutivas de quien es.”⁴ Y ello por más que, después de todo, estemos inevitablemente formados y sometidos por circunstancias sobre las que no tenemos control. “Las causas a las que estamos sujetos pueden también cambiarnos radicalmente, sin que ello implique que no somos agentes moralmente responsables.”⁵

En cierto sentido, esta posición se asemeja a la de Aristóteles, en tanto que incluso si el punto de vista fundamental de alguien está fijado por factores sobre los que uno no puede ejercer control, esto no hace que el propio carácter deje de estar bajo el control de uno mismo, porque pueden aún realizarse cambios en él.⁶ Incluso si la conducta del agente está últimamente causada por factores externos, puede haber una contribución

⁴ Frankfurt (1988), pp. 171-2. En realidad, Frankfurt no se ocupa directamente de la cuestión de la suerte moral (constitutiva o formativa), sino que ha sido principalmente Daniel Statman quien lo ha interpretado de esta manera. Véase Statman (1993), pp. 12-3. Su sugerencia es que las teorías de la responsabilidad moral basadas en la idea de identificación reflexiva pueden dar solución al problema de la suerte moral constitutiva, tal y como he sugerido al principio de esta sección. Sin embargo, hay otros tipos de teorías que, en principio, son igualmente candidatas a dar respuesta a la cuestión, y que cabrá considerar.

⁵ Frankfurt (2002a). Esta cita forma parte de la respuesta de Frankfurt a objeciones externistas, como la planteada por los casos de manipulación por medio de estimulación electrónica directa del cerebro, o por los llamados casos de Mundo Feliz. Sin comprometerme en absoluto acerca de este tipo de objeciones o contraejemplos, resultará evidente que la postura de Frankfurt no parece ser la de negar tal cosa como la suerte constitutiva.

⁶ Aristóteles, EN, libro III, epígrafe 5.

personal al resultado, entendida en términos de la identificación con los resultados de la acción, suficiente para hacer la conducta atribuible al agente.

La teoría jerárquica de la autonomía de Frankfurt apela a la distinción entre deseos y voliciones de primer y segundo orden, o de orden inferior y superior. Aquello que define a las personas es su capacidad para autoevaluarse, esto es, el poder de reflexionar sobre sus propios deseos y voliciones y permitir o no su permanencia, o perseguir su reajuste. Pensemos en nuestro ejemplo del final del capítulo anterior. Ana era una persona con una baja autoconfianza somática, que se sentía alienada o desposeída de su poder, lo cual la llevaba a actuar de modo tentativo y ansioso, a limitar sus deseos, a evitar confrontaciones, etc., cosas todas ellas que no deseaba realmente hacer. Esto constituye, en expresión del propio Frankfurt, un tipo de *división interna* del sujeto. Específicamente, nos encontramos ante un conflicto entre cómo uno quiere estar motivado y el deseo que de hecho le mueve a actuar. La solución de Frankfurt consiste en apelar a la noción de *identificación*. Una persona se identifica propiamente con un deseo de orden inferior cuando tiene, *sin oposición*, una volición para actuar de orden superior, con la que concuerda el deseo de orden inferior, y juzga que ulteriores deliberaciones, relacionadas con deseos de orden superior acerca de la cuestión, no influirían en su resolución.⁷ En todo caso, lo fundamental es la actitud del agente hacia su propia voluntad (técnicamente: los deseos *efectivos* de primer orden); la perfecta integración entre el agente y su voluntad. Cuando la voluntad de un

⁷ Son relevantes para estos puntos Frankfurt (1971) y (1987). Ver también otros ensayos en la antología Frankfurt (1988).

agente muestra esta configuración, el agente revela su yo auténtico en ella, sosteniéndola incondicionalmente.⁸

En este sentido, Ana tendría que solucionar el problema de su baja autoconfianza somática mediante un proceso semejante de autorreflexión, conducente a algún tipo de reajuste de sus estados mentales. Sin embargo, esto no es tan sencillo, pues dejar de tener una baja autoconfianza somática no es algo que dependa directamente de su voluntad. Así, puede que, tras el proceso de reflexión y lucha contra este rasgo, quizá llegue a eliminarlo, si bien también es posible que finalmente no tenga más remedio que resignarse a él.⁹

⁸ Frankfurt (1994). Frankfurt ha desarrollado sucesivamente diferentes nociones de *identificación* para intentar superar la objeción de regresión, que diferentes autores le han planteado; en especial, Watson (1975): para que un agente sea moralmente responsable se requiere la conformidad de su voluntad con una volición de segundo orden, que a su vez ha de ser libremente querida; pero sucede que ello requiere de la conformidad con alguna volición de orden más superior, etc., etc.

⁹ David Zimmerman ha defendido que este tipo de solución —consistente en afirmar que, al final, puede que a la persona no le quede más remedio que *asumir la responsabilidad* por los deseos que de hecho posee, llegando simplemente a identificarse incondicional y decisivamente con ellos, al margen de su origen— es una solución que compromete a Frankfurt con la “repugnante idea estoica” de que “la resignación ante la necesidad es un camino hacia la liberación” (2000, p. 38). Sin embargo, esta acusación me parece injusta, pues la teoría de Frankfurt alienta decididamente el esfuerzo del agente porque sean los deseos de primer orden los que el sujeto modifique. Ciertamente, si al final lo único que le queda a nuestra amiga es resignarse a su baja autoconfianza somática y adecuar sus deseos de segundo orden a esta condición, el resultado será descorazonador. Pero: *C'est la vie!* Si no hay más, si resulta imposible deshacerse de ella, más le valdrá que la acepte e intente establecer una nueva relación, no patológica, para con ella. Saber aceptar las propias limitaciones, como propias, no deja de ser una virtud. (Véase Velleman (2002) para una discusión de este tipo de cuestiones en la obra de Frankfurt.) Estas palabras, que Marguerite Yourcenar pone en boca del emperador romano Adriano, resultar muy iluminadoras al respecto: “Y es de esta manera, con una mezcla de reserva y audacia, de sumisión y revuelta cuidadosamente concertadas, de exigencia extrema y de concesiones prudentes, que finalmente me he aceptado a mí mismo.” (*Memorias de Adriano*; “Varius multiplex multiformes”, I.)

5.1.2. Reflexividad racional

Por otro lado, los estados y propiedades mentales que conforman el carácter de una persona son de muy diferentes tipos, en relación al grado de control que el agente puede ejercer sobre ellos. No es lo mismo una tendencia temperamental, que un deseo o una creencia. Una tendencia temperamental es, por lo general, algo a lo que uno está pasivamente sujeto, algo que *se sufre*; como también lo son las sensaciones y ciertos deseos. Son estados respecto a los que el agente es predominantemente pasivo. Por el contrario, hay otros tipos de actitudes, fundamentalmente activas, que conllevan un sentido de agencia y autoridad muy diferente. Puede ser interesante a este respecto caracterizar la diferencia entre dos tipos de deseos, en relación a nuestro grado de control sobre ellos.¹⁰ Hay deseos, como pueden ser los asociados con el hambre o la fatiga, que se experimentan como sensaciones (como, por ejemplo, la sensación de dolor, de vértigo, etc.) que a uno simplemente le sobrevienen. En el caso de Ana, parece que el deseo de huir rápidamente del lugar del atropello es una fuerza que la asalta contra su voluntad. Sin embargo, no todos los deseos son de este tipo —aunque así lo sugiera la tradición filosófica, empezando por Platón. Hay otros deseos que constituyen estados de gran complejidad intelectual, actitudes que articulamos, revisamos, razonamos, y a las que sólo llegamos tras una larga reflexión. Estos últimos deseos más sensibles al juicio dependen de ciertas creencias acerca de lo que hace deseable el objeto del deseo. En este sentido, si la creencia relevante para el deseo en cuestión resulta ser falsa, o simplemente cambia, el deseo también debe variar o apagarse. Finalmente, las creencias son, en general,

¹⁰ La distinción puede formularse en diferentes términos: deseos motivados y no motivados, sensibles al juicio y no sensibles, dependientes de creencias relevantes o no dependientes, etc. Sigo a Moran (2001) pp. 113-20, que a su vez se basa en Nagel (1970) y Scanlon (1998), cap. 1.

el tipo de actitudes más sensible a la revisión y reflexión; aunque, por supuesto, el control que ejercemos sobre ellas sea también distinto al que ejercemos sobre las acciones básicas, del tipo coger un vaso, levantar una pierna, etc.

Podemos, pues, avanzar y afirmar que es posible que lo fundamental no sea el autocontrol reflexivo meramente volitivo, sino el autocontrol reflexivo *racional*; es decir, que la capacidad reflexiva debe basarse en la consideración de razones, y no sólo en la mera jerarquía de deseos.¹¹ Lo que se requiere, para acabar de dar forma a los poderes de autocontrol reflexivo, es el autocontrol racional, a saber, la capacidad del agente para adaptar las propias actitudes a sus razones, y dirigir su propia conducta de acuerdo a ellas.¹² Así, y más allá de la mera identificación reflexiva, lo importante sería que el deseo de orden superior de Ana de luchar contra su tendencia natural a la introversión y a evadir sus problemas, sea adecuadamente sensible a lo que su razonamiento práctico le dicta que debe hacer, o como lo mejor que puede hacer. (Si, finalmente, la misma fuerza del razonamiento no es suficiente para que Ana se deshaga de este rasgo, hay también técnicas indirectas de autorreforma a su disposición, que hacen uso del condicionamiento, la terapia, etc. —justificadas, en última

¹¹ Gary Watson (1975) ha destacado en la defensa de que un agente es autónomo cuando los deseos de primer orden, o *meros* deseos, están regidos, no por otros deseos (de orden superior), sino por el sistema de valores del agente. Véase también Taylor (1976). Por su parte, Angela Smith ha defendido que para atribuir correctamente un estado mental a una persona hay que situarlo en el sistema racional de otros juicios y creencias que la persona en cuestión acepta (2005, p. 255).

¹² En esta línea, Aristóteles apela la autoridad de la razón para llevar a cabo este autoexamen: es por medio del ejercicio de la capacidad racional que convertimos nuestras tendencias naturales positivas en disposiciones afirmadas y abandonamos otras tendencias menos favorables. Sin duda, también para Kant la razón es el elemento fundamental para la autonomía, pero su posición es mucho más fuerte, como veremos. Para un estudio comparado del papel de la razón y su relación con el tema de la suerte moral en Aristóteles y Kant, véase Anathaloussis (2005), cap. 2-4 y 5-6, respectivamente.

instancia, por la conclusión racional de que este rasgo es algo que hay que eliminar.)

Aun así puede ser que, en última instancia, Ana sea incapaz (o sea imposible para ella) deshacerse de este rasgo. No obstante, y aunque lo conserve, hay estrategias adicionales que pueden ayudarla a tenerlo neutralizado, como ser consciente de esta falta de confianza e intentar sobreponerla en los momentos clave. Por otro lado, el mismo ideal de integración psicológica completa y del control racional de todos nuestros deseos y creencias no puede ser más que un *ideal*. Pues un sujeto ni suele ni parece que pueda haber reflexionado sobre todos sus deseos y estados mentales, ni siquiera sobre una parte significativa de ellos,¹³ y además en muchas ocasiones actuamos impulsivamente sobre la base de algunos deseos, sin que quede coartada nuestra autonomía por ello; e incluso en otras muchas ocasiones el sujeto no tiene una respuesta inequívoca en relación a lo que realmente quiere respecto a una cuestión determinada, es decir, se dan multitud de casos de ambivalencia sostenida. Además, el tratar los deseos o rasgos rechazados, pero que de alguna manera permanecen en nosotros, como algo completamente ajeno a uno mismo puede llevarnos a una cierta esquizofrenia. (Frankfurt acepta que hay deseos no queridos que no son eliminables y con los que el agente debe aprender a convivir. Lo mismo vale para otras actitudes igualmente recalcitrantes.¹⁴)

¹³ Respecto a esto, Richard Moran ha afirmado que lo fundamental, en todo caso, es que “una creencia es la conclusión posible de algún razonamiento teórico”, al igual que “lo esencial para que un deseo cuente como ‘motivado’ en el sentido relevante es ser la conclusión posible de algún razonamiento práctico.” (2001, p. 116) Sin embargo, esta posición, aunque en principio plausible, no está tampoco libre de problemas.

¹⁴ Sobre esta cuestión, Frankfurt (1976).

5.1.3. Dos proyectos

En cualquier caso, estas consideraciones en torno a la posibilidad de revertir nuestro carácter siguen sin bloquear el principal sentido de suerte moral en cuestión: de nuevo, el problema que presenta el fenómeno de la suerte moral constitutiva no es (directamente) el de anular la responsabilidad del agente, sino el hecho de que partir de unas condiciones más o menos favorables, hace que alcanzar una determinada configuración (moralmente relevante) del carácter resulte más o menos difícil. Por ejemplo, tener una tendencia temperamental que impulsa al egoísmo, sea innata o sea formada en los primeros años de vida, hace normalmente más difícil la conducta altruista. Obviamente, ello no excluye la posibilidad de que uno llegue a superar esta tendencia, ni le exime de su responsabilidad; sino que simplemente supone una desigualdad de origen, que inevitablemente influye en la evaluación que se haga de él. Si bien la posibilidad de revertir nuestro carácter parece, por lo menos, hacer menos traumática la existencia del tipo constitutivo de suerte moral.

De hecho, el adversario de la suerte moral constitutiva puede acogerse a esta última consideración. En particular, el kantiano concede que cada cual debe afrontar, y superar, tendencias instintivas o tentaciones adversas de diferente fuerza, lo cual es fundamental para determinar la valía moral del agente. Con ello está reconociendo cierta suerte moral; por bien que, dentro del sistema kantiano, esto no sea ceder mucho: pues, si bien puede ser cierto que cada persona aparece en el mundo con unos rasgos constitutivos diferentes, y se encuentra en su desarrollo con diferentes tentaciones, que le dificultan o facilitan la meta de actuar correctamente; no obstante, dada la libertad y autonomía inherente a toda persona en tanto que ser racional, está siempre en su mano oponerse a aquellos rasgos negativos, así como reforzar los positivos, de modo que el juicio moral

que merezca dependa finalmente de lo que está bajo su control. De hecho, cuanto más poderosas sean las tendencias instintivas o tentaciones que uno ha tenido que vencer para actuar correctamente, mayor será su valía moral; lo que al final supone que encontrarse con mayores obstáculos no es un factor de *mala* suerte, sino de *buena* suerte. Para Kant, éste es un poder intemporal que se fundamenta en el seguimiento de los dictados de la Razón.

Este proyecto, que llamaremos *superación racional* de la suerte moral, depende de una concepción fuerte de la razón (la Razón), que contrasta con la que hemos considerado anteriormente al hablar del autocontrol racional. Aun sin llegar a la excesiva crudeza de Kant, es plausible la defensa de una concepción de las razones, de qué cuenta como una verdadera razón, más fuerte u objetivista que la anterior. Mientras que aquella era una capacidad racional de *autorrevisión*, lo que se demanda ahora es una capacidad racional de *autocorrección*. Esta parece ser la posición de aquellos que abogan por el requisito de la *competencia normativa*.

Pero, hay además otra cuestión, la del optimismo exacerbado que supone semejante confianza en el poder del agente, que a través del ejercicio de su razón puede revertir su carácter, o resolver cualquier problema moral, en cualquier momento. No hace falta tener mucha experiencia para darse cuenta de la gran diferencia existente, por ejemplo, entre tratar de prevenir la adquisición de un vicio e intentar deshacerse de uno ya asentado como hábito conductual. Pero el adversario de la suerte moral no necesita ser tan optimista (o falta de realismo) como para negar esto, pues tiene otro movimiento a su disposición: los factores antecedentes, que conforman la historia formativa particular, que llevaron al agente a tener el carácter que actualmente tiene, son fundamentales para la exclusión de la suerte del juicio moral que éste merece. Llamaré a este otro proyecto la

superación histórica (por la importancia que concede a la historia singular del agente) de la suerte moral.

Por otro lado, una vez reconocemos la importancia de la historia formativa de cada persona en relación a la responsabilidad por su carácter, nos percatamos también de que la misma capacidad racional reflexiva está sometida a los avatares de las circunstancias de su desarrollo. La imagen que hemos considerado en este apartado venía a desdoblar al sujeto en dos: aquello sobre lo que el agente reflexiona, foco de su atención, y la naturaleza de la propia capacidad reflexiva. Pero hemos de reconocer que la misma capacidad reflexiva, incluso el mismo acceso a la razón, es algo que depende de la historia formativa particular de cada individuo, de sus capacidades mentales, de sus circunstancias familiares y sociales particulares, así como de sus distintivas experiencias vitales. Sin embargo, el racionalista tiene una respuesta a esto, que insiste en la idea anterior: la capacidad normativa es algo que se tiene o no se tiene; habrá quienes no la alcancen, pero quienes lo hagan estarán, *eo ipso*, en la situación de superar los factores contingentes de su desarrollo.

En el siguiente apartado exploraré el primer proyecto, el de la superación racional, que tendrá que afrontar el problema del sometimiento de la misma adquisición de la capacidad racional al proceso histórico. Abordaré el proyecto de la superación histórica en las secciones subsiguientes.

5. 2. Competencia normativa y condiciones sociales

Cabe remarcar que la teoría de la autonomía centrada en el autocontrol reflexivo, que hemos considerado, pretende superar las concepciones que se limitan a exigir la posesión de un poder o habilidad para

actuar como queremos actuar, para adecuar nuestras acciones a nuestra voluntad. Este era el rasgo definitorio de teorías como las de Hume, Hobbes o Mill, para quienes lo fundamental era que el agente esté libre de los diferentes tipos de limitaciones externas, como son la retención física, la coerción, la parálisis, la intimidación, la amenaza, la opresión, etc. La insistencia en el autocontrol reflexivo viene a enfatizar otras limitaciones, éstas de tipo interno, que han de superarse para alcanzar la condición de agente autónomo. En concreto, la idea de Frankfurt era que “Una persona no puede ser identificada con todo lo que pasa en su mente... como no puede ser identificada con todo lo que pasa en su cuerpo”.¹⁵ Para ello, además del poder anterior, se requiere que el sujeto sea capaz de comprender y reflexionar sobre el conjunto de sus estados mentales, evaluarlos y modificarlos o aprobarlos convenientemente, superando compulsiones, obsesiones, adicciones y neurosis. Además, el autocontrol racional (como autorrevisión) venía a insistir en la consideración de razones para la revisión de deseos, creencias y otras actitudes. Ahora, el nuevo requisito para la autonomía es la *competencia normativa*, o control racional como autocorrección o automejora.¹⁶

En la presente sección abordaré esta nueva condición, considerando la cuestión de su génesis en el agente, pero haciendo hincapié en qué es poseer esta competencia normativa y en la diferencia existente entre poseerla y no poseerla en relación a la responsabilidad moral. La cuestión última será la de si la posesión de esta competencia neutraliza suficientemente las diferencias en las circunstancias particulares de la historia de cada agente en cuanto a sus tendencias morales.

¹⁵ Frankfurt (1976), p. 240-2.

¹⁶ Sobre la *competencia normativa*, además de Wolf (1987) y (1990), véase también Fischer y Ravizza (1998).

5.2.1. Apreciar lo Verdadero y lo Bueno

Para ello, podemos recurrir a la teoría de Susan Wolf, según la cual “un agente no puede tener la clase de libertad y control necesarios para la responsabilidad a menos que, cuando toma sus elecciones sobre valores y acciones, pueda comprender los rasgos significativos de su situación y de las alternativas entre las que escoge.”¹⁷ Ser un agente autónomo comportará ser capaz de apreciar las razones existentes para corregirse a uno mismo, haciendo “lo *correcto* por las razones *correctas*” o “actuar de acuerdo a lo Verdadero y Bueno”. También John McDowell ha insistido recientemente en que “uno se autodetermina en la medida en que piensa o actúa como lo hace por razones que, pensando y actuando como lo hace, responden a las razones que hay.”¹⁸ Como en el caso de Wolf, la sensibilidad a las *razones que hay* se refiere tanto a las razones teóricas, esto es, las razones para creer, como a las razones prácticas, o razones para actuar.

En todo caso, para que el agente sea capaz de tal cosa, su educación y el tipo de entorno social en el que se forma resultarán fundamentales. Para ilustrar su posición, Wolf nos propone la historia de JoJo, el hijo favorito de Jo I, un malvado y sádico dictador de un pequeño país subdesarrollado. A causa del especial amor que su padre le profesa, JoJo recibe una educación especial y se le permite acompañar a su padre y observarle en su rutina diaria. Dado este trato, no es extraño que el pequeño JoJo tome a su padre como modelo y desarrolle sus mismos valores. Ya como

¹⁷ Wolf (1990), p. 117.

¹⁸ McDowell (2007), p. 1. McDowell insiste correctamente en que lo que hay en juego es la afirmación de una capacidad: “*la capacidad de dar un paso atrás y evaluar si las supuestas razones garantizan la acción o creencias, más bien que su ejercicio.*” (p. 4; mi subrayado.)

adulto, hace muchas de las mismas cosas que hizo su padre, incluyendo el enviar a gente a prisión o a la muerte o a cámaras de tortura por mero capricho. Por otro lado, no hace estas cosas por *coacción*, sino que actúa en virtud de sus propios deseos y con sus razones. Son, además, deseos sobre los que reflexiona y que aprueba, con los que se identifica plenamente, y que responden a sus mismos valores. Afirma Wolf: “no podemos decir de JoJo que su identidad, como agente, no es la identidad que él desea tener. Es quien quiere ser. Desde dentro, se siente tan integrado, libre y responsable como nosotros.”¹⁹

El problema con JoJo está en su educación. Para Wolf la cuestión es que el tipo de educación recibido le impide ser una persona moralmente *cuerda* —lo que no significa necesariamente que se trate de alguien mentalmente enfermo. Wolf entiende la cordura como “la habilidad cognitiva y normativa mínimamente suficiente para reconocer y apreciar el mundo tal y como es.”²⁰ Sin embargo, el ser o no una persona cuerda escapa completamente a nuestro control, depende del tipo de educación recibido y del entorno familiar y social. Ser o no como JoJo es algo que no está en nuestro poder determinar y, por lo tanto, es cosa de suerte para cada uno de nosotros. Sin embargo, si la teoría de Wolf es correcta y los que son como JoJo deben quedar excluidos de las atribuciones de responsabilidad moral, entonces JoJo y sus semejantes han sufrido una muy mala suerte formativa, pero no *suerte moral formativa*. (Para que haya suerte moral uno deber ser moralmente responsable.)

Con todo, la distinción de casos como el de JoJo y casos *normales* en los que la educación recibida no impide las atribuciones de responsabilidad, no está carente de problemas. La misma Wolf considera otros casos

¹⁹ Wolf (1987), p. 54.

²⁰ Wolf (1987), p. 56.

de agentes que “falsamente creen que los modos en que actúan son moralmente aceptables [...por lo que] su conducta es expresiva de o concuerda con el verdadero yo del agente.”²¹ Sus ejemplos incluyen a los ciudadanos normales de la Alemania nazi, a los norteamericanos blancos hijos de propietarios de esclavos alrededor de la década de 1850 y a todas las personas educadas durante la primera mitad del siglo XX para aceptar los convencionales roles de sexo como rasgos profundos de la naturaleza humana o la sociedad civilizada. Todas estas personas son víctimas, en relación a las actitudes por las que las censuramos, de sus sociedades. Sus sociedades, equivocadas respecto a estas cuestiones, les llevaron a sostener creencias equivocadas y a actuar de manera moralmente errónea.

Esta teoría ofrece una explicación de por qué las víctimas de infancias con privaciones severas y las víctimas de sociedades equivocadas no pueden ser responsables por sus acciones, de la misma manera que lo son aquellos que no han sufrido estas privaciones en su infancia, ni han sido llevados a error por la sociedad en la que viven. No obstante, hay una diferencia importante entre ambos grupos: a unos, como a JoJo, se les niega por completo la condición de agentes autónomos; mientras que a los otros sólo se les retira esta condición en relación a uno de sus rasgos o creencias (¿o es que alguien ya no es autónomo por meramente creer en la superioridad de una raza, un género, etc.?). El segundo grupo está constituido por agentes que, aunque capaces de apreciar muchas actitudes y juicios morales, muestran huecos en su comprensión moral que se explican, de manera más natural, por elementos de su desarrollo y ambiente social, por presiones y normas sociales que obstaculizan los juicios morales correctos. La misma Wolf duda de si retirarles por completo la condición de agentes moralmente responsables en relación a estas creencias o sólo

²¹ Wolf (1987), p. 57.

disminuirla. Además, resulta “difícil establecer, incluso en los casos individuales, que estas personas no eran capaces de ver y apreciar la injusticia de algunas de las prácticas, actitudes y instituciones de sus comunidades”.²²

Al margen de los méritos particulares de esta propuesta, son dos las cuestiones que nos conciernen aquí. La primera es la determinación de qué personas pueden ser consideradas como moralmente responsables o no, es decir, a quienes podemos culpar o debemos excusar, a partir de las condiciones propuestas. La segunda es la de qué grado o tipo de objetivismo se asume al usar la expresión “lo Verdadero y lo Bueno”, o “las razones que hay”.

Pues bien, la primera cuestión parece ser un problema básicamente epistémico. Wolf, y McDowell,²³ nos recuerdan que no hay ningún método para determinar con claridad qué educación e influencias son consistentes con el requisito de la competencia normativa, con la habilidad para ver qué debe hacerse y actuar en conformidad. Puede que esto no preocupe en exceso a quienes se desentiendan del problema de la aplicabilidad concreta de la teoría. No obstante, hay un sentido de esta primera cuestión, que la conecta estrechamente con la segunda, y la convierte en un problema también para estos. Pues, dado que la condición de agente autónomo depende de la capacidad para *la correcta comprensión y apreciación de cómo es el mundo y de las consideraciones morales correctas*, en resumen, de la comprensión y apreciación de “lo Verdadero y lo Bueno”, dependerá de qué sentido concreto se le dé a esta expresión para saber qué

²² Wolf (1990), p. 121.

²³ Para McDowell, tras reconocer que no hay un criterio, en el sentido de una fórmula, para distinguir las razones genuinas de las aparentes, sostiene que lo único que cabe añadir es que es uno mismo quien “ha de resolver *por sí mismo* la cuestión de si la manera en que está inclinado a pensar es el modo correcto de pensar.” (2007, p. 7.)

agentes (y qué proporción del conjunto total) pueden ser considerados agentes autónomos. Si la concepción de lo Verdadero y lo Bueno por la que se aboga es fuertemente objetivista, la condición de la competencia normativa se volverá muy restrictiva y será mucho menor el número de persona autónomas —de candidatos adecuados para las atribuciones de responsabilidad moral— que hay en el mundo. Por el contrario, si la concepción favorecida es más pluralista, mayor será el número de personas incluidas en el grupo de agentes autónomos, pero menos *protegidos* estarán en relación al grado de control o autonomía requerido. (Lo mismo sucede en relación a la censura por un rasgo o actitud determinado.) Este es un dilema importante que nadie preocupado por estas cuestiones puede eludir.

Wolf afirma que su competencia normativa sólo implica la existencia de criterios no arbitrarios de corrección moral, que no exceden los compromisos implícitos en la mayor parte de nuestro discurso moral concreto y corriente. En particular, “no necesita asumir que estos criterios determinan un sistema de valores y juicios de valor único, universalmente aplicable, completo y óptimo, ni necesita asumir la accesibilidad o incluso la inteligibilidad de la existencia de un punto de vista independiente de toda cultura desde el que se establecen estos criterios.”²⁴ Lo único necesario, según Wolf, para su teoría es un Pluralismo Normativo, que acepte tan solo que hay juicios de valor más correctos que otros.²⁵ Sin embargo,

²⁴ Wolf (1990), pp. 124-5.

²⁵ Wolf parece vacilar respecto al nivel de objetivismo moral que su propuesta implica. Algo parecido sucede en el caso de McDowell, quien muy vagamente afirma que “la diferencia entre responder a razones genuinas cuya fuerza racional uno aprecia y responder a algo que meramente le parece a uno ser una razón marca la diferencia entre un ejercicio de la autonomía y una mera apariencia de autonomía.” (2007, p. 7.) Y añade: “la carga de la autonomía: la responsabilidad para reflexionar por uno mismo acerca de las credenciales de supuestas razones para pensar y actuar, con completa conciencia de que uno no puede confiar acríticamente en ninguna supuesta sabiduría [de la comunidad

esta solución es completamente insuficiente para la *superación racionalista* de la suerte moral. Asegurar la superación racionalista de la suerte moral formativa comportaría asegurar que la misma posesión de la competencia normativa por parte del agente, le da un poder racional tal que le permite superar toda contingencia que en su desarrollo le pueda conducir a formarse un carácter moralmente negativo en los aspectos relevantes. Para esto, para que el disfrute mismo de la competencia normativa comportara la exclusión de la suerte moral formativa (o la mayor exclusión posible de suerte moral formativa), se precisaría de una concepción mucho más objetivista de lo Bueno. Pero esto supondría, como he dicho, una restricción amplísima del número de personas moralmente responsables existentes en el mundo. Puede que haya teóricos que estén dispuestos a aceptar este resultado y, sin duda, entre los adversarios de la suerte moral habrá bastantes dispuestos a hacerlo. Sin embargo, estaríamos de nuevo ante una propuesta revisionista que desautorizaría gran parte de nuestras prácticas comunes.²⁶

No cabe duda de que este tipo de disquisiciones nos puede llevar a desconfiar de nuestra propia autonomía. Ciertamente, el riesgo de que seamos mucho menos responsables por nuestros rasgos y acciones de lo que creemos es algo que siempre está ahí. Pero negar la responsabilidad moral de aquellos que no han apreciado lo Verdadero y lo Bueno, en una interpretación excesivamente objetivista, o incluso meramente reducirla,

a la que uno pertenece] que uno meramente ha recibido. No se puede aceptar confiadamente nada de nadie. [...] Nada es inmune al escrutinio reflexivo.” (p. 9.) Si nos centramos en la frase de que “no se puede aceptar confiadamente nada de nadie”, su racionalismo parece ser enormemente exigente.

²⁶ Ya di algunas razones en el capítulo anterior para rechazar el revisionismo inmoderado acerca de las prácticas; para un tratamiento más detallado habrá que esperar al capítulo 8.

nos encaminaría hacia la exención o excusa de aquellos que precisamente, por sus acciones, parecen más censurables.²⁷

5.2.2. Afectividad y relaciones interpersonales en la formación del carácter

Por otro lado, tampoco hay que exagerar el papel de la Razón en la autodefinición del propio carácter. Hay también otros tipos de consideraciones, no estrictamente racionales, como es el caso de las emocionales, que juegan un papel muy relevante a este respecto. Considérese el caso de Bazárov, protagonista de *Padres e hijos* de Turguénev.²⁸ Bazárov es un joven radical, comprometido con el nihilismo como ideología total, que afecta tanto a la política, la organización social, como la ética individual; que, a su vez, desprecia profundamente el romanticismo y toda actitud afectada ante la vida. Convengamos que sus ideas nihilistas se siguen del conjunto de razones que considera correctas, con sus pesos y anclajes particulares. Y, obviamente, estas ideas son compatibles con el desprecio de la *vida afectada* —de hecho, se sigue de las primeras. Además, traicionar estos valores (que son sus valores) es algo *impensable* para él. Sin embargo, un día conoce a Ana Serguéievna y su trato continuado acaba haciendo que se enamore de ella. Bazárov comienza a despegarse progresivamente de sus ideas nihilistas y a comportarse más y más emocionalmente. Siente, entonces, una tensión entre sus ideas y su sentir más profundo, y se torna incapaz de evitar este nuevo comportamiento, que le revela algo que genuinamente quiere. En realidad, no puede decirse que se trate cla-

²⁷ Mi sospecha es que, si presionamos en esta idea, en último término sólo los moralmente (y epistémicamente) virtuosos podrán ser moralmente responsables. Aunque no estoy seguro de que esto se siga necesariamente del punto de vista considerado. También Kant hace alarde, en ocasiones, de un cierto *socratismo*.

²⁸ Es posible que mi descripción de la historia de Bazárov no se corresponda exactamente con el relato original de Turguénev. En todo caso, esto no es importante, aquí.

ramente del descubrimiento de algo que ya estaba en él pero que ignoraba. Más bien, ha sido el conjunto del proceso el que le ha conducido a esta, por así decir, *conversión*. Bazárov podría incluso seguir creyendo que el nihilismo es la conclusión correcta de sus razones, que es la posición que debe mantener; sin embargo, ya no es algo que sienta como importante para él, ya no puede dirigir su vida en virtud de sus consignas. Así, es posible que cuando reflexione o discuta de política y ética con otras personas se muestre todavía como un fuerte defensor de las ideas nihilistas; pero cuando se trata de actuar, de vivir su vida diaria, éstas ya no constituyen un patrón de conducta y pensamiento para él. En realidad, el mismo hecho de dejar de experimentar estas ideas como algo por lo que se siente realmente preocupado, como el proyecto que articula su vida, es muy probable que conlleve una debilitación progresiva de su consideración de ellas como necesarias, como racionalmente concluyentes. Aquello por lo que, cada vez más, se siente preocupado Bazárov es por Ana Serguéievna. En este caso, es su misma influencia (quizá no querida ni por la misma Ana) la que le ha llevado a la conversión de su carácter.²⁹

En realidad, estos procesos de conversión, en un sentido lato del término, no son cosas excepcionales o literarias, sino que abundan en nuestras vidas, y resultaría muy implausible decir que impiden las atribuciones de responsabilidad moral. Piensa en la evolución de tus ideales políticos, por ejemplo; en la gran mayoría de casos es fruto de muchas

²⁹ Hay una cierta relación entre casos de este tipo y lo que Frankfurt ha llamado “necesidades volitivas”. Véase Frankfurt (1999b), p. 110; también Frankfurt (1988c), pp. 80-94. (Algunos trabajos previos de Frankfurt sugieren que lo necesario para la autonomía es, más bien, la identificación, donde identificación no implica necesidad volitiva.) No obstante, esta noción se refiere a algo que el agente descubre acerca de su verdadero yo, mientras que en este caso se produce, como digo, un proceso de conversión. El problema de la noción de Frankfurt es que, como su perspectiva general, es esencialmente estática. Ver Velleman (2002) para una discusión de este problema, así como la réplica de Frankfurt en Frankfurt (2002a). En Rosell (en prensa) abordo esta cuestión explotando, precisamente, el caso de Bazárov.

experiencias vitales, que llevan a uno a considerar otras cosas como importantes, como *cosas que deben considerarse*. Y no hay duda de que estos procesos de conversión tienen una dimensión moral. En el caso de Bazárov, la conversión le ha evitado posibles acciones terroristas (supongamos que su nihilismo incluye el nihilismo político activo, que se caracteriza por el activismo terrorista indiscriminado); así, al abandonar el nihilismo, o perder éste su fuerza vital, Bazárov se aparta del camino que podría haber conducido (muy probablemente, dado el alto grado de su convencimiento) a convertirse en un asesino de masas. En definitiva, éste y otros avatares del mismo tipo nos vuelven a mostrar cómo de dependiente de nuestro entorno es nuestra identidad como agentes morales.

Hay todavía otra cuestión, que va en paralelo a lo anterior, que no me resisto a comentar. Creo que hay que empezar a reconocer que la mayor parte del debate sobre la responsabilidad moral está dominado por la desconfianza generalizada respecto a las influencias externas sobre el agente —de hecho, parece que por la aspiración a su total neutralización. Normalmente, se toman casos de influencias externas negativas (neurocientíficos malignos, manipuladores, etc.) que de algún modo han de bloquearse o neutralizarse para que la autonomía del agente no se resienta. Sin embargo, habría que destacar también lo contrario, que no por obvio deja de ser cierto: que sin influencias externas, no hay autonomía ni responsabilidad ni nada. Esto es, para la formación y desarrollo de una persona son fundamentales cosas tales como la influencia de sus padres, familiares, maestros, amigos, colegas, entorno comunitario, organización social, etc. En concreto, para una buena formación se requieren buenos padres, buenos maestros, relaciones sociales afortunadas, etc. No hay duda de que la formación adecuada de alguien depende en gran medida del *buen control ajeno*, es decir, del control benéfico de las personas de nues-

200

tro entorno que tienen poder (en diferentes grados) para ello. Nuestro carácter e identidad como agentes depende también de, es inconcebible sin, la red de relaciones interpersonales en la que estuvimos y estamos inmersos. Creo que este es un hecho que normalmente no se enfatiza lo suficiente.³⁰

En este sentido, dado que nuestro mundo está constituido por una extensa red de interacciones humanas plurales, en las que unos influyen sobre otros de modos más o menos racionales y equitativos, un elemento que debería incluir toda teoría satisfactoria de la responsabilidad moral es la regulación del tipo de influencias adecuadas que los demás pueden ejercer sobre nosotros. Lo relevante aquí sería algo así como que el agente conserve su *control discursivo* en el mayor número de sus relaciones interpersonales y grupales.³¹

En definitiva, ante el resultado anterior, algunos autores han insistido más en el hecho de que el agente debe ser el origen adecuado del proceso mismo por el que uno llega a constituirse en la persona que es, que en la competencia normativa misma. Aunque, por supuesto, ello no supone la desatención a las condiciones de racionalidad de este desarrollo; sino, más bien, un cambio de énfasis. (El énfasis puede ir de considerar exclusivamente uno u otro aspecto, a defender un equilibrio entre ellos, o a incidir más o menos en uno u otro.) Me ocuparé a continuación de esta

³⁰ Para casos (excepcionales) de teóricos que sí que han incidido en esta cuestión; véase especialmente, en el debate específico sobre la suerte moral, Browne (1992), p. 353; y Arpaly (2003), esp. pp. 128ss, y Moya (2007), cap. 6, en el contexto más amplio del debate en torno a la responsabilidad moral. Cfr. con la afirmación de McDowell de que “no puede aceptarse confiadamente nada de nadie” (véase *supra*, nota 25).

³¹ Para un intento, muy tentativo y vago, de elaboración de la noción de “control discursivo”, véase Pettit (2001), capítulo 4.

otra perspectiva, que se corresponde con lo que anteriormente llamé *superación histórica* de la suerte moral formativa.

5. 3. Autocreación: decisiones y acciones autoformativas

5.3.1. ¿Autocreación?

Cabe remarcar que la cuestión en que estamos inmersos es, en buena medida, lo que David Zimmerman ha llamado *el rompecabezas de la autocreación naturalizada en tiempo real*.³² Esto es, la cuestión de cómo es posible que a partir de un estadio en el que el sujeto es un niño que no tiene casi ninguna capacidad mental y, atravesando un proceso de crecimiento progresivo, termine poseyendo la capacidad para *decidir por sí mismo* y para elegir sus propias preferencias y valores. Esta historia tiene dos versiones. La pesimista dice, más o menos, así: cuanto mayor es el grado de desarrollo de alguien, como es el caso de los adultos, más formado está ya su carácter y menores son las posibilidades para modificarlo; mientras que cuando hay mayores posibilidades de moldeado, como en los niños y adolescente, menor es su capacidad reflexiva y su control racional sobre qué y cómo desarrollar o modificar. Por su parte, la versión optimista apela a la naturaleza gradual y progresiva del desarrollo: conforme va aumentando la capacidad mental e intelectual del individuo, aumenta su control racional y éste puede afrontar decisiones de mayor complejidad e importancia. Daniel Dennett afirma que “si nos esforzamos podemos imaginar [...] un proceso de autocreación que empieza con un agente no responsable y se constituye, gradualmente, en un agente res-

³² D. Zimmerman (2003), p. 638.

ponsable por su propio carácter.”³³ Es un proceso largo y costoso, que de hecho nunca termina, y en el que son fundamentales cosas diversas como la influencia positiva de otras personas, el acceso a ejemplos positivos, la presencia de tentaciones, o de oportunidades para llevar a cabo determinadas acciones, para la habituación, etc. Pero, en todo caso, lo fundamental es que es un proceso en el que *la misma persona emergente participa activamente en su emergencia*.³⁴

Podemos convenir en que *este desarrollo es un hecho* —que, por supuesto, se da con mayor o menor éxito según los casos. Por lo que la cuestión a investigar será, más bien, la del grado de control del agente sobre el proceso en su conjunto.

En la sección anterior, vimos que la posesión de la competencia normativa o componente racional en la formación del agente, era insuficiente para neutralizar la influencia de la suerte en el proceso. No obstante, el adversario de la suerte moral puede aún esgrimir la que llamé estrategia de *superación histórica* de la suerte moral, que incide en la necesi-

³³ Dennett (1985), p. 170. Aún más optimista es Mark Bernstein, quien explica la historia de este modo: “Gracias a nuestra dotación genética y a los factores ambientales, incluso en este estadio [inicial, “sin naturaleza”] podemos tener disposiciones a formarnos una naturaleza particular, pero hay suficiente plasticidad como para que nuestro carácter pueda instanciarse en una gran variedad de opciones. No hay ningún primer momento en el que una persona tenga un carácter, del mismo modo que no hay ningún primer momento en que uno empiece a leer. Igual que la imitación y la repetición evolucionan finalmente en la lectura, otros fenómenos que comparten esta naturaleza (“protoelecciones”, “protointenciones”, “protocreencias”) se desarrollan finalmente en elecciones, intenciones y creencias hechas y derechas. Como sus maduros contrapartes, esos protofenómenos carecen de un nítido punto de origen. El protocontrol evoluciona durante este período indeterminado de tiempo. [...] Con el tiempo, la complejidad se incrementa resultando en una naturaleza que está producida y modelada, al menos en parte, por nosotros; esto es, por la estructura motivacional que constituye nuestra naturaleza. Así, dentro de los límites de los factores genéticos y ambientales respecto a los que no tenemos ningún control, las opciones que escogemos y que nos hacen lo que somos, son seleccionadas intencional e intencionadamente por nosotros.” (2005, pp. 11-2.)

³⁴ Ver Zimmerman (2003), p. 648.

dad de que el agente sea el *origen* apropiado de sus rasgos y acciones, a través de ser el creador del propio carácter. Lo que se demanda es que el carácter sea producto de la propia agencia, que surja de manera adecuada del yo del agente; por lo que el sentido de control que está en juego es el de *origen*.

A este respecto, algunos teóricos han considerado que lo que se precisa para que uno sea verdadero dueño de su capacidad reflexiva o competencia normativa es que siga un proceso de *asunción de responsabilidad* por su propia capacidad. Este proceso, paralelo al de su desarrollo intelectual, incluiría el *verse a sí mismo* como capaz de responder a las expectativas de carácter moral que la comunidad moral sitúa sobre uno mismo, en tanto que límites para su conducta, así como el espectro de actitudes reactivas que sus acciones susciten en los demás.³⁵ Además, parece claro que el agente debe formarse su propio esquema de valores, por oposición a que éste sea ajenamente forjado por adoctrinamiento u otros medios de manipulación. Un esquema de valores *auténtico* se diferenciaría de uno inauténtico en que surge por medios que facilitan la habilidad del agente para evaluar sus valores y, especialmente, permitir su libertad de actuar o no actuar conforme a ellos. Por ejemplo, el niño fanático del fútbol (o de la música clásica), cuya afición es inducida, no actúa a partir de un esquema evaluativo auténtico, dado que sus educadores no han tenido ningún interés por fomentar la comprensión y sensibilidad del niño, que es lo que permite su maduración moral.³⁶

³⁵ Fischer y Ravizza (1998), pp. 207-39. Por su parte, Susan Hurley ha afirmado que la responsabilidad precisa de un proceso “por el cual se adquiere el mecanismo de respuesta a razones y la autopercepción en relación a estos mecanismos”, que equie al agente con unos mecanismos que respondan de manera suficientemente adecuada a razones objetivas (2003, p. 51-2).

³⁶ Para este tipo de explicación que incide en la distinción entre autenticidad e inautenticidad en el origen de los propios valores, ver Haji (1998), pp. 124-39.

Sin embargo, otros teóricos se han preocupado principalmente por el problema de la regresividad de la responsabilidad; esto es, el hecho de que para ser responsables por algo, hemos de serlo también de las causas de ese algo. El control regresivo de x exige el control de las causas de x , además de x mismo. Resolver este problema, supone poner un freno adecuado a esta regresión.³⁷

Una manera de hacerlo es la de apelar a decisiones o acciones especialmente relevantes, o determinantes, en la formación del carácter personal. Si el agente es el responsable último de este tipo de decisiones, formadoras de su carácter, entonces frenaremos la regresión de un modo adecuado y, a su vez, tendremos una justificación para atribuirle responsabilidad por su carácter y por sus acciones futuras, que sean consecuencia de este carácter. De este modo, nuestro carácter será el resultado de nuestras decisiones, más allá de lo heredado.

5.3.2. Robert Kane: decisiones/acciones autoformativas

Consideremos el famoso caso de Robert Kane.³⁸ Otilia se dirige a una importante reunión de negocios cuando se da cuenta de que una joven está siendo asaltada en un callejón. En virtud de su complejo carácter, su voluntad está dividida. Por un lado, Otilia es una mujer ambiciosa y sabe que la reunión es crucial para su carrera; por el otro, es una persona que se preocupa por los demás y se siente inclinada a ir en auxilio de la agredida. Por lo tanto, posee (y piensa que posee) fuertes razones (egoístas) para continuar su camino y, a la vez, posee (y piensa que posee) fuertes razo-

³⁷ Como vimos, este carácter regresivo de la responsabilidad era, de hecho, lo que motivaba a aquellos que negaban la coherencia de la noción misma de suerte constitutiva — en especial, de Susan Hurley.

³⁸ Kane ha contado su historia de la mujer de negocios en numerosos lugares. Véase, por ejemplo, Kane (1994), pp. 44ss; (1996), 126ss; y (2002), pp. 417ss.

nes (morales) para parar y ayudar a la víctima. Este conflicto da origen a una nivelada lucha interna. Como consecuencia, está bajo el poder efectivo de Otilia tanto pararse y prestar auxilio, como continuar su camino hacia la reunión. Finalmente, decide ayudar a la víctima a expensas de su carrera. Kane describe a Otilia como pretendiendo realizar simultáneamente dos acciones incompatibles, por razones que el agente igualmente comparte. Una lucha interna tiene lugar en la mente de Otilia, que desemboca en la realización de una acción y no otra, lo cual dependerá de sucesos cuánticos indeterminados de su cerebro.³⁹

La acción que realice, sea cual sea, satisfará las condiciones de Kane para la libertad directa. Aparte del hecho de que un proceso indeterminista (los esfuerzos en conflicto de su voluntad) produjo su decisión, Otilia tenía un *control voluntario plural* sobre el conjunto de sus opciones: podía causar ambas opciones voluntariamente, intencionalmente (con conocimiento) y racionalmente (teniendo razones para ello). Dado que está indeterminado lo que hará, pero cualquier cosa que haga será algo que ha intentado hacer y que aprobará como lo que ella quería hacer, Otilia es responsable de lo que haga; además, dado que somos responsables (según Kane) sólo por lo que hacemos libremente, su responsabilidad demuestra que actuó libremente.

Kane llama a esto *decisiones, quererres o acciones autoformativos*. Pero hay que remarcar que, en realidad, la elección de Otilia se produce durante la acción misma. Es actuando de un modo u otro que finalmente

³⁹ Kane describe esta *lucha interna* como la interrupción del equilibrio termodinámico del cerebro del agente, que inicia un proceso caótico, que amplifica el nivel de indeterminación cuántica que es usualmente demasiado escaso como para tener ningún efecto en la conducta o el pensamiento. Pero, a mi juicio, esto es irrelevante para nuestro propósito. Si alguien objeta que es precisamente esto —que posibilita el carácter indeterminista de la decisión— lo fundamental para la meta de acabar adecuadamente con la regresividad, entonces lo concedo por completo. Como se verá, mi objeción no depende de este punto.

se decide por una acción u otra. Por otro lado, es a partir de este tipo de acciones que el agente adquiere un carácter (nuevamente) formado que determina numerosas de sus acciones posteriores, de las cuales el agente es plenamente responsable en tanto que surgen de su carácter y su carácter fue su propia creación. Esto explicaría, por ejemplo, los famosos casos Lutero: aunque Lutero no pudiera más que hacer lo que hizo, ante el tribunal de la Dieta de Worms, pues su carácter le impidió poder actuar de otro modo, es responsable por su conducta en tanto que su carácter es el resultado de sus acciones autoformativas anteriores, que realizó de manera directamente libre.⁴⁰

Para Kane, este tipo de acciones, conceden al agente el poder de ser origen último de sus acciones, del modo adecuado que niega a otras teorías (compatibilistas), pero sin exigir poderes sobrehumanos o cuasimágicos de autocreación, imposibles de alcanzar. Sin embargo, muchos autores han destacado lo que para ellos es un problema importante de esta teoría: si la causa de la diferencia entre la resolución/acción por la que Otilia de hecho opta y su alternativa se halla, por hipótesis, fuera del agente mismo, pues no hay nada en su carácter, disposiciones, creencias, etc., que la incline a ayudar a la mujer en lugar de apresurarse a llegar a la reunión; entonces la resolución/acción real se debe, en último término, a la suerte.⁴¹ En un sentido, Otilia no controla qué opción es finalmente la suya. La respuesta de Kane es esta: realice una u otra acción, Otilia es no obstante responsable por la acción realizada. Si pretendía hacer tanto una cosa como la otra, entonces haga la cosa que haga (de entre las dos opciones igualmente queridas) la aprobará como aquello que pretendía hacer.⁴²

⁴⁰ Los casos Lutero han sido discutidos, entre otros, por Dennett (1984), Williams (1995c) y Kane (1996).

⁴¹ Para esta objeción, véase esp. Haji (2002), Clarke (2003) y Mele (2006).

⁴² Para la respuesta a sus críticos, véase, Kane (1999) y (2002).

Estoy de acuerdo con Kane en que Otilia es responsable de aquello por lo que se decida o haga, sea lo que sea. No obstante, sus críticos también tienen razón en el hecho de que será cuestión de suerte (puede que limitada, pero suerte al fin y al cabo) que haga una cosa u otra. En realidad, creo que esta propuesta es de muy poca ayuda para los adversarios de la suerte moral. Démonos cuenta de que estas decisiones imponen una dirección en el agente, y cierran caminos que (previamente a la decisión) eran posibles y, en gran medida, también caminos futuros. (No digo que otras teorías no tengan que reconocer igualmente este hecho; pero en el caso de Kane, su teoría parece incidir especialmente sobre él.) Propongo que reflexionemos sobre los presupuestos y consecuencias de este tipo de explicación.

Por un lado, la idea de *acciones o decisiones autoformativas* tiene importantes reminiscencias de las famosas “elecciones radicales” de los existencialistas. En particular, Jean-Paul Sartre defendía que mediante una “elección radical” uno se elige a sí mismo *completamente*, como creándose a sí mismo *desde la nada*.⁴³ Su lenguaje es siempre exagerado, pero la idea no es ajena a la posición de Kane: el agente se crea a sí mismo, crea o recrea su propio carácter, de un modo realmente incondicionado (sin negar, por supuesto, que hay razones a favor de cada opción). En particular, si comparamos su caso con el famosísimo ejemplo sartriano del joven que ha de decidir entre permanecer con su madre o unirse a la Resistencia, vemos que en ambos estamos ante una elección dilemática. En el caso de Kane, tal y como él lo plantea, parece que la elección no es entre dos opciones morales, sino entre una opción moral y otra inmoral o egoísta; sin embargo, también podría describirse entre dos tipos de obligaciones (mo-

⁴³ Sartre (1946).

rales).⁴⁴ Esto es más obvio si modificamos el ejemplo ligeramente: Otilia* tiene que escoger entre ayudar a la mujer o llegar a hora a su trabajo y evitar que la despidan. Añadamos que de su trabajo depende su familia. Kane enfatiza la *inconmensurabilidad* entre ambas opciones. No obstante, me parece que lo fundamental en las elecciones autoformativas de Kane es tener que optar entre una opción moral y una inmoral o amoral, que resulte en la definición de la cualidad moral del carácter del agente. Sin embargo, esta dualidad fundamental (digámoslo así) está ausente de muchas de las acciones o decisiones que pueden resultar autoformativas. Y, además, se basa en el presupuesto de que el agente es conocedor del carácter moral o no de una y otra opción en la decisión y de la repercusión que cada una puede tener en su futuro, a través de la definición de su carácter en una dirección determinada.

En todo caso, si estamos ante una elección dilemática, entre dos cursos de acción incompatibles, que son igualmente valiosos para Otilia —ésta tiene razones fuertes para optar por ambas alternativas— es porque no es una elección en el vacío, sino que toda decisión autoformativa presupone necesariamente un carácter previo, con el que el agente llega al momento de estas decisiones. Si esto es así, difícilmente pueden suponer un freno a la naturaleza regresiva de la responsabilidad moral. Ciertamente, podemos retrotraernos al pasado del agente, a decisiones previas donde el carácter estaba menos formado, pero en todo caso siempre que para la decisión el agente disponga de razones fuertes para ambas opciones se precisará un carácter previo para el que estas razones son vinculantes. Además, cuanto más joven es el agente, menos formado está su carácter y

⁴⁴ En el caso de Sartre, la elección se presenta como un dilema moral entre dos tipos de obligaciones —para con su madre y para con la causa por la libertad. Véase Taylor (1976), pp. 293-4, para una crítica del ejemplo de Sartre, en cierto modo, paralela a mi crítica del de Kane en el siguiente párrafo.

menor es su control racional sobre sus decisiones. Por supuesto, Kane respondería insistiendo en que, bueno, si bien cabe reconocer estos compromisos evaluativos previos y el propio carácter del agente con el que llega al momento de decisión, esto no determina la elección. Pero mi objeción es que, ciertamente, no la determina, pero la constriñe de una manera muy importante. Sólo sobre la base de la existencia de un carácter previo, las dos opciones en liza se presentan como queridas y deseadas por el agente.⁴⁵

Por otro lado, hemos de preguntarnos por la articulación de las diferentes acciones o decisiones autoformativas de cada individuo. Es implausible pensar que el carácter de una persona pueda depender de una sola acción autoformativa primera. Más bien se requiere un conjunto de ellas. Sin embargo, si éstas constituyen una secuencia temporal, entonces unas condicionarán a otras, de modo que las primeras serán más fundamentales, esto es, más significativas para la formación del carácter subsiguiente. Pero, si la acumulación de decisiones formativas del carácter tiene lugar mediante su interconexión secuencial, en último término sólo un pequeño grupo (al final, sólo una), podrán cumplir la exigencia del igual peso de ambas opciones para el agente. Con todo, cabe también tener presente que las diferentes decisiones podrían ser independientes entre sí, pero entonces el carácter que se forma sería un todo caótico, inarticulado, o basado en proporciones. Finalmente, podría hablarse de diferentes decisiones vinculadas a diferentes aspectos del carácter del agente, pero en-

⁴⁵ Las palabras siguientes de Gerald Dworkin al respecto, me resultan muy sugerentes: en el amanecer de nuestra autoconciencia racional “simplemente nos encontramos a nosotros mismos motivados de una determinada manera; la noción de elegir, desde cero, no tiene sentido. Tarde o temprano, nos hallamos como en la metáfora de Neurath del barco en medio del océano, que es reconstruido mientras navega, a mitad de la historia. Pero [en tanto que autónomos] siempre retenemos la posibilidad de dar un paso atrás y juzgar dónde estamos y dónde queremos estar.” Citado en Fischer (2006), p. 110.

tonces no pueden presentarse paradigmáticamente como decisiones entre una opción moral y otra no moral.

5.3.3. La “decisión” de Raskólnikov

Pero, volvamos a la cuestión de la suerte en cuanto a la decisión o acción que el agente finalmente lleva a cabo. Acepté que se decida por lo que se decida, el agente será responsable por ello. Sin embargo, esto puede conllevar una concepción *trágica*, tanto en relación a la dependencia del agente respecto a acciones muy puntuales de su historia, como en relación a situaciones incontroladas con repercusiones muy significativas para su estatus moral posterior.

Para ilustrar este punto, propongo que consideremos la “decisión” de Raskólnikov, tan magistralmente narrada por Dostoyevski en *Crimen y castigo*. Raskólnikov es un joven que acaba de terminar sus estudios y que vive preso de las penurias económicas. Para comer y poder pagar a su casera, tiene que recurrir a una vieja usurera. No obstante, es capaz de realizar acciones altruistas, como darle a su amigo Marmeladov el poco dinero que acaba de recibir por un empeño, ante la extrema precariedad de su familia (tan extrema que Sonia, la hermana de éste, se ha visto obligada a prostituirse). Raskólnikov se arrepiente, nada más hacerlo, de haberle dado el único dinero que tenía, pero no intenta recuperarlo. Podemos decir que, aunque dudase entre darles el dinero o no, finalmente se lo da y asume la responsabilidad de su acción. Pero hay más.

Sus problemas económicos aumentan: su familia está igualmente necesitada de dinero y su hermana puede verse obligada a casarse por dinero. Compara el *sacrificio* de ésta con el de Sonia. Tiene además un sueño, que interpreta como un signo sobre su plan de matar a la usurera. De hecho, su mente se halla inmersa en el proceso de planear el asesinato de

la usurera, aunque el mero pensamiento directo en el crimen le resulta repugnante. Podemos decir que Raskólnikov tiene razones (egoístas e inmorales, pero buenas razones) para acabar con la vida de la usurera y robarle la riqueza que acumula “sin desdén”. Pero también tiene sus razones (morales) para no hacerlo: asesinar a alguien es un crimen imperdonable. Durante unos días está inmerso en esta ambigua indecisión. Considera aún más razones a favor y en contra del crimen. Pero un día, se dirige sin más al apartamento de la usurera. Cuando ésta le abre, entra, se interesa por sus cuentas y, de repente, se abalanza sobre ella y la golpea hasta matarla.

Si hay una verdadera indecisión en Raskólnikov, de modo que igual podría haber matado a la usurera como no, dependiendo del peso determinado que otorga a cada grupo de razones en un momento determinado (que muy bien podría depender de un estado de ánimo particular), entonces estamos ante un claro caso de suerte moral, que paradigmáticamente cumple con las condiciones de Kane para las decisiones y acciones de máxima libertad de un agente.⁴⁶ No obstante, desde la posición de Kane podría responderse que si Raskólnikov ni decidió ni se planteó ambas opciones en el momento mismo de actuar, el paralelismo con el caso de Otilia no se da. Ante esto, sólo puedo replicar que si las acciones autoformativas requieren una ponderación consciente y reflexiva en el momento de actuar, entonces es irrealistamente exigente —lo cual no sería tan raro, dada la pobreza de la psicología moral de los casos de Kane.

⁴⁶ Ciertamente, otra interpretación posible es que Raskólnikov ya se había formado (inconscientemente) la intención de cometer el asesinato desde muy antes, pero encubriéndola mediante fenómenos de autoengaño, que le facilitarían esta conducta. Sin embargo, creo que mi lectura es más adecuada, o al menos hace más interesante el episodio. De hecho Raskólnikov no toma, en ningún momento, *la* decisión consciente de cometer el crimen, sino que la decisión se deduce de su acción. Parece que ni en el momento mismo de cometer el crimen se puede decir que Raskólnikov está plenamente convencido de que eso, y no lo contrario, es lo que *quiere* hacer.

5.3.4. La repercusión futura de las decisiones formativas

Pero hay más, el factor suerte va amplificándose con el aumento de las decisiones formativas del agente y su repercusión en la formación progresiva de su carácter. Es decir, la objeción anterior incide en cómo una decisión individual influye, más o menos dramáticamente, en la identidad subsiguiente del agente. Pero cabe también destacar cómo la acumulación de un conjunto de estas decisiones será todavía más determinante en el establecimiento de un carácter concreto e individualizado para el agente; que, además, comportará (según las propias especificaciones de Kane) la responsabilidad moral del agente por las acciones futuras que se sigan de este carácter —incluidos hábitos y adquisiciones de creencias y otras actitudes, posteriormente— que se presentarán como necesarias para él, y que pocas o ninguna opción tenía de anticipar en el momento de llevar a cabo las acciones autoformativas.⁴⁷

Esto es, además del factor azaroso, generalmente reconocido, de este tipo de decisiones, y de la amplificación de esta característica con la acumulación de ellas; además, se introducen nuevos factores de suerte en relación a las consecuencias futuras de estas decisiones (todo aquello en su conducta y pensamiento que se siga del carácter que uno se ha formado) difícilmente anticipables en el momento en el que se suscitan —por la distancia temporal y por la enormidad potencial de estas consecuencias.⁴⁸

⁴⁷ Habría aquí espacio para algo así como el *pesar-por-el-carácter* (*character-regret*), un tipo de pesar-del-agente en el que el implicado se lamenta del conjunto de su conducta pasada que le llevó a convertirse en la persona que es. Véase Rorty (1980) y el capítulo 7 de este trabajo.

⁴⁸ Para aquellos inmersos en el debate acerca del libre albedrío que duden de este resultado, que recuerden la liberalidad de la que puede llamarse condición de *transferencia irrestricta* (TI): la responsabilidad del agente por sus Decisiones Autoformativas se transfiere a todas sus acciones subsiguientes que (a) satisfacen condiciones compatibilistas razonables para la libertad y (b) tiene entre sus condiciones suficientes el estado de su

Un ejemplo del fenómeno que pretendo explicitar podría ser el siguiente. Imagina que alguien en un momento determinado de su adolescencia opta, se esfuerza y consigue desarrollar en sí mismo el rasgo de carácter de la moderación en el trato interpersonal. En concreto, logra habituarse a ser una persona dialogante, que suele tomar en consideración las razones de los demás, incluso de los que se oponen a él o le atacan. Este rasgo, que le permite comportarse virtuosamente en muchas circunstancias, puede incapacitarla para adoptar una actitud de firmeza, de rechazo radical, ante ciertas exigencias que merecerían tal actitud. Y hay circunstancias en las que la falta de firmeza puede conllevar el fallo moral, pudiendo llegar a ser muy dramático. Hanna Arendt ha narrado la singularidad de la actitud del gobierno danés ante las órdenes de los ocupantes nazis para que empezaran a aplicar distinciones entre sus ciudadanos.⁴⁹ Es conocida la hábil estrategia adoptada por los nazis de aumentar muy gradualmente sus demandas, para *hacer factible lo impensable*, podemos decir; la cual parece que sólo pudo ser adecuadamente resistida en virtud de una oposición radical, incluso irrazonable, a sus más pequeñas exigencias inmorales. El gobierno danés, a diferencia del de los otros países invadidos por los nazis, se opuso radicalmente a la inicial aplicación de las distinciones entre sus ciudadanos, evitando caer en la trampa nazi. En resumen, el tema es que entre el momento en que uno desarrolla un rasgo en sí mismo (en la medida en que esto depende realmente de uno mismo)

voluntad que se establece a partir de sus Decisiones Autoformativas. Como ha afirmado Neil Levy (en prensa), TI no requiere que el agente pretenda, en el momento de su decisión indeterminada, ni producir las consecuencias deterministas de su elección que se produzcan, ni que apruebe en ese momento las consecuencias posteriores, ni siquiera que sea consciente de la mera posibilidad de éstas. Levy considera que la teoría de Kane, que califica de *restrictivismo*, es en realidad un compatibilismo encubierto.

⁴⁹ Arendt (2006), p. 252ss. El rasgo en cuestión pertenece a un grupo o colectivo. Pero no es difícil pensarlo igualmente como rasgo de un individuo; por ejemplo, del presidente del gobierno danés.

y las circunstancias futuras en las que este rasgo puede resultar virtuoso o vicioso, media una distancia temporal y una indeterminación tal, que ciertas exigencias de anticipación o previsión de las consecuencias futuras es del todo irrazonable; pero, aun así, uno sigue siendo responsable por estas consecuencias.

En este ejemplo vemos también que las diferencias moralmente relevantes en los sujetos no se limitan a una diferencia en, lo que podemos llamar, la cualidad de su voluntad, sino que se extienden a rasgos de muy diverso tipo. Recuérdese, asimismo, las diferencias entre los asesinos Smith y Hickock de *A sangre fría* (o entre Höss y Eichmann), relatadas en el capítulo anterior. Con esto pretendo reforzar algo a lo que me referí anteriormente: que la mayoría de las decisiones fundamentales que uno toma no suelen consistir entre optar por una opción moral o altruista y una amoral, inmoral o egoísta. Más bien, las decisiones de ese tipo nos obligan a menudo a elegir entre opciones de muy diferente tipo o calidad, que uno considera valiosas y que en el momento de actuar o decidir aparecen como incompatibles. En otras palabras, la diferencia fundamental no se halla tanto en la mera distinción entre rasgos buenos y rasgos malos, sino en que diferentes rasgos pueden propiciar actos de diferente signo sin que se pueda prever con seguridad qué rasgos llevarán a qué acciones, dadas las circunstancias. En todo caso, el ejemplo anterior es sólo una pequeña muestra en relación al gran número de consecuencias futuras que las propias decisiones y acciones pueden tener.

Por otro lado, no podemos pensar que la identidad o carácter del agente depende por completo de decisiones autoformativas del tipo en cuestión (lo que, definitivamente, supondría una idealización), ni de cualquier otro tipo de proceso deliberativo; ni, por otro lado, que no sea posible la reevaluación y reforma posterior; pero sí que hay decisiones o ac-

ciones —en las que, antes de la decisión o acción misma, ninguna de las opciones es singularmente *la tuya*, aunque ambas sean tuyas— que especialmente comprometen al agente respecto a su identidad (sus deseos, planes y valores) subsiguiente y los cursos de acción futuros, por lo que será responsable aún a pesar de la tenue comprensión de sus repercusiones posibles.

Este resultado está eminentemente relacionado con los casos de Gauguin y Ana Karenina, que Bernard Williams presentaba como casos de suerte moral en relación a proyectos particulares. La decisión de ambos está indeterminada de tal modo que no pueden saber cuál será su resultado, hasta que éste se dé; pero, en todo caso, ambos serán responsables del resultado particular de sus respectivos proyectos. (En el caso de Gauguin, hay un factor claro de suerte constitutiva, a saber, su posesión o no de un genio para el arte; que posibilitará o imposibilitará el éxito de su proyecto.) Pero, dado que lo que principalmente se dirime en estos casos es el papel que cabe otorgar a las consecuencias en la determinación del merecimiento del agente, dejo esta cuestión para el capítulo 7, en el que trataré el tema de la suerte moral resultante.

Podemos, pues, reafirmar la principal conclusión de la presente sección de este modo: las decisiones o acciones especialmente relevantes para la formación del carácter constituyen umbrales que comprometen la responsabilidad futura del agente, aún a pesar de la lejanía temporal y del alto grado de indeterminación entre éstas y sus consecuencias o repercusiones futuras. Las limitaciones para su anticipación son reflejo de la esencial *falta de transparencia* entre decisiones autoformativas, y en general el carácter que uno va desarrollando, y las repercusiones en su iden-

tidad y acciones futuras. La relación sólo alcanza su mayor grado de transparencia retrospectivamente.⁵⁰

En todo caso, está claro que la suerte juega inevitablemente un papel central en el proceso formativo del agente, por lo que no podemos hablar de una superación histórica de la suerte moral, mediante la toma de control de decisiones especialmente autoformativas. Más aún, sin la suerte —esto es, sin todo el cúmulo de factores externos e internos que no controlamos y que posibilitan y colaboran en nuestro desarrollo— la misma formación y autonomía personal sería imposible.

5. 4. Rastreo de la responsabilidad moral y la condición epistémica

Pero hay aun una manera más directa de motivar dudas acerca del *control efectivo* de rasgos y actitudes propios, a través del rastreo histórico de la satisfacción de la condición epistémica para la responsabilidad moral por parte del agente. La estrategia será ahora, en vez de considerar las condiciones generales para ser un agente autónomo —o, como dijimos en el capítulo primero, un Agente Moralmente Responsable (AMR)—, centrar más bien la atención en situaciones particulares y tratar de dictaminar si el agente puede o no ser excusado. Para ello, propongo que consideremos el siguiente caso. Hans es un joven fervientemente antisemita. Cuando nos encontramos con él y algo en su conducta o habla nos muestra este rasgo suyo, no dudamos en censurarlo. Pero, cuál es nuestra base para censurar su actitud antisemita. ¿Lo hacemos simplemente por el

⁵⁰ Para otras defensas de la idea de que carecemos de una clara visión de cómo nuestras acciones presentes afectarán a nuestros caracteres y acciones futuras, véase Arpaly (2003), pp. 139-44; Vargas (2005) y Sher (2006). Por otro lado, estos dos últimos, junto con Adams (1985) y Rosen (2004) son referencias importantes de mis disquisiciones en la sección siguiente.

hecho de que posee esa actitud? ¿O resultan relevantes las consideraciones acerca de cómo ha adquirido esta creencia? En la mayoría de los casos simplemente censuramos a alguien por su actitud, y no podemos entrar en la historia particular que le llevó a tener el estado mental censurado. Sin embargo, parece que en nuestra censura está actuando el presupuesto de que estuvo en su poder no adquirir esta creencia. Su responsabilidad dependería de los actos y omisiones voluntarios del pasado que le llevaron a adquirir y mantener la creencia que ahora tiene. A algunos teóricos les gusta distinguir, a este respecto, entre *responsabilidad moral originaria o directa* y *responsabilidad moral derivada o indirecta*, en virtud del tipo de control respecto a aquello por lo que es juzgado, al que el agente tenga acceso. Por supuesto, la segunda depende de la primera.

El rastreo desempeña, así, la función de anclar la responsabilidad en acciones o decisiones anteriores, que pueden incluir la adquisición de creencias, actitudes, disposiciones o hábitos, de las que se deriva la responsabilidad actual del sujeto, que a su vez pueden depender de otras anteriores. Pero a pesar de su atractivo, esta idea no está libre de problemas; en especial, en relación al rastreo de la condición epistémica, necesaria para la responsabilidad moral.

Volviendo a Hans, puede decirse que éste es responsable de no haber cumplido con sus *obligaciones epistémicas procedimentales*.⁵¹ Por la misma naturaleza de la creencia, Hans no tiene un control directo sobre su creencia (no puede cambiarla directamente), pues estas no son voluntarias,⁵² pero sí un control indirecto, que depende del correcto ejercicio de

⁵¹ Tomo esta expresión de Rosen (2004), p. 301.

⁵² No cabe duda de que el término “voluntario” es ambiguo. Su uso primario tiene que ver con nuestras acciones externas básicas, y su aplicación a los estados mentales y rasgos de carácter es derivada y más problemática. Aquí, su significado está cerca de la mera decisión (arbitraria): nuestras creencias serían voluntarias si igual que creemos que

nuestras habilidades epistémicas procedimentales.⁵³ Así, uno es indirectamente censurable por una actitud moralmente negativa que mantiene, debido a que no tomó los pasos directamente voluntarios necesarios para haber evitado su adquisición, ni para su conveniente revisión.

En particular, sus defensores más coherentes se comprometen con que, en algún punto de la historia causal, el agente realmente llevó a cabo una acción o decisión intencional y voluntaria, esto es, bajo su conocimiento y control, que le llevó a adquirir, culpablemente, el rasgo x por el que se le juzga responsable. Sin embargo, no está nada claro que, en la mayoría de los casos, esta acción o decisión intencional y voluntaria haya nunca tenido lugar. En particular, la acción o decisión por la que el agente es originariamente responsable no puede ser fruto de su mera ignorancia, pues entonces el agente sería, con mucho, derivadamente responsable por ella, y deberemos indagar más atrás en su historia causal para tratar de hallar una acción o decisión previa por la que el agente pueda ser considerado directa u originariamente responsable, una acción o decisión que resulte de su ignorancia culpable. Así, una condición necesaria (pero no suficiente) para que una persona sea moralmente censurable (o elogiable) sería la realización de una acción en la que la persona intencional y voluntariamente hace algo incorrecto (o correcto) por razones incorrectas (o correctas). Esto es, *una persona es censurable sólo si la persona cree que ha actuado mal deliberada e intencionalmente o ha contribuido a ello, o*

p , estuviera en nuestra mano creer que no p , en virtud de la misma evidencia; pero esto es psicológica y conceptualmente imposible. El lugar clásico al respecto es Williams (1973c); en Rosell (2008) ofrezco mi propia defensa de esta imposibilidad.

⁵³ Algo parecido sucede en el caso de las acciones. Por ejemplo, podría decirse que el conductor borracho es moralmente responsable por el daño que pueda causar en virtud de una negligencia previa, coger el coche estando borracho —o incluso anterior: decidir beber cuando podía prever que iba (a tener que) coger el coche. No obstante, situar el locus de la responsabilidad *meramente* en la negligencia es problemático, como veremos en el capítulo 7.

*cree que ha cultivado una disposición negativa, esté o no en lo cierto acerca de esta creencia.*⁵⁴

Hay aquí un presupuesto internista acerca de (la condición epistémica de) la atribución de responsabilidad moral: uno es responsable sólo en virtud de tipos de acciones que son (o fueron) internamente accesibles al agente de modo que por reflexión (reflexión acerca de las obligaciones morales propias) pudo (o podría) ver que debía realizarlas. Este internismo es una subclase del más general *internismo normativo*: una atribución de creencia justificada o responsabilidad por una acción es verdadera (o falsa) en virtud de la presencia (o ausencia) de ciertos hechos, estados o sucesos accesibles a la persona mediante reflexión, incluyendo la reflexión introspectiva, o obtenible a voluntad por la persona.⁵⁵

Sin embargo, en el caso de Hans, es muy dudoso que en algún momento cometiese un acto deliberado e intencional por el que adquiriese sus creencias antisemitas, a sabiendas de su perniciosidad moral (o incluso que dejase de mejorar, o deteriorase, su posición cognitiva intencional y voluntariamente). Por lo que, si una persona sólo puede ser censurada moralmente cuando sabe que aquello que hace o decide es moralmente incorrecto, entonces Hans no puede ser censurado. Hans no cree que haya nada moralmente incorrecto en su antisemitismo, más bien al contrario. Parece así que, aplicadas coherentemente, las condiciones epistémicas para la responsabilidad moral se tornan demasiado rigurosas.⁵⁶

⁵⁴ Ver, a este respecto, Zimmerman (1997). Holly Smith (1983) considera que lo que se requiere es que el agente cometa un “acto de ignorancia”, por el cual “deja de mejorar (o deteriora positivamente) su posición cognitiva”, de modo que crea el riesgo del acto incorrecto del tipo que precisamente ocurre después (p. 547).

⁵⁵ Ambas caracterizaciones son de Audi (1991), p. 316.

⁵⁶ Zimmerman (1997) reconoce que “las condiciones son bastante restrictivas”, con lo que “la ignorancia culpable tendrá lugar menos frecuentemente, quizá mucho menos frecuentemente, de lo que se supone comúnmente” (p. 411).

Por otro lado, también podríamos preguntarnos si aquello de lo que le culpamos es por tener creencias *erróneas* o por tener creencias *moralmente perniciosas*. Si fuera por lo primero, el tipo de culpabilidad atribuida sería principalmente epistémica. Pero lo que está aquí particularmente en cuestión es la responsabilidad moral. Hans no es meramente un mal agente epistémico, sino que además es un mal agente moral, alguien que merece ser moralmente censurado por las actitudes o creencias en cuestión. (No hay duda de que ambos tipos de responsabilidad están estrechamente conectados. De hecho, por lo que parecemos culparle es por tener creencias moral y epistémicamente erróneas.⁵⁷) En todo caso, lo fundamental es que si retrocedemos en la historia causal del agente en búsqueda de acciones y omisiones intencionales y voluntarias previas que podrían haber causado su actitud actual, no encontraremos más que otras *faltas involuntarias* anteriores —o mejor, *fallos involuntarios*, pues la falta parece implicar el actuar con conocimiento de la maldad del acto.

Parece, así, que en realidad nuestras atribuciones comunes de responsabilidad moral por actitudes y rasgos apuntan primariamente a los fallos morales; esto es, a la posesión misma de actitudes y rasgos moralmente negativos, y no a la de las acciones, decisiones y procesos anteriores por los que llegamos a adquirirlos —fuesen o no voluntarios, y existiesen o no en algún momento. Pero, antes de establecer esta conclusión, consideremos todavía otro ejemplo.⁵⁸ Imaginemos el caso de una joven que es ingrata con una persona con quien no debería serlo, pues ha hecho mucho por ella. Si indagamos de dónde procede esta ingratitud, veremos que ella no ha decidido nunca ser ingrata, ni ha aprobado esta actitud su-

⁵⁷ Un trabajo interesante sobre la conexión entre responsabilidad moral y epistémica es Montmarquet (1993).

⁵⁸ Tomo el ejemplo de Adams (1985), pp. 13-4.

ya, que muestra en su conducta y palabras. Puede que ni sepa que se comporta ingratamente. Parece que su deber sería acabar con este comportamiento ingrato; pero si no es consciente de ello, resulta incluso insensato pensar que debería luchar contra su ingratitud. ¿Cómo podría hacerlo, si ni sabe que es ingrata? Pero, con todo, *debería*; y la culpamos, por no hacerlo. Me explico. Internamente, dados sus estados mentales, ignora su actitud, por lo que no puede luchar contra ella y, en este sentido, no debe. Pero externamente, dado lo que consideramos correcto, y lo que ella consideraría correcto, y lo que ella creería que debe hacer si tuviese el acceso cognitivo necesario, sí que debe luchar contra su ingratitud. Así, como pensamos que es su obligación no ser ingrata, y que debería ser consciente de su misma conducta ingrata —y, en tanto que no es consciente de ello, su ignorancia es culpable—, la censuramos moralmente por ello. Esta consideración parece implicar que es una obligación moral la posesión de cierto grado de autoconocimiento —particularmente en relación a aquellos estados mentales y actitudes que tienen que ver con el (in)correcto trato de los demás.

Con todo, puede resistirse mi caracterización insistiendo en que lo que uno realmente censura (o debe censurar) es que los agentes no hayan tomado los pasos necesarios para eliminar estas actitudes ofensivas,⁵⁹ pues por lo que nunca deberíamos censurar a las personas es por meramente tener actitudes que no pudieron evitar tener. Ante el problema de por qué solemos censurar a la joven y a Hans, aunque puede que nunca cometieran ninguna *falta moral* que les haga responsables de tener las actitudes por las que les censuramos, uno puede ofrecer una explicación del siguiente tipo. Nuestra reacción, nuestro acto de censura, no se fun-

⁵⁹ Alternativamente, puede también defender que lo que realmente es censurable son las manifestaciones en el agente de sus actitudes negativas.

damenta en su culpabilidad, sino en un *propósito*: influir o provocar un cambio de actitud en ellos haciéndoles patente, por medio de nuestra *actitud reactiva negativa*, cuán contrarios somos a su actitud actual, debido a cuán moralmente incorrecta es (o nos parece). Sin embargo, parece que esto no se corresponde fielmente a nuestra experiencia del juicio moral: la influencia no es *siempre* el propósito, o el *único* propósito, por el que les censuramos.

Pero esta posición puede sofisticarse con un nuevo movimiento: articulando una especie de explicación disyuntiva del juicio moral. Así, cuando decimos que “Juan es moralmente censurable”, podríamos querer decir, o bien (a), que Juan es moralmente culpable, o bien (b), que la actitud de Juan debe ser reprobada para conseguir que la abandone o modifique, siendo el fundamento de (b) claramente consecuencialista. Esto funcionaría igual para el elogio: cuando elogiamos a alguien estamos, o bien (a) reconociendo sus méritos reales, o bien (b) reforzando (consecuencialmente) su actitud o conducta. Sin embargo, resulta que en la mayoría de casos no podemos saber cuál es el fundamento real de nuestro juicio de censura, puesto que (en la mayoría de casos) no tenemos el suficiente conocimiento de la historia del agente. Pero esto no debe quitarle el sueño al defensor de la explicación disyuntiva: no es cosa de su teoría decir cuándo hacemos una y cuándo otra, en cada caso particular, sino que simplemente la disyunción es verdadera. Esto es compatible con el hecho de que nuestro juicio aspira siempre a ser del tipo (a), pero el no tener un conocimiento detallado de la historia del agente, lo justifica al menos en virtud de (b). (No obstante, también dudamos de poder emitir un juicio del tipo (a) cuando tenemos un conocimiento muy detallado de su historia. Pues, a menudo, el mayor conocimiento de la historia personal de alguien, los diferentes avatares que le inclinaron a ir tomando los pasos que tomó;

cuando comprendemos por qué hizo cada cosa de las que hizo, podemos tender a excusarlo, ya que nos damos cuenta de cómo el curso mismo de los acontecimientos *le llevó* a hacer las cosas que hizo. Este fenómeno, particularmente presente cuando aquello por lo que se acusa a alguien es especialmente grave, es si duda un acicate para el escepticismo acerca de la responsabilidad moral.⁶⁰)

Esta explicación me resulta *prima facie* atractiva. No obstante, en casos como el de Hans, en el que llevamos a cabo una indagación minuciosa de la historia causal que le llevó a tener la actitud por la que le censuramos, y que no encontramos ningún obstáculo particularmente relevante para que Hans no hubiese adquirido esa actitud, parece que le censuramos y que lo hacemos en el sentido de (a) —en tanto que somos capaces de resistir el escepticismo anterior. Y aún cuando Hans no está presente, ni estamos ante nadie (en quien podamos querer reforzar su rechazo, o el del grupo, de este tipo de actitudes), aún así tendemos a pensar que Hans es culpable por su actitud. (Cabe aún la posibilidad de que queramos reforzar el rechazo de esta actitud en nosotros mismos. Pero entonces la distinción entre (a) y (b) se vuelve tenue; lo que puede que no sea muy descabellado.) Lo mismo parece ocurrir en el caso de la joven ingrata. Por el contrario, si negamos que ambos merezcan realmente ser censurados, resultará que no podemos censurar verdaderamente a casi nadie por sus rasgos negativos. Y, *mutatis mutandis*, elogiar verdaderamente a casi nadie por sus rasgos positivos.

⁶⁰ Pensemos en juicios por asesinato y en las reflexiones de los jurados que pueden imponer la pena de muerte. *A sangre fría* de Capote o el estudio que Watson (1987) hace del caso del asesino Robert Harris, son apasionantes en este sentido. Por otro lado, tenemos lo que podemos llamar la *paradoja del mal extremo*, esto es, el hecho de que cuanto más malvado es alguien y, por ello, más deseamos censurarlo, más difícil suele ser que éste satisfaga las condiciones para la responsabilidad moral.

Volvamos, pues, a la cuestión de las condiciones para el juicio de culpabilidad moral correcto —es decir, del tipo (a). La tesis que discuto es que nadie merece ser moralmente censurado por poseer rasgos negativos, a menos que haya *causado o permitido, intencional y voluntariamente, el desarrollo* de esos rasgos. Hemos visto que esta idea, que *prima facie* parece correcta, cuando la analizamos detalladamente, se vuelve muy problemática. Tomada el pie de la letra, nos conducía a la conclusión de que parece que finalmente es muy difícil censurar o elogiar a nadie por sus rasgos; lo que nos conduciría a un serio escepticismo o, al menos, agnosticismo.⁶¹

Evitar esta conclusión, podría exigir que *nos resignemos* a la tesis opuesta: uno merece ser censurado *por el mero hecho de tener esos rasgos negativos*. Por ejemplo, Robert Adams, refiriéndose a un joven nazi formado en las juventudes hitlerianas, afirma:

No importa cómo llegue a tenerlas, sus creencias malvadas son una parte de lo que él es, moralmente, y lo hacen un objeto de reproche adecuado. Puede también ser víctima de su educación; y si lo es, eso da una razón para tratarlo con merced —pero no es una exención de la censura.⁶²

Podría ser que finalmente nos viésemos obligados a aceptar una posición de este tipo, si no hay otro modo de salvar el grueso de nuestras prácticas cotidianas de juicio moral. No obstante, para defender las dudas que aquí planteo acerca del rastreo de la condición epistémica, no es necesario aceptar una posición tan drástica.

⁶¹ Rosen (2005) defiende que, en último término, sólo podremos considerar culpable al acrático, lo cual es un resultado intuitivamente chocante.

⁶² Adams (1985), p. 19.

Hay una posición más moderada, intermedia, según la cual *uno es moralmente censurable por sus actitudes y rasgos negativos aunque no tomase ninguna decisión intencional y voluntaria al respecto, sino sólo en virtud de no haber evitado el desarrollo y solidificación de los hábitos que subsecuentemente le llevaron a adquirir la actitud porque se le juzga o que le impiden incluso darse cuenta de que está haciendo mal manteniendo esa actitud*. Obviamente, si queremos evitar el resultado anterior, esta posición no puede defenderse sobre la base de que el agente deliberadamente reforzó en sí mismo tal hábito, anticipando sus consecuencias y teniéndolo como incorrecto. Las creencias antisemitas de Hans, ni mucho menos la ingratitud de la joven, no parecen deberse a ninguna disposición deliberadamente formada por uno mismo, a ningún acto de voluntad.

(La forma más clara de internismo es el volicionismo, según el cual nuestros actos básicos son voliciones o intentos. Este puede también combinarse con diferentes teorías morales; el candidato más obvio es la moral kantiana, según la cual la reflexión puede revelarnos nuestras obligaciones morales. Sin embargo, también puede combinarse con la ética de la virtud de inspiración aristotélica, caracterizando voluntaristamente la adquisición de hábitos y virtudes.⁶³ Alternativamente, el internismo también podría ser caracterizado de un modo no voluntarista: el agente es moralmente responsable por lo que le es internamente accesible y obtenible por reflexión, pero sin la necesidad de tener que retrotraer causalmente el origen de la responsabilidad a un acto interno.⁶⁴)

⁶³ Jacobs (2001) defiende una concepción voluntarista de la habituación. Trianosky (1990) y Sher (2006) critican y rechazan la concepción aristotélico-voluntarista de la formación del carácter.

⁶⁴ Audi (1991): “La acción culpable básica no tiene porqué ser un acto de la voluntad, incluso si mi responsabilidad por él depende de mi acceso interno a un acto de la voluntad compensatorio” (p. 318).

Así, contrariamente, mi posición es que *el agente es culpable porque simplemente adquirió y reforzó en sí el hábito por el que la adquisición de cierta actitud que no le parece incorrecta, cuando pudo no haberlo hecho, aunque no es claro en qué medida tuvo el control efectivo para ello*. No obstante, en los casos en los que definitivamente el agente no tuvo ninguna opción razonable de evitar la adquisición o de conseguir deshacerse del rasgo en cuestión, su responsabilidad moral por ese rasgo quedará cancelada —al menos desde la perspectiva del *juicio moral de tercera persona* (que es el que exclusivamente considero en este trabajo).

Por un lado, esta posición se basa en la idea de que hay consideraciones morales cuya fuerza racional no puede ser ignorada por ningún agente autónomo, o moralmente responsable. Esto es, hay normas y consideraciones morales de especial relevancia que todo agente autónomo tiene la obligación de conocer, por lo que su ignorancia al respecto es siempre culpable. En este sentido, mi posición es externista en relación a la condición epistémica: soy responsable por mis creencias moralmente incorrectas con independencia del mi acceso interno a su incorrección moral. El agente puede no tener acceso reflexivo al hecho de que, dada la existencia de ciertas exigencias morales, su carácter necesita un cambio. Según el internista, esto cancelaría su responsabilidad. Por el contrario —es mi posición—, el agente tiene la obligación (externistamente) de cambiarlo, y es culpable por no hacerlo. La cuestión es en qué sentido no le son accesibles.

Por otro lado, lo que aparece como crucial en este punto son las posibilidades efectivas, dependientes del control del agente, para impedir la formación o modificar los estados mentales por los que es censurado. No obstante, en la mayoría de los casos, la misma posibilidad real de que tomase los pasos necesarios para eliminar sus actitudes negativas depen-

dió, en gran parte, de factores ajenos a su control. (Puede que el único modo, o el más efectivo, de que la joven ingrata se dé cuenta de que es ingrata sea gracias a la censura de alguien ante su ingratitud. Quizá no está de más recordar aquí una obviedad: el juicio de censura moral está inmerso en nuestras mismas prácticas, no es ni externo a ellas ni el final de la historia; cómo reaccionemos a él también es muy relevante.) Es posible, incluso, que no exista una respuesta única a cuán grande es la posibilidad de disentir o de resistir la opinión unánime, en la sociedad, en un grupo o en un colectivo; o cuál es la fuerza máxima de una tentación, o cuán resistente un rasgo heredado o involuntariamente adquirido, para esperar razonablemente que el agente supere la situación. De hecho, puede que la misma respuesta a esta cuestión sea necesariamente imprecisa —a causa, posiblemente, de la misma naturaleza imprecisa de la ética.⁶⁵

En todo caso, la conexión entre las situaciones en las que uno podría haber hecho por no ser cómo luego fue y cómo finalmente resultó ser, no es nada transparente. Puede incluso que el agente no dispusiese de un *control efectivo*. Pero dejar, por ello, de tenernos por agentes moralmente responsables no parece una alternativa plausible. De hecho, los casos considerados muestran cómo estamos inclinados a ver a los demás (y a nosotros mismos) como responsables por rasgos y actitudes personales, sobre los que el control del agente es ambiguo.

⁶⁵ Esta idea se encuentra ya en Aristóteles: “es propio del hombre instruido buscar la exactitud en cada género de conocimientos en la medida en que la admite la naturaleza del asunto; evidentemente, tan absurdo sería aprobar a un matemático que emplea la persuasión como reclamar demostraciones a un retórico.” (EN I, 3; 1094b-1095a.) También Williams ha incidido en esta idea; véase esp. Williams (1985).

5. 5. Carácter y evaluación. Conclusiones

En la parte final de este capítulo, quiero abordar directamente la cuestión misma de la evaluación o atribución de responsabilidad moral por el carácter del agente, con el fin de clarificar algo más la posición que he defendido, así como situarla en un contexto más amplio. Hasta ahora, hemos simplemente asumido que es apropiado atribuir directamente responsabilidad moral por el carácter (por ciertos estados mentales: creencias, actitudes, deseos, etc.). Sin embargo, la posición ortodoxa afirma que el objeto primario de evaluación moral son las acciones. La razón principal para no evaluar directamente los rasgos y el carácter de una persona es que la posesión y desarrollo de unos rasgos de carácter u otros es algo que depende en exceso de factores externos al control del agente, como son su constitución original y los avatares formativos.⁶⁶

Sin embargo, cotidianamente también juzgamos a las personas por sus mismos rasgos de carácter. Alguien que es envidioso, pero que evita comportarse envidiosamente en una situación determinada, no deja por ello de ser envidioso y parece que se le censura en ocasiones por el hecho mismo de su disposición a la envidia. (Aunque, por supuesto, si alguien logra refrenar su tendencia a la envidia merece un elogio que no merece el que se deja llevar por ella.)⁶⁷ Es decir, los rasgos internos pueden ser el objeto *directo* de juicio moral. Además, algunos adversarios de la suerte moral, como Richards o Rescher (defensores, recordemos, de lo que llamé la Estrategia Moderada) afirmaban que el locus apropiado del merecimiento moral es el carácter del agente.

⁶⁶ En último término, la idea es que no podemos censurar a nadie por sus rasgos, en tanto que uno no tiene control sobre sus rasgos y nadie merece ser censurado por algo que está más allá de su control. Véase Sher (2001) para un análisis de este argumento y las opciones para resistirlo.

⁶⁷ Más sobre esta cuestión en el capítulo siguiente.

Alternativamente, la posición ortodoxa apela a la estrategia de distinguir entre dos tipos de juicios: los *juicios aretáicos*, que evaluarían el carácter del agente, sus rasgos (más ajenos al control inmediato), y los verdaderos juicios de responsabilidad moral, que tendrían que ver con el elogio o laudabilidad y la censura o culpabilidad moral de las personas, cuya base serían voluntades o decisiones explícitas del agente.⁶⁸

Cabe remarcar que estos dos tipos de evaluación parece que se corresponden con *las dos caras de la responsabilidad* de Gary Watson.⁶⁹ Es conocida la distinción de Watson entre los que considera dos aspectos relativamente diferenciados de la responsabilidad moral: una cosa es considerar responsable a alguien en el sentido de *juzgar sus actos como expresión de sus valores o compromisos* [*attributability*] y otra considerarle responsable en el sentido de *juzgarle censurable o elogiabile* [*accountability*]. Lo primero, el aspecto aretaico o atribucional, estaría íntimamente conectado con una concepción autorreveladora de la responsabilidad moral; de quién se es y de la propia posición en relación a cuestiones de valor. La segunda cara de la responsabilidad es la que Watson llama el aspecto de *rendir cuentas*. Una persona ha de rendir cuentas por su acción si hay sanciones o beneficios que se aplican justamente a ella como consecuencia de su acción.

Sin embargo, reservar los términos censurable y elogiabile para el segundo tipo o aspecto de la responsabilidad implica ya privilegiar éste

⁶⁸ Como ya avancé en el capítulo 3, la meta de esta distinción es deshacerse de la suerte moral (constitutiva) restringiéndola a los juicios aretaicos, de modo que queden libres de suerte constitutiva los de responsabilidad moral, que serían los únicos relevantes para determinar el merecimiento moral del agente.

⁶⁹ Watson (1996).

último aspecto sobre el primero.⁷⁰ De hecho, esta parece la posición más estrictamente kantiana. Kant distinguía entre *temperamento* (rasgos heredados) y *carácter* (rasgos desarrollados voluntariamente) y afirmaba que el agente sólo es responsable por lo segundo y no por lo primero. Ahora bien, también es cierto que el carácter mismo incorpora aspectos que escapan asimismo al control del agente. Por ello, Kant añade:

las excitaciones de sus inclinaciones e impulsos (y, por tanto, la naturaleza entera del mundo sensible) no pueden menoscabar las leyes de su querer [de la persona] como inteligencia, *hasta el punto de que no se considera responsable de esas inclinaciones e impulsos y no los atribuye a su propio yo*, esto es, a su voluntad, aunque sí es responsable de la complacencia que pueda manifestarles si les concede influjo sobre sus máximas, con perjuicio de las leyes racionales de la voluntad.⁷¹

Así, aquello por lo que se es realmente responsable, según Kant, sería sólo la parte del carácter que es fruto de decisiones, pero no por los deseos e inclinaciones o, en general, por cualquier cosa meramente experimentada. Respecto a éstos, se es responsable solamente por cómo se elige responder a ellos.⁷² En última instancia, aquello por lo que uno es respon-

⁷⁰ Aunque Watson, oficialmente, defiende que son dos aspectos igualmente importantes de nuestra concepción de la responsabilidad moral. Esta, a mi modo de ver, es la posición correcta. En general, si llanamente distinguimos entre juzgar a alguien por lo que es expresión de sus valores o compromisos y juzgarlo sólo por aquello que es resultado de sus elecciones voluntarias, parece claro que los defensores de la Estrategia Moderada consideran el primer aspecto como fundamental, mientras que los que abogan por la Estrategia Radical priman el segundo.

⁷¹ *Fundamentación*, § III (Ak. IV: 457-58). Traducción de García Morente, ligeramente modificada.

⁷² Aunque la rechace, Angela Smith (2005) considera que el atractivo de esta distinción se halla en que “nos permite reconocer los modos en que nuestra naturaleza biológica, nuestra cultura y las condiciones en que fuimos criados pueden servir para formar nuestros patrones de evaluación y juicio. Ya que muchos de estos patrones fueron adquiridos antes de la “edad de la razón”, puede parecer inapropiado tomar estas cosas como base de la estimación moral [*moral appraisal*] de una persona. En cambio, debemos centrarnos en cómo la persona adulta elige lidiar con esas propensiones, más que en las propen-

sable se limitaría al “conjunto de compromisos con, o la continuada aprobación de, principios morales por la voluntad (racional)”.⁷³

En esta dirección, por ejemplo, Michael Zimmerman (representante principal, recordemos, de la Estrategia Radical de negación de la suerte moral) considera que la “evaluabilidad” [*appraisability*], esto es, la evaluación de la valía moral, del merecimiento de ser interiormente elogiado o censurado, recae sobre el *yo* del agente, y no sobre su carácter. La evaluación del carácter produce la mera admiración o reprensión. Por ejemplo, Claggart, personaje de la novela *Billy Budd* de Herman Melville, es un ser malvado, que nos produce una reprobación total; pero en tanto que su carácter es el resultado de su misma naturaleza, no es censurable como lo es Dorian Gray, quien ha formado autónomamente su carácter. Aunque posean los mismos estados mentales, Dorian es responsable y Claggart no, porque la maldad del primero depende de sí mismo, mientras que el segundo no es malvado libremente.⁷⁴ En la misma línea, Neil Levy defiende que cabe mantener que la evaluación de las actitudes consiste simplemente en la atribución de cualidades al agente, que son buenas o malas, admirables o repugnantes. “Después de todo, parece que *prima facie* debe haber espacio conceptual para esa evaluación. Debemos poder decir que algo es *malo* sin decir que es *censurable* [*blameworthy*].”⁷⁵ En resumen, una cosa es ser malvado y otra censurable por ello.

siones mismas, que pueden deberse a afectos y experiencias tempranos sobre los que la persona no tuvo control racional o volitivo. Esta perspectiva informa muchos de los libros de autoayuda, así como las tendencias de terapia en boga hoy en día, que insisten en trazar una marcada distinción entre cómo uno “siente” y cómo “actúa” (siendo sólo lo último aquello por lo que se es responsable.” (p. 266). Sin embargo, el mismo control racional y volitivo depende de “nuestra naturaleza biológica, nuestra cultura y [de] las condiciones en que fuimos criados”.

⁷³ Trianovsky (1990), p. 104.

⁷⁴ Zimmerman (2002), pp. 556-7.

⁷⁵ Levy (2005), p. 5.

Sin embargo, la posición que defendí al final de la sección anterior no es la de que uno es culpable meramente por poseer unos estados mentales objetables, sino que además debe ser un agente autónomo y haber tenido alguna posibilidad de modificar estos estados mentales por los que es juzgado, por bien que tal posibilidad permanezca esencialmente indeterminada.⁷⁶ De todas maneras, el caso de Claggart no es un buen caso, pues Claggart parece ser como JoJo, alguien que no satisface los criterios mínimos de *cordura* —por utilizar la expresión de Wolf— y, por lo tanto, no es un agente autónomo o candidato adecuado para la atribución de responsabilidad moral.

Volviendo a Kant, una concepción tan pobre y restringida del carácter y de nuestra responsabilidad por él —como hemos visto a lo largo de este capítulo— resulta, en primer lugar, totalmente insatisfactoria como descripción de nuestras prácticas cotidianas de evaluación moral. En segundo lugar, desatiende completamente los procesos de habituación y de adquisición real de creencias, deseos y otras actitudes, que no son tipos de volición y que sólo indirectamente dependen de la voluntad. Y, finalmente, establece una exclusión injustificada (e ininteligible) de la propia facultad racional o Voluntad de los avatares formativos y empíricos, en general, en los que se mueve el resto de los atributos mentales. Como hemos podido comprobar, el carácter es en realidad mucho más que la continuada aprobación de los principios morales por parte de la voluntad racional. Es también el complejo conjunto de actitudes, deseos, emociones, hábitos, valores, etc., cuya relación con la voluntad es significativamente diversa y, en todo caso, ésta última no siempre tiene la última palabra.

⁷⁶ Retornaré a la distinción entre bueno/malo y elogiable/censurable en el capítulo 8.

Por el contrario, David Hume afirmaba que uno sólo es responsable por aquellas acciones que reflejan su carácter o alguna de sus cualidades mentales duraderas, rechazando a su vez que el carácter dependa de una elección o proceso de adquisición voluntario. Así, además de situar en el carácter el locus originario de la responsabilidad moral, rechaza su limitación a la esfera de lo voluntario. Las personas son consideradas responsables por sus cualidades mentales o carácter, incluso aunque se opongan manifiestamente a su voluntad. Por lo que la voluntariedad no sería un requisito de todos los casos de atribución de responsabilidad moral.⁷⁷

En esta línea, pero más matizadamente, diferentes autores contemporáneos han defendido que para que un estado mental pueda ser correctamente atribuido a la persona (como fundamento legítimo para la evaluación moral) basta con que pertenezca a la clase de actitudes que (en agentes idealmente racionales) son sensibles al juicio, aunque nunca haya sido aceptado ni considerado. El criterio para establecer por cuáles de mis rasgos y estados mentales soy responsable es que se me pueda requerir que los justifique, que los reconozca y defienda o desaprobe. No obstante, la completa ausencia de control puede bloquear la atribución de actitudes. No tendría sentido que me pidieran que justificara mi altura, el color de mi piel o mis compulsiones, sí que lo tiene que justifique mis creencias políticas, mi deseo de que mi amigo consiga superar una enfermedad o mi actitud hacia ti. En tanto que somos agentes racionales, nuestras actitudes sensibles al juicio son el producto y la expresión de nuestra identidad como agentes prácticos, revelan dónde nos situamos en cuestión de valores.

⁷⁷ Hume, *Tratado de la naturaleza humana*, §§ 2.3.1-2 y 3.3.; así como *Investigación del entendimiento humano*, § 8. Para una interpretación naturalista de la posición de Hume acerca de la responsabilidad moral, véase Russell (1995).

Esta perspectiva desborda la propuesta basada en la voluntad en el siguiente sentido: no expresamos nuestra agencia moral y actividad sólo en nuestras elecciones y decisiones explícitas, sino también por medio de lo que irreflexivamente pensamos, sentimos y deseamos. Somos también responsables por los juicios implícitos con que respondemos al mundo que nos rodea (en tanto que susceptibles de justificación).⁷⁸

Parece claro que ambas posiciones tratan de alcanzar la misma meta, a saber, que nuestras atribuciones de responsabilidad moral sean justas, en el sentido de que toquen algo central en el agente. Sin embargo, ofrecen soluciones contrarias, pues sitúan el locus último de éstas en aspectos diversos del agente: o bien sus elecciones voluntarias o bien sus cualidades mentales duraderas, que no son necesariamente coincidentes. Es decir, puede que una acción o actitud sea fruto de una decisión voluntaria que surge de nuestros valores y compromisos, esto es, de cualidades mentales duraderas; pero también es posible que nuestra voluntad adopte decisiones que no responden a nuestro carácter o que nuestro carácter incluya elementos que rechazamos. El problema reside, a mi entender, en que ambas posiciones se presentan como excluyentes: hay algo que es el locus último o fundamental de la atribución de responsabilidad moral; pero esto es simplemente falso, diferentes aspectos de mi yo son igualmente relevantes para la evaluación moral, y igualmente “últimos” (con muchas comillas, por su puesto). El error reside en pensar esta cuestión en términos fundamentalistas. En realidad, sólo contextualmente un tipo de

⁷⁸ Esta es la posición de Scanlon (1998), cap. 1; y A. Smith (2005), pp. 256-7 y *passim*; que Levy (2005) ha llamado *atribucionismo*. Hay, no obstante, una clara ambigüedad entre sí para que un rasgo pueda ser correctamente atribuido a alguien es necesario que esté integrado en sus sistema de valores o simplemente que sea sensible al juicio. A mi parecer, la opción elegida debería ser la segunda.

evaluación resulta más relevante que el otro. Por ejemplo, tras hablar bruscamente y gritar a alguien, puedes decir: “perdón, es que es mi forma de hablar, no *quería* hablarte mal”. Aquí puede pesar más la voluntad, lo que quieres, que cómo eres. Pero el caso contrario es igualmente posible. Si tu amigo te dice: “¡me marchó, estoy harto de lo que haces!” , otro amigo puede decirte que el no es realmente así. Ahora, lo fundamental sería, más que su decisión voluntaria, cómo es. (Esto funciona completamente al contrario cuando se quiere acusar a alguien.) Y, por otro lado, ambos aspectos están irremediable confundidos.

Es decir, nuestra idea de responsabilidad moral incluye diferentes tipos de juicio (dos, al menos) que son igualmente fundamentales. En concreto, uno u otro aspecto será más o menos importante o fundamental en relación a un contexto determinado de evaluación. Así, una conclusión a la que podemos llegar tras lo dicho, que dejaré sin más defensa por el momento, es que las atribuciones de responsabilidad moral son contextualmente sensibles. Además, la cuestión de si las atribuciones de responsabilidad moral han de recaer sobre elecciones voluntarias o sobre el carácter del agente volverá a surgir en los dos capítulos siguientes.

En definitiva, la conclusión a este capítulo es evidente: la pretensión de eliminar por completo los factores incontrolables envueltos en la constitución personal y en las circunstancias de su desarrollo y formación del reino de la responsabilidad moral, no es más que una vana pretensión, condenada al fracaso. No es aventurado afirmar que se trata, simplemente, de un hecho acerca de la condición humana, de cómo adquirimos nuestra capacidad para la responsabilidad moral en el mundo real. Y, en todo caso, la suerte no es sólo un factor inevitable en la constitución y formación del agente, sino que resulta imprescindible.

Cabe remarcar, finalmente, que el mismo estudio de la suerte y de cómo está presente en nuestra constitución y se inmiscuye en nuestro desarrollo y formación como agentes autónomos resulta altamente relevante en el afán por alcanzar una comprensión más exacta y realista de nuestro carácter moral y de los juicios de responsabilidad moral acerca de él.

Dónde estamos y adónde vamos

Con lo anterior, la meta de conseguir un rechazo general de la suerte moral en todos sus aspectos (bien mediante el Caso global solamente, o con la colaboración de la negación de la coherencia de la noción de suerte constitutiva) ha quedado ya desacreditada. No obstante, todavía es posible atacar individualizadamente los tipos circunstancial y consecuencial de suerte moral. De hecho, los argumentos contra la suerte moral parecen especialmente efectivos en relación a los tipos circunstancial y resultante. No es de extrañar, así, que la mayor parte de los participantes en el debate sobre la cuestión se hayan centrado en estos dos tipos, algunos incluso ignorando los tipos antecedentes. Otros abogan por una estrategia híbrida: aceptar la suerte antecedente pero negando la circunstancial y la resultante.

En todo caso, los tipos de suerte moral circunstancial y consecuencial merecen una discusión individualizada. Es lo que haré en los dos capítulos siguientes, dedicados respectivamente a cada una de ellas.

6. Suerte circunstancial

Atrapados por las circunstancias

- 6.1. El rechazo de la suerte circunstancial
 - 6.1.1. El Principio de Control
 - 6.1.2. Carácter y acción
- 6.2. Disposiciones y la fuerza de las circunstancias: el reto situacionista
 - 6.2.1. Los experimentos de Milgram sobre la obediencia
 - 6.2.2. Otros experimentos situacionistas
 - 6.2.3. Repercusiones: escepticismo acerca de los rasgos de carácter
 - 6.2.4. Resistir el eliminativismo
- 6.3. Circunstancias y evaluación
 - 6.3.1. Intenciones, planes e indeterminación. El significado de la suerte circunstancial
 - 6.3.2. Más repercusiones: la opacidad de los propios motivos e intenciones
 - 6.3.3. Sobre los dilemas morales

No sé si en mis profundidades acecha un asesino, ni me interesa mucho saberlo, pero sí sé que he sido víctima inocente y no asesino. Sé que los asesinos existen... y que confundirlos con sus víctimas es una enfermedad moral, una afectación estética o un signo siniestro de complicidad; pero por encima de todo es un precioso servicio que (intencionalmente o no) se presta a quienes niegan la verdad.

Primo Levi, *Los hundidos y los salvados*.

El resultado de los dos capítulos anteriores ha sido que la suerte moral constitutiva y formativa no son fenómenos ilusorios, sino que son reales. Tanto la influencia de los desiguales rasgos constitutivos, como las diferencias en la formación, son fundamentales para las posibilidades de desarrollo del propio carácter. Estos factores, sin duda, condicionan los rasgos y todo tipo de actitudes del agente, incluyendo sus intenciones, decisiones

y acciones. Con ello, la pretensión de excluir por completo de la esfera de la responsabilidad moral todos aquellos factores incontrolables o incontrolados envueltos en la constitución personal y en las circunstancias que rodearon la formación del propio carácter —aunque loable—, no puede ser más que una pretensión vana, condenada al fracaso. Se trata simplemente de un hecho acerca de la condición humana, de cómo adquirimos nuestra capacidad para la responsabilidad moral en el mundo real.

Sin embargo, uno puede admitir este hecho y negar todavía los tipos circunstancial y resultante de suerte moral: ciertamente, afirma este razonamiento, cada cual es fruto de un origen y de unas experiencias y avatares formativos, pero lo que realmente importa para el juicio moral es la voluntad o intenciones actuales del agente. De hecho, muchos de los que se han ocupado de la cuestión de la suerte moral se han limitado a abordar estos dos tipos de suerte moral —ignorando o aludiendo tímidamente a los otros—, en tanto que son los tipos más característicos de esta cuestión.

En el capítulo presente me ocuparé específicamente de la suerte moral circunstancial y de los argumentos dirigidos contra ésta. Para ello, no tendré más remedio que recuperar argumentos y distinciones que ya vimos en el capítulo 3, pero que ahora cabrá desarrollar específicamente en relación al tema de la suerte circunstancial. En concreto, someteré a consideración la opción de situar el locus último de la responsabilidad moral en los rasgos de carácter del agente, como el modo de evitar la suerte circunstancial (y resultante). Para evaluar esta posición será fundamental examinar lo que podemos llamar *la dialéctica carácter-acción*, mediada por la *situación*. En este contexto, presentaré y discutiré también una serie de experimentos psicológicos pertenecientes a la tradición de investigación conocida como *situacionismo*. Consideraré y rechazaré la

tesis de que estos experimentos arruinen la noción de rasgo de carácter, pero mantendré que los resultados suponen un desafío importante al autocontrol y autoconocimiento del agente y, con ello, a la responsabilidad moral. En general, defenderé la importancia de la acción en la misma determinación del carácter y la necesidad de considerar intenciones particulares y acciones efectivas en la evaluación del agente. Si esto es así, las circunstancias marcarán necesariamente diferencias morales.

Antes de pasar a ello, cabe añadir que, para defender mi posición, en este (y el siguiente) capítulo no apelaré ya a factores antecedentes, en el estado del mundo y el sujeto, anteriores al momento en que se presenta un caso determinado de suerte moral —lo cual no sería más que una repetición o ampliación de los argumentos de los capítulos anteriores; en especial del capítulo 3—, sino a la diferencia que puede suponer que suceda una u otra cosa a partir del momento en consideración. Es decir, me centraré en *las diferencias en los hechos posteriores, que los negadores de la suerte moral excluyen de la evaluación del agente*. Esto se entenderá mejor en breve.

6.1. El rechazo de la suerte circunstancial

6.1.1. El Principio de Control

Consideremos un caso positivo de suerte circunstancial, es decir, un caso por el que uno puede merecer ser elogiado o no debido a factores circunstanciales más allá de su control. Es este. Diversos operarios del ayuntamiento están pintando las rayas del carril bici que discurre junto a una gran avenida. En un momento determinado, un niño irrumpe en la calzada. A continuación, y repentinamente, el niño se da cuenta de que, a unos doscientos metros, los coches comienzan a acelerar, debido a que el

semáforo ha cambiado a verde. El niño se queda paralizado. Su padre le grita desde el otro lado. Uno de los operarios (Operario 1) decide ir al rescate del niño, pero otro operario (Operario 2), que se encontraba más cerca del niño, se le adelanta. Operario 2 logra devolver al niño a la acera antes de que lleguen los coches. Al día siguiente, Operario 2 recibe de manos de la alcaldesa una condecoración municipal y diferentes periódicos locales lo entrevistan y publican su foto en portada. Operario 1 perdió su oportunidad de ser elogiado y de convertirse en el *héroe* local que por unos días será Operario 2, debido a la mala suerte de encontrarse más lejos que su compañero del lugar donde el niño se paró y permaneció inmóvil. En este tipo de casos, parece vano decirle a Operario 1 que “bueno, tú también lo habrías hecho si hubieras estado más cerca de él”. Más aún, parecería del todo inadecuado que Operario 1 insistiera en que él merece el mismo elogio y reconocimiento que ha recibido Operario 2.¹ Parece evidente que un rasgo fundamental para el elogio aquí es salvar realmente al niño, y no meramente *querer salvarlo*.

Recordemos que Nagel afirmaba que “[j]uzgamos a las personas por lo que de hecho hacen o dejan de hacer, no sólo por lo que habrían hecho si las circunstancias hubieran sido diferentes.”² Por supuesto, los negadores de la suerte moral se oponen frontalmente a esta conclusión: si lo que Nagel quiere decir es que no juzgamos a las personas por lo que habrían hecho, se equivoca meridianamente. De hecho, como vimos en el capítulo 3, lo que deberíamos juzgar, según ellos, sería *lo que el agente habría hecho*.

¹ Para ser justos, existe una asimetría entre la autoevaluación y la evaluación de otros (entre la primera y la tercera persona) que impide o hace muy implausible una reclamación de este tipo. En el capítulo siguiente, veremos cómo funcionan y cuáles son las repercusiones de este tipo de asimetría.

² Nagel (1979), p. 34.

Consideremos ahora un caso de presunta suerte circunstancial en relación a la responsabilidad negativa y su evaluación desde la perspectiva de los negadores. El juez Real y el juez Contrafáctico trabajan juntos en un juzgado. Ambos son igualmente corruptos: ambos aceptarían un soborno si se les ofreciera y fuera suficientemente importante, pero hasta hoy nadie les había hecho ninguna propuesta de este tipo. En un juicio del que se ocupa el juez Real, el acusado le ofrece, por medio de intermediarios, una cantidad importante de dinero por fallar a su favor, y éste la acepta gustosamente. Si el juez hubiera sido Contrafáctico, éste hubiera aceptado igualmente el soborno, pero debido a los horarios y distribución de los juicios, fue el juez Real y no el juez Contrafáctico a quien le correspondió el juicio en cuestión, por lo que fue sólo él quien recibió la oferta y quien aceptó el soborno.³ Para los negadores de la suerte moral el juicio moral que ambos merecen es, sin duda, el mismo.

Pero consideremos aún otro caso: el de dos personas que pretenden cometer un asesinato. George dispara a Henry y el disparo acaba con su vida, mientras que cuando Georg se dispone a disparar a Henrik, un camión se sitúa en su línea de tiro impidiéndole disparar. Con ello Georg pierde la oportunidad de actuar como quería. No obstante, Georg tiene en común con George que, de no haber ocurrido este hecho, también él habría libremente escogido disparar sobre Henrik. La conclusión, de nuevo, para los negadores de la suerte moral es que uno es tan culpable como el otro.⁴ Pero, en realidad, ¿de qué se supone que es culpable Georg? ¿De asesinato?, si no mató a nadie. ¿De intento de asesinato? Pero el mero intento de asesinato parece menos grave que el asesinato. Si bien el negador de la suerte moral puede, estrictamente, limitarse a decir que moral-

³ El caso es de Thomson (1989), pp. 206-7.

⁴ Ejemplo de Zimmerman (2002), p. 563.

mente ambos son igualmente culpables, sin más (que un camión se situase en la línea de tiro de Georg, salvando la vida Henrik, no es excusa para aquél); parece que nos debe una explicación positiva de qué es aquello que hace que ambos, supuestamente, sean igual de culpables. Un obvio candidato aquí son las intenciones de ambos: tanto George como Georg actuaron guiados por la intención efectiva, anclada en la configuración real de su carácter, de querer acabar con la vida de sus respectivas víctimas. El locus fundamental de la responsabilidad moral estaría en una cualidad mental duradera de los agentes.

Sin embargo, Michael Zimmerman, interpretando este caso, rechaza que el locus de la responsabilidad se halle en la manera como el agente es, en su carácter, ni en su intención. Incluso la realización concienzuda de todos los preparativos necesarios para que el intento resultase exitoso no agota su responsabilidad. En realidad, si Georg es responsable *en virtud de* algo, lo es meramente en virtud de que habría matado libremente a Henrik, si hubiera tenido la cooperación de ciertos rasgos del caso. No hay *nada* de lo que podamos decir que Georg es responsable, ya que no mató a nadie. Además, es dudoso que Georg controlara el hecho de estar en la situación de querer matar a Henrik. Es, simplemente, “responsable *tout court*”.⁵ Según Zimmerman, Georg y George tienen el mismo *grado* de responsabilidad porque ambos son responsables en virtud del mismo hecho: el hecho de que cada uno habría matado libremente a alguien, si hubiera tenido la cooperación —que George tuvo y Georg, no— de ciertos rasgos del caso, por bien que el *alcance* de la responsabilidad de Georg quede diluido a nada.

⁵ Zimmerman (2002), pp. 564-5. Según éste, con ello se enfatiza aún más drásticamente la distinción entre el grado (*degree*) de la responsabilidad y su alcance (*scope*), a la que ya me referí en el capítulo 3.

Creo decididamente que esta manera de conceptualizar el caso que propone Zimmerman confunde más que aclara. Sin embargo, las intuiciones que pone aquí en juego son claras, y es justo atender a ellas. En la sección siguiente intentaré reelaborar el tipo (o tipos) de explicación positiva de la atribución de responsabilidad en los casos de suerte circunstancial a que el negador puede apelar. En todo caso, cabe adelantar que el marco general lo constituye, sin lugar a dudas, la idea de que las circunstancias no pueden marcar una distinción en el merecimiento del agente. Y en ayuda de esta idea vienen tanto el *argumento epistémico*: en todos los casos considerados la suerte afecta a nuestro conocimiento del merecimiento moral del agente, pero no a su merecimiento moral como tal;⁶ como la idea de que una persona merece ser tratada de una cierta manera, no por lo que ha hecho, sino por lo que *es plausible pensar que habría hecho si hubiera tenido la oportunidad*.⁷ En el capítulo 3, repliqué que no podemos abusar demasiado de esta última idea separando en exceso sus condiciones contrafácticas de aquellas en las que de hecho actúan los agentes; es decir, apelando a mundos posibles demasiado lejanos. Pero, evitados los excesos, la idea sí que podría suponer una solución adecuada para los casos de suerte moral circunstancial.

Antes de seguir adelante, cabe decir que el último ejemplo de suerte circunstancial —el de George y Georg— difiere del caso anterior de los dos jueces, y más aún del de los dos adeptos al nazismo —ejemplo canónico de suerte circunstancial, como sabemos—, que abordaremos en seguida. En general, cabe distinguir entre casos de suerte circunstancial en los que la acción del agente es *inmediata* o está muy *próxima* de llevarse

⁶ Véase capítulo 3, nota 1; sobre sus principales defensores.

⁷ Ver Richards (1986), pp. 173-4. Estos son los dos rasgos más fundamentales de la que llamé Estrategia Moderada, en el capítulo 3.

a cabo, y de producir o no los resultados que, en última instancia, repercutirían en el juicio moral; y casos más *distantes* de suerte circunstancial, en los que hay una separación temporal más amplia entre la situación que marca la diferencia entre los agentes y el momento posterior de la acción. En este sentido, los primeros se asemejan más a (o están más cercanos de) casos de suerte resultante; mientras que los segundos se acercan más a la (ya considerada) suerte antecedente o *formativa*. Esto, sin embargo, no es óbice para el reconocimiento de que la suerte circunstancial constituye por sí misma un tipo de suerte moral de pleno derecho; sino que, más bien, muestra —como ya dije en el capítulo 2— los límites borrosos entre los diferentes tipos de suerte moral, lo cual arruina la idea de que los diferentes tipos de suerte moral puedan constituir unidades homogéneas y aisladas.

En este sentido, dado que los casos de suerte moral resultante son los más favorables a las intuiciones de los negadores —cosa que debe ser reconocida—, su estrategia respecto a la suerte circunstancial es a menudo discutir ejemplos de suerte circunstancial próxima, muy semejantes a los casos de suerte resultante. En realidad, el caso de George y Georg es más semejante a los ejemplos paradigmáticos de suerte resultante que a los de suerte circunstancial. Con ello el negador puede hacer más fuerte su posición. Sin embargo, considerando sólo este tipo de casos se desatiende a la diferencia fundamental entre suerte resultante y suerte circunstancial. Pues lo crucial de los casos de suerte circunstancial más interesantes es que las circunstancias en las que uno se encuentra repercuten en sus propios estados mentales —por decirlo de un modo rápido, que más adelante desarrollaré— y, a través de ellos, acaban marcando una diferencia en las acciones que el agente realiza. Para casos como el de

George y Georg, mi respuesta será más bien la que dé en el siguiente capítulo, en relación a la suerte resultante.

6.1.2. Carácter y acción

He aquí, de nuevo, el caso de los dos adeptos al nazismo. Rudolf y Adolf son dos alemanes igualmente simpatizantes de la política que el nazismo está empezando a desplegar en Alemania y Austria; pero Adolf, por motivo de negocios, tiene que emigrar a Argentina antes de que empiece la II Guerra Mundial y no puede participar activamente en la política nazi. Rudolf, por el contrario, permanece en Alemania y tiene la oportunidad de enrolarse en las SS y acabará dirigiendo un campo de exterminio. Lo relevante para el defensor de la suerte moral circunstancial es que las diferentes circunstancias en las que cada uno se mueve le abren y cierran diferentes oportunidades de acción que acabarán marcando una diferencia en su propio carácter y, finalmente, en su merecimiento. Parece claro que el juicio que Rudolf nos merece (miembro de las SS y comandante de un campo de exterminio) es mucho más duro que el de Adolf (mero simpatizante *intelectual*). Sin embargo, el adversario de la suerte moral afirma que lo realmente importante para su evaluación moral no es cómo cada uno de ellos finalmente actúa, sino que si Adolf hubiese estado en las mismas circunstancias que Rudolf, se habría integrado en el nazismo en la misma medida y habría cometido los mismos horribles crímenes. De hecho, estamos suponiendo que el emigrado comparte con su contraparte los mismos rasgos de carácter moralmente relevantes. Siendo esto así, para el negador ambos son igualmente responsables.

La solución puede venir, pues, de la mano de la idea de situar el locus último de la evaluación moral en las disposiciones estables o carácter del agente —en su naturaleza virtuosa o viciosa. Norvin Richards de-

fiende explícitamente esta posición. También Linda Zagzebski, cuyo argumento general, que vimos al final del capítulo 3, apela a la minimización del que, según ella, era el *efecto acumulativo* de la suerte.⁸ Sin duda, esta estrategia parece especialmente prometedora para el fin de resistir la suerte moral circunstancial.⁹ Y digo *especialmente* porque la estrategia, en principio más poderosa, de situar el locus de la responsabilidad en la volición, decisión o intención inmediatas del agente, no es aplicable a la suerte circunstancial —pues el agente no ha tomado ninguna decisión ni se ha formado una intención concreta en relación a los actos por los que se condena a su contraparte (salvo en casos como los de George y Georg, que asimilo a los de suerte resultante).

Sin embargo, esta posición no esta falta de problemas; más allá incluso de la obvia cuestión de que los rasgos internos de carácter del agente dependen crucialmente de situaciones y acciones anteriores que en su mayoría están igualmente fuera del control del agente —esto es, de la suerte constitutiva y formativa, sobre las que ya he insistido suficientemente con anterioridad. Pero hay aún otro tipo de problemas, que tiene que ver con la secuencia que va de los rasgos de carácter en adelante; esto es, desde éstos y las intenciones específicas a que dan origen, hasta las acciones particulares y los resultados de éstas. Por supuesto, acepto que los rasgos de carácter son un objeto legítimo de evaluación moral (como

⁸ Recuérdese que Zagzebski (1996) describía así este efecto: “Los rasgos internos de carácter originan disposiciones en circunstancias específicas, que llevan a la formación de intenciones particulares, que llevan a la realización de acciones, que a su vez producen consecuencias externas. En cada estadio se añaden nuevos elementos de suerte, de modo que el mayor grado de suerte se encuentra en las consecuencias, el menor en los rasgos de carácter.” (p. 72.)

⁹ Esta estrategia, como veremos en el próximo capítulo, está sobretodo a disposición del adversario de la suerte resultante. Por supuesto, éste último tiene también disponible la opción de basar la responsabilidad en los rasgos de carácter. De hecho, si se consigue anular la suerte moral circunstancial de este modo, automáticamente se anulará también la resultante.

insistí en el capítulo anterior); mi problema es más bien con la tesis de que deban ser *el locus último* o único. Ya avancé mi rechazo de la idea misma de que exista un locus último en el agente a partir del cual debamos dictar sentencia en cuanto a su estatus moral.¹⁰ Pero veámoslo más detenidamente. La cuestión es: ¿por qué no fundar la atribución de responsabilidad moral en los rasgos de carácter moralmente relevantes?

Para responder a esta pregunta cabrá, en primer lugar, tratar de indagar la naturaleza misma de los rasgos de carácter. Parece que mucho de lo que está en juego aquí depende de qué concepción de los rasgos de carácter sea la más correcta. De hecho, la idea misma de limitar la evaluación del agente a sus rasgos internos descansa fundamentalmente en una concepción relativamente estable o robusta del carácter. Es esto lo que idealmente podría autorizar el desatender incluso a su misma exteriorización. Téngase en cuenta que si la estrategia de fundar la asignación de responsabilidad moral en los rasgos de carácter del agente ha de superar el *peligro* de la suerte moral circunstancial, aquella tiene que renunciar a la consideración de cómo estos rasgos son exteriorizados en circunstancias determinadas.

Puede ser interesante, empezar considerando el siguiente pasaje de Nagel:

¹⁰ Esta idea podría compararse con el fundamentismo en epistemología. Creo que es igualmente equivocado que seamos indirectamente responsables de nuestras acciones, en tanto que originadas correctamente en nuestro carácter, por el que seríamos directamente responsables; ni tampoco, a la inversa, que seamos responsables de nuestro carácter, porque últimamente lo somos de las acciones que fueron el origen de ese carácter, etc. Mi alternativa es que cualquier elemento del conjunto de nuestros rasgos, acciones, etc. son susceptibles de ser la base legítima de atribuciones *directas* de responsabilidad moral. Recuérdense mis consideraciones de la sección 5.5, en el capítulo anterior.

Una persona puede ser codiciosa, envidiosa, cobarde, fría, poco generosa, poco amable, vanidosa o presumida, pero comportarse perfectamente gracias a un *esfuerzo enorme de la voluntad*. Poseer esos vicios supone no poder evitar tener ciertos sentimientos dadas ciertas circunstancias, así como sufrir fuertes impulsos espontáneos en esas circunstancias. *Incluso si uno controla los impulsos, sigue teniendo el vicio*. Una persona envidiosa odia el mayor éxito de los demás. Puede ser moralmente condenada como envidiosa incluso si les felicita cordialmente y no hace nada para denigrarlos o arruinar su éxito. La vanidad, asimismo, *no es necesario que se muestre*. Está completamente presente en alguien que no puede evitar regodearse secretamente en la superioridad de sus logros, talentos, belleza, inteligencia o virtud.¹¹

Esta posición podría radicalizarse sosteniendo que una persona puede ser, por ejemplo, cobarde, aunque nunca haya manifestado este rasgo, incluso aunque nunca haya experimentado internamente una sensación que cabría calificar de actitud cobarde. Esta sería la posición de un, digamos, disposicionalista extremo. Pero lo que Nagel parece defender aquí no es tan fuerte; lo que dice es que podemos juzgar los rasgos internos de un agente con independencia de sus mismas acciones. Es decir, que los rasgos internos pueden ser el objeto *directo* de juicio moral. Alguien que es envidioso, pero que evita comportarse envidiosamente en una situación determinada, no deja por ello de ser envidioso y se le puede censurar por ello — en particular, por haber adquirido, o no acabar con, esa disposición—; aunque, por supuesto, si alguien logra refrenar su tendencia a la envidia merece un elogio que no merece el que se deja llevar por ella. Podría darse también el caso de que éste consiguiera refrenar siempre su inclinación a la envidia, y sólo nos confesara verbalmente el constante acecho de ésta.

¹¹ Nagel (1979), p. 32-3. Mis cursivas.

Pero en este caso es dudoso que esa persona pueda ser considerada como *realmente* envidiosa.¹²

Particularmente, diría que es absurdo mantener que uno puede tener un rasgo de carácter aún sin que nunca lo manifieste, ni interna ni externamente. Además, alguien que nunca se ha encontrado en una situación determinada, por ejemplo una situación real de guerra, que reclama la manifestación de cierto rasgo, el valor, no sólo no conoce cuál sería su valor o falta de él en esas circunstancias, sino que parece que en realidad su carácter ni posee este rasgo ni el contrario, sino que es más bien *indeterminado* respecto a él. Ciertamente, un contraste tan radical entre la inclinación interna y la conducta externa es, cuanto menos, altamente inusual en el mundo real.¹³ Además, hablar de carácter presupone hábitos, *maneras características* en que una persona cree, siente y actúa, disposiciones *normales* a actuar y sentir. En todo caso, constataciones como la de Nagel apuntan al hecho de que tanto podemos juzgar una acción como un rasgo (como defendí en el capítulo anterior), dependiendo del propósito particular del juicio.

Una consideración preliminar en contra de fundar la atribución de responsabilidad moral exclusivamente en los rasgos de carácter es la de que el hecho mismo de manifestar en la conducta unos rasgos de carácter, y no meramente poseerlos, supone una diferencia moralmente significativa. En otras palabras, y pensando en rasgos negativos, como la envidia, llevar a acto la disposición a la envidia es peor —*intrínsecamente* peor, diría yo— que sólo poseer la tendencia interna. Si esto es así, parece que

¹² No pretendo negar que alguien pueda ser cobarde sin saberlo; pero entonces debe haber algo en su conducta que muestre su cobardía.

¹³ Véase sobre este punto, Rosebury (1995), p. 509-10; quien considera la afirmación anterior de Nagel y distingue hasta ocho alternativas de lo que podríamos querer censurar en la relación rasgo-acción.

estar en una situación u otra, que propicie o no la manifestación del rasgo, marcará en muchos casos una distinción clave. La idea puede aun generalizarse: llevar a acto cualquier tipo de tendencia, intención o volición es significativamente distinto de sólo tener o mantenerlas internamente. En lo que sigue, trataré de defender esta idea.

Sin embargo, hay un ataque previo a esta posición cuyo blanco es la concepción robusta del carácter asumida; a saber, que parece que esta concepción ha sido empíricamente desautorizada.¹⁴ Es por ello, que en la sección siguiente, presentaré y discutiré los experimentos más relevantes de la tradición de investigación sociopsicológica que presuntamente ha alcanzado tales resultados y que se conoce como *situacionismo*. Puedo adelantar que me opondré al escepticismo fuerte acerca de los rasgos de carácter, pero enfatizaré las repercusiones del situacionismo para la autonomía y la responsabilidad moral y, así, en última instancia, para la suerte moral.¹⁵

6.2. Disposiciones y la fuerza de las circunstancias: el reto situacionista

Cabe empezar advirtiendo que la tradición situacionista se apropia de muchos experimentos que no tienen por qué haber sido explícitamente

¹⁴ Una réplica temprana, en este contexto, a la concepción excesivamente estable de los rasgos de carácter, contra la evidencia empírica, es Adler (1987). Adler discute a Richards, que es quien principalmente apela al carácter como base para la asignación del merecimiento. Para Adler: “La consistencia en la conducta a través de las situaciones es la excepción, no la norma.” (1987, p. 249.)

¹⁵ Esto último justifica por sí solo —a mi modo de ver— el largo espacio dedicado en este capítulo sobre la suerte circunstancial a esta cuestión. No obstante, para aquellos que aun duden de la conveniencia de tratar aquí y tan por extenso este tema, debo añadir que mi réplica al eliminativismo sobre los rasgos de carácter pretende ser, en sí misma, una contribución a este debate reciente.

diseñados para poner a prueba hipótesis situationistas. Más bien, esta tradición recoge todos aquellos experimentos que muestran que factores situacionales de diferente tipo pueden ejercer una influencia “sorprendente” (inesperada) en la génesis de nuestras acciones. A continuación, describiré la metodología y resultados de los más destacados experimentos y debatiré su interpretación.

6.2.1. Los experimentos de Milgram sobre la obediencia

El grupo de experimentos más conocidos en esta tradición son los ideados por Stanley Milgram sobre la obediencia y motivados por la entonces reciente e inquietante conducta de los ciudadanos alemanes bajo el nazismo. En particular, el propósito de Milgram fue poner a prueba la famosa tesis, esgrimida por Hanna Arendt, de *la banalidad del mal*.

Recordemos que, en una controvertida y discutida serie de artículos publicados en el periódico *New Yorker* —y después convertidos en libro— Arendt relata y analiza el proceso judicial a Eichmann en Jerusalén y usa la expresión “banalidad del mal” para caracterizar la conducta del acusado, arquitecto en jefe y ejecutor de la Solución Final de Hitler.¹⁶ A primera vista, en el contexto de semejante genocidio el término “banal” no puede más que parecer inapropiado. Pero con él Arendt pretende rebatir las interpretaciones prevalecientes, según las cuales las atrocidades nazis emanan de una malvada voluntad de hacer el mal, incluso de un delito gratuito en el asesinato.

La tesis general, que Milgram tomará en consideración, es la de que “en determinadas circunstancias, la más decente persona podría convertirse en un criminal”. No hay duda de que la tesis es, al menos, vaga. Por supuesto, en sentido literal es incorroborable, pues habría que poder

¹⁶ Arendt (2006).

situar a todas las personas en todas las circunstancias posibles relevantemente diferentes. Pero sí que puede pensarse en ciertos resultados, experimentalmente comprobables, que serían relevantes para evaluar esta idea *matizadamente* —como lo sería, en particular, que determinadas circunstancias hagan que ciertos rasgos corrientes se desaten, que muestren que las personas carecen de un rasgo que esperábamos que tuvieran, o que muestren el carácter limitado de una disposición que creíamos más amplia y extendida.¹⁷ Así, Milgram se propuso convertir la tesis en una hipótesis empírica y ponerla a prueba.

En particular, Milgram estudió a cerca de mil sujetos durante un período de tres años (1960-63). Sus experimentos, además, han sido posteriormente replicados en muy diversos países, obteniéndose los mismos resultados. El experimento paradigmático es como sigue. A partir de anuncios en prensa en los que se buscaban personas para participar en un experimento sobre aprendizaje y memoria, y por lo que se ofrecían 4,50 dólares como pago (una cantidad nada despreciable en aquel tiempo), se reclutaron sujetos de edades comprendidas entre los veinte y los cincuenta años, de diversa procedencia socioeconómica y educativa. Tras informar a los participantes de que el propósito del estudio era comprobar los efectos del castigo en el aprendizaje, se les distribuyó en tres roles: el de Experimentador, el de Enseñante y el de Aprendiz. El sujeto real de experimentación era siempre el Enseñante, los otros dos eran miembros del equipo de experimentación. El Enseñante tenía que leer primero pares de palabras al Aprendiz y después sólo la primera palabra del par con cuatro asociaciones posibles. El Aprendiz tenía que elegir la pareja correcta de la asociación original. Cuando se equivocaba, el Enseñante debía presionar

¹⁷ Véase Flanagan (1991), cap. 14, esp. pp. 293ss.

un botón que produciría una descarga eléctrica en el Aprendiz. La descarga era en realidad simulada, pero el Enseñante lo desconocía. Tras cada error, el Enseñante debía aumentar el voltaje de la descarga. Si durante el experimento el Enseñante expresaba preocupación —y prácticamente todos lo hicieron— el Experimentador decía, progresivamente, “por favor, siga”, “el experimento requiere que continúe”, “es esencial que continúe” y “no tiene elección, debe seguir”. El generador de descargas eléctricas tenía un panel de treinta palancas asociadas a cada descarga que iban de 15 a 450 voltios, etiquetadas como descarga ligera, moderada, fuerte, muy fuerte, intensa, extrema intensidad, peligro y descarga severa. Las últimas dos palancas tenían inscrito “XXX”. Cabe añadir que antes de empezar el experimento, el Enseñante explicó que aunque las descargas eléctricas podían ser extremadamente dolorosas, no causarían daños en su piel.

A lo largo de una gran variedad de protocolos similares, el 65 por ciento de los sujetos llegaron a hasta los 450 voltios, incluso a pesar de las protestas y gritos del Aprendiz y de los insistentes golpes en las paredes, a partir de los 300 voltios. Incluso en variaciones en las que la administración de la descarga requería que el Enseñante situara enérgicamente la mano del Aprendiz en la bandeja de la descarga, un 60 por ciento de los sujetos llegaron a la máxima descarga. En la Figura 1 se puede apreciar la distribución del número de sujetos y el punto del estudio al que llegaron en el experimento original; en él, 26 de 40 llegaron hasta el final.¹⁸

¹⁸ Para el experimento original, véase Milgram (1963). Milgram explica las diferentes variaciones del experimento, junto con el estudio de sujetos individuales que participaron en él, y ofrece las claves de su interpretación de las fuerzas psicológicas operantes en Milgram (1974).

Voltage level	Learner response		Number for whom maximum shock
Slight Shock	15–60	No Response	0
Moderate	75–120	“Ugh! Hey this really hurts.”	1
Strong	135–180	“Ugh!!! Experimenter! That’s all . . . heart trouble . . . let me out!	6
Very Strong	195–240	. . . You have no right to keep me in here . . . Let me out! Let me out!	1
Intense	255–300	(Agonized Scream.) I absolutely refuse to answer . . .	3
Extreme Intense	315–360	(Intense and prolonged agonized scream.) . . . heart . . .	2
Danger: Extreme Shock	375–420	No response.	1
XXX	435–450	No response.	26

Figura 1. De Nelkin (2005), adaptación de Doris (2002) a partir de una tabla original de Milgram (1974).

El resultado es especialmente robusto, y chocante incluso para los académicos consultados. En uno de los estudios, Milgram reunió a un grupo de treinta y nueve psiquiatras de Yale, que predijeron, en relación a un grupo diverso de sujetos, que no más de un 50% superaría el décimo nivel de 150 voltios, menos de un 4% llegaría al nivel veinte, y menos de uno entre mil administraría la descarga máxima. Se pensaba que sólo los sádicos podrían llegar a los 450 voltios. En general, tanto los psiquiatras, como los estudiantes y profesionales consultados se mostraron seguros de que ellos mismos habrían abandonado pronto el experimento. Por otro lado, Milgram consiguió maximizar la obediencia introduciendo más de un Enseñante en cada experimento: cuando un Enseñante observaba antes a otro Enseñante cumplir obedientemente las instrucciones, el porcentaje de obediencia llegaba al 90%, y descendía al 10% cuando un compañero

Enseñante desobedecía (o cuando la víctima actuaba como un masoquista que pedía más descargas).¹⁹

Cabe tener en cuenta que una significativa minoría (un tercio) rechazó obedecer, aunque es de remarcar que ninguno de ellos denunció el experimento ante autoridades superiores. Como ha afirmado John Doris, “los experimentos de Milgram muestran cómo factores situacionales aparentemente no coercitivos pueden inducir la conducta destructiva más allá de la evidente presencia de estructuras evaluativas y disposicionales contrarias.”²⁰ Milgram pensó que el principal hallazgo del estudio era la extrema voluntad de los adultos a obedecer hasta casi cualquier extremo las instrucciones de una autoridad, lo que da apoyo a la tesis de la banalidad del mal de Arendt.

Existen ciertas claves de cómo los sujetos fueron llevados a obedecer que no puedo dejar de comentar. Pero empecemos viendo lo que tiene que decir Lee Ross, un eminente psicólogo social, al respecto:

Quizá la característica más obvia y reconocida del paradigma específico de Milgram fue la naturaleza gradualmente creciente de la complicidad del enseñante. Éste no obedeció la sola y simple orden de producir una poderosa descarga eléctrica en una víctima inocente. Al principio, empezó meramente a producir suaves refuerzos negativos —respuestas, en realidad— a un aprendiz que estuvo de acuerdo en recibir esta respuesta como ayuda para realizar su tarea. También aceptó, como lo hizo el aprendiz, un procedimiento en el que el nivel de refuerzos negativos se iría incrementando ligeramente tras cada error; pero lo hizo sin ni siquiera imaginarse las implicaciones a largo plazo de ese acuerdo inicial. La progresión continuó paso a paso, y con cada incremento del nivel de descarga, el dilema psicológico del enseñante se fue tornando más y más difícil. En cierto sentido, el enseñante tuvo que buscar una justificación (que fuese satisfactoria para sí mismo, para el

¹⁹ Los experimentos de Solomon Asch (que describo más abajo) corroboran este efecto.

²⁰ Doris (2002), p. 39.

experimentador y quizá para el aprendiz) que explicara por qué tenía que desistir *ahora*, cuando no lo había hecho antes; cómo podía ser ilegítimo producir la siguiente descarga, pero legítimo haber producido una de magnitud sólo ligeramente menor momentos antes. Esta justificación es difícil de hallar. De hecho, es claramente asequible sólo en un punto del procedimiento —el punto en el que el aprendiz retira su consentimiento a recibir descargas y, aún así, se ve obligado a continuar en el experimento— y es significativamente éste el punto en el que fue más probable que los sujetos dejaran de obedecer.²¹

Podemos resumir las claves en cuatro: 1) La más importante es el efecto gradual del experimento. El primer paso era insignificante, y después el nivel de agresión fue aumentando muy poco a poco. 2) Sin embargo, con el aumento de las descargas aparecía un problema de justificación para el sujeto. Poner la línea en un punto determinado no era nada fácil, más aun cuando suponía autodesautorizarse a un mismo, desautorizar la propia conducta anterior. 3) Además, el sujeto había aceptado, en un contrato al menos verbal, comportarse de acuerdo a cómo se le pediría. El experimentador le había presentado una justificación aceptable para actuar de un modo tan indeseable, además de darle un papel fundamental en el experimento y unas reglas claras de lo que tenía que hacer, de modo que el valor de sus acciones parecía invertido: el error era no hacer lo que se le exigía. Con el castigo ayudaba al aprendiz. 4) Por último, los costes de abandonar eran altos. Muchos protestaron pero se les dijo que el experimento exigía que continuasen, que no les estaba permitido abandonar. Cabe remarcar que el mero disentimiento verbal no contaba como desobediencia para los diseñadores del experimento, sólo la conducta efectiva. La gran dificultad era convertir la intención en acción. Además, las exigencias del

²¹ Ross (1988), pp. 102-3.

experimentador facilitaban la difusión de la responsabilidad. El experimentador les recordaba que ellos no eran responsables.²²

La verdad es que resultados inesperados de este tipo resultan alarmantes o, por lo menos, chocantes. Esto es, contradicen determinadas intuiciones o sesgos de nuestra psicología de sentido común. En particular, contradicen nuestra excesiva tendencia a atribuir todo el peso de la conducta a rasgos internos y robustos del agente. A esto se le conoce en psicología social como el Error Fundamental de Atribución:

(EFA) La creencia exagerada de la gente en la importancia de los rasgos de carácter y disposiciones, junto con la falta de reconocimiento de la importancia de los factores situacionales, en la producción de la conducta.²³

Es de resaltar, a este respecto, el fallo estrepitoso de los psiquiatras en la predicción de la conducta de los sujetos del experimento.

Algunos filósofos han ido más lejos y han afirmado que resultados experimentales como el anterior, y otros que veremos a continuación, vendrían a poner en solfa la robustez de los rasgos de carácter, cuando no su misma existencia. En lo que siguiente intentaré clarificar y evaluar esta tesis. Para ello, empezaré describiendo, más sucintamente, otros experimentos particularmente relevantes al respecto.

²² En la elaboración de esto cuatro puntos he seguido a Flanagan (1991), p. 297; y Zimbardo (2004), pp. 26-29. En general, el paralelismo de los sutiles efectos psicológicos registrados con las tácticas políticas de los nazis es más que notable.

²³ Ver esp. Ross y Nisbett (1991).

6.2.2. Otros experimentos situacionistas

Empecemos por el grupo de experimentos en principio más relevante, que es el constituido por los estudios sobre influencia social, como el mismo experimento de Milgram. Vimos cómo la manipulación de ciertas variables situacionales sutiles tenía efectos poderosos en la producción de conductas determinadas, aun en contra de las convicciones previas del agente. Pero la acción puede ser también inhibida.

En 1964, Kitty Genovese estuvo pidiendo auxilio a gritos cerca de media hora mientras era violada y finalmente apuñalada en un parque de Queens, sin que ninguno de los cuarenta testigos que observaron los hechos desde sus viviendas acudiese en su ayuda, ni siquiera llamase a la policía. Los medios de comunicación atribuyeron los hechos a la apatía e insensibilidad de los neoyorquinos. Sin embargo, los psicólogos sociales empezaron a estudiar posibles factores situacionales causantes de inhibir la intervención, y numerosos experimentos mostraron que la presencia o el aumento del número de personas que era testigo de situaciones que reclamaban su intervención repercutía en el porcentaje de gente que intervenía. En un experimento se detectó que cuando los sujetos estaban solos un 70% ayudaba a una mujer que había sufrido una mala caída; mientras que cuando estaban junto a un colaborador del experimento que interpretaba el papel de compañero y que hacía caso omiso de los gritos de la víctima, sólo el 7% de ellos intervenían.²⁴

Una explicación de esto es que cuando estamos entre otras personas nuestra percepción de la situación se altera. Este efecto fue confirmado por los estudios de Solomon Asch sobre influencia de grupo en relación a errores visuales.²⁵ Asch descubrió que de un 50 a un 80 por ciento

²⁴ Latané y Darley (1968); véase también Latané y Rodin (1969).

²⁵ Asch (1951).

de los sujetos estudiados, dependiendo del estudio particular, se amoldaron a la visión mayoritaria incorrecta por lo menos una vez. Los sujetos tendieron a ajustar su opinión a la de los demás aún en contra de la evidencia de sus sentidos.

Si nadie reacciona, no debe ser tan serio después de todo. La percepción de nuestra responsabilidad de actuar se difumina, así, por la posibilidad de que otros intervengan. Lo dramático aquí es que usualmente no nos damos cuenta de estos efectos de grupo. Los experimentadores han registrado que “los sujetos persistentemente afirmaron que su conducta no estaba influenciada por la presencia de otras personas. Esta negación se daba ante las pruebas mismas que mostraban que la presencia de otros inhibían la ayuda.”²⁶

En otro experimento, conocido como el *experimento del buen samaritano*, Darley y Baston dividieron a estudiantes seminaristas de Princeton en dos grupos y a los miembros de uno de ellos se les sugirió que prepararan una breve charla sobre la parábola del buen samaritano y a los del otro, sobre posibles ocupaciones para seminaristas, que todos tenían que ir a exponer a un edificio colindante. Por el camino, un hombre echado en el suelo parecía necesitar ayuda médica. Los investigadores hallaron que era 6 veces más probable que los transeúntes sin prisa ayudasen a esta persona en apuros que los transeúntes con prisa (63% v 10%). El contenido de las charlas no tuvo una clara incidencia en la conducta. Tampoco se encontró ninguna correlación significativa entre la conducta de aquellos que ayudaron y las autoatribuciones de rasgos de carácter. Fue, básicamente, el factor de la prisa el que hizo que los sujetos cambiaran su percepción de la situación: “a causa de la presión del tiempo, no

²⁶ Latané y Darley (1970), p. 124.

percibieron la escena como una ocasión para la decisión ética”.²⁷ De nuevo, estos efectos pasaron desapercibidos para los sujetos participantes en el estudio.

Hay otro famoso experimento, ideado a partir del experimento de Milgram, conocido como el experimento de la Prisión de Stanford. En 1971, Phillip Zimbardo y sus colegas reclutaron un grupo de estudiantes universitarios varones adultos que dividieron al azar en dos grupos: los que interpretarían el papel de “presos” y los que harían lo propio con el de “carceleros”. Aunque el experimento fue diseñado para que durara dos semanas, tuvo que ser interrumpido a las seis horas debido al creciente abuso por parte de los carceleros y el desaliento y la extrema ansiedad de los presos. Los resultados fueron asombrosos e inquietantes, al mostrar el gran uso de la humillación y otros abusos psicológicos por parte de algunos carceleros y la falta de voluntad de intervenir por parte de los demás.²⁸ Los paralelismos con los abusos de los soldados norteamericanos a los reclusos de la prisión iraquí de Abu Ghraib son evidentes.

La investigación se centró aquí en el fenómeno de la desindividua-
lización. El experimento difiere del paradigma de Milgram en que en él no hay una figura de autoridad presente que inste al sujeto a obedecer. Más bien, la conducta de los sujetos viene guiada por la configuración de la situación, sin que éstos piensen en el significado o las consecuencias de sus acciones. Las acciones no responden tanto a una guía cognitiva, sino que están dirigidas por las acciones de otros próximos a ellos, o por los fuertes estados emocionales que éstos y los rasgos de la situación les suscita, tales como la presencia de armas.²⁹

²⁷ Darley y Batson (1973), p. 108.

²⁸ Véase Haney, Bank y Zimbardo (1973); también se puede consultar el sitio web del experimento, en www.prisonexp.org.

²⁹ Zimbardo (2004), pp. 32-3.

Por otro lado, existe también un importante número de experimentos que muestran que variaciones calificables como menores bajo cualquier estándar pueden afectar a los estados de ánimo y conducta de las personas. Isen y Levin observaron que para sujetos que acababan de encontrarse una moneda de 10 céntimos en la ranura de retorno de una cabina telefónica era 22 veces más probable que ayudaran a una mujer a quien le habían caído unos papeles al suelo, que sujetos que no se habían encontrado la moneda (88% v 4%).³⁰ Por su parte, Mathews y Canon hallaron que era 5 veces más probable que personas a quienes se les habían caído unos libros fuesen ayudadas cuando el sonido ambiente estaba a un nivel normal que cuando un gran cortacésped estaba en marcha en las inmediaciones. Estos resultados parecen especialmente inquietantes, pues nadie espera ni desea que la conducta moral dependa de semejantes factores.³¹

Finalmente, otro grupo de estudios tiene que ver con la falta de correlación conductual entre diferentes situaciones. Hartshorne y May llevaron a cabo un ambicioso experimento con la participación de cerca de 8.000 escolares, en el que observaron que las correlaciones en la conducta honesta de los niños en diferentes situaciones era muy baja (.2 y .3). Aunque actuaron de manera similar en casos repetidos de la misma situación, no actuaron consistentemente a lo largo de las diferentes situaciones planteadas. Casi ninguno de los escolares actuó honestamente de manera consistente, incluso entre situaciones muy similares, como mentir, copiar en un examen y robar monedas de encima de una mesa. La conclusión que alcanzaron los investigadores es que la honestidad no es una “entidad interna” sino “una función de la situación”.³²

³⁰ Isen y Levin (1972).

³¹ Mathews y Canon (1975).

³² Ver Hartshorne y May (1928). “Nuestra opinión [...] es que este factor común no es una entidad interna que opera independientemente de las situaciones en las que los indi-

En realidad, estos experimentos son sólo una pequeña muestra de todo lo que pueden agruparse bajo la etiqueta de *situacionistas*. No obstante, son los más conocidos y constituyen una buena muestra de los resultados más importantes obtenidos por esta tradición de investigación.

Tenemos, pues, que en cada uno de estos experimentos hay ciertos factores situacionales que juegan un destacado papel en la determinación de la conducta de los sujetos, que minimiza la importancia de las diferencias en sus disposiciones internas. Esto es, si un factor situacional particular —tal como encontrar una moneda— puede provocar una conducta similar de la mayoría de sujetos, entonces, las diferencias entre los caracteres de los individuos serán correspondientemente poco significativas en la producción de su conducta. Ello parece cuestionar la creencia de que la consistencia conductual que encontramos en los demás y en nosotros mismos sea convenientemente explicada por medio de la atribución de rasgos de carácter robustos. Más bien, abonarían la idea de que los factores situacionales a menudo predicen mejor la conducta de una persona que los factores personales.

6.2.3. Repercusiones: escepticismo acerca de los rasgos de carácter

Concretamente, estos resultados han sido interpretados como una amenaza para las atribuciones de rasgos de carácter. La Tesis Situationista acerca de los Rasgos de Carácter podría formularse así:

viduos se encuentran, sino una función de la situación en el sentido de que el individuo se comporta de modo similar en situaciones diferentes en proporción a lo semejantes que sean esas situaciones, hayan sido experimentadas como oportunidades para la conducta honesta o deshonestas, y sean comprendidas como oportunidades para el engaño o la honestidad.” (Volumen 1, p. 381.)

(SRC) Los rasgos de carácter (o personalidad) tradicionales juegan un papel menos relevante en la predicción y explicación de la conducta que los factores situacionales particulares.³³

En mi opinión, esta tesis se sigue sin problemas de los resultados experimentales. Pero, como salta a la vista, es demasiado indeterminada. La cuestión sería en qué medida nuestra conducta es menos dependiente de nuestros rasgos personales. ¿Lo es de una manera tan importante como para justificar, por ejemplo, el abandono de nuestras atribuciones cotidianas de rasgos de carácter? La meta primaria debe ser tratar de precisar qué consecuencias pueden tener los resultados.

Algunos filósofos, en particular Gilbert Harman y John Doris, consideran que estos resultados muestran que los rasgos de carácter son ficciones de las que debemos que prescindir. En principio, un buen argumento con base empírica contra los rasgos de carácter debería parecerse a éste:

Argumento para la eliminación de los rasgos de carácter

- (1) Si la conducta de las personas es de hecho fruto de rasgos de carácter (robustos), entonces la observación sistemática revelará la consistencia y fuerte estabilidad en su conducta.
- (2) De hecho, la observación sistemática no revela *ni* la consistencia entre situaciones *ni* la fuerte estabilidad temporal en la conducta de las personas.
- (3) Por lo tanto, la conducta de las personas no es de hecho fruto de rasgos de carácter (robustos).

³³ Existen ciertas diferencias entre los rasgos de *carácter* y *personalidad* —por ejemplo, los primeros incluyen una dimensión evaluativa que los últimos no tienen. En todo caso, lo que nos interesa son los rasgos con repercusiones morales. Para más sobre esto, véase Doris (2002), cap. 2; y Kupperman (1991), cap. 1.

No obstante, a partir de los experimentos anteriores no parece posible montar este argumento. Pues, ¿cómo pueden mostrarnos esos experimentos la falta de consistencia y fuerte estabilidad si no son reiterativos, sino que examinan situaciones puntuales? El experimento ideal para este propósito debería, más bien, basarse en una amplia muestra de sujetos cuya conducta fuera estudiada individualmente a lo largo de un gran número de situaciones diferentes. A partir de estos datos, se llegaría a obtener expedientes individuales con los cuales comprobar la consistencia y estabilidad conductual o falta de ellas para cada sujeto. En particular, podríamos comprobar en qué medida la conducta es o no situacionalmente variable. Sin embargo, este experimento parece logística y éticamente inviable, o por lo menos es muy diferente de los experimentos reales de que disponemos.³⁴ Por ello, el argumento tiene que ser modificado. Creo que la siguiente versión revisada del argumento puede funcionar mejor:

Argumento revisado para la eliminación de los rasgos de carácter

- (1) Si las personas poseen rasgos de carácter diferentes, actuarán de maneras significativamente distintas en la misma situación (en virtud de poseer rasgos de carácter diferentes).
- (2) En los experimentos, los participantes no actuaron de maneras significativamente distintas en las mismas situaciones.
- (3) Por lo tanto, las personas no poseen rasgos de carácter diferentes.

Posiblemente este sea el argumento que podría realmente justificar la eliminación de los rasgos de carácter, en virtud de los datos experimentales disponibles. Un supuesto de este argumento es que la noción de rasgos de

³⁴ Es cierto que algún experimento ha avanzado en esta dirección —en especial, el experimento sobre la honestidad de Harthorne y May (1928)— pero sólo de una manera muy limitada —en particular, no se estudió la conducta de los escolares individualizadamente.

carácter depende de la variedad o pluralidad de caracteres. Si todo el mundo tiene exactamente los mismos rasgos de carácter, nadie tiene ninguno; pues su funcionalidad para explicar o predecir la conducta sería cero. El argumento presupone, así, que los rasgos de carácter son entidades teóricas, cuya existencia depende de su funcionalidad en la predicción o explicación de la conducta. El argumento, como inferencia a la mejor explicación de los datos experimentales, justificaría una reducción explicativa, que en último término conllevaría la eliminación ontológica.

Una cuestión importante, antes de seguir, es la de qué se entiende exactamente por rasgo de carácter. Harman describe la que podemos llamar *concepción tradicional*, cotidianamente asumida, de los rasgos de carácter basándose en Aristóteles. Los rasgos de carácter son *disposiciones relativamente estables a largo plazo a actuar de un modo característico*. Las disposiciones relevantes han de conllevar hábitos y no sólo destrezas, incluyendo hábitos de deseo. Es importante resaltar que los rasgos de carácter de una persona ayudan a explicar por lo menos algunas de las cosas que la persona hace. Harman concluye que se trata de “amplias disposiciones a actuar que ayudan a explicar la conducta a la que disponen. Las disposiciones limitadas no cuentan.”³⁵

Sobre una base similar, John Doris ofrece esta definición general: *si una persona posee un rasgo, esa persona exhibirá la conducta relevante a ese rasgo en condiciones relevantes para la manifestación de ese rasgo, con una probabilidad p marcadamente superior al mero azar.*³⁶ La precisión probabilística es importante, pues nadie espera que poseer un rasgo determinado conlleve actuar siempre del modo relevante en todas las condiciones propicias para ello. Asimismo, Doris descompone la con-

³⁵ Harman (1999), p. 318.

³⁶ Doris (2002), cap. 2.

cepción tradicional de los rasgos de carácter —que llama *globalismo*— en tres características:

- 1) *Consistencia*: los rasgos de carácter se manifiestan de manera fiable en la conducta relevante al rasgo en cuestión a lo largo de una diversidad de condiciones que pueden variar ampliamente en su capacidad de conducir a la manifestación del rasgo en cuestión.
- 2) *Estabilidad*: los rasgos de carácter y personalidad se manifiestan reiteradamente de manera fiable en la conducta relevante a esos rasgos en condiciones relevantemente similares para la manifestación del rasgo.
- 3) *Integración evaluativa*: la ocurrencia de un rasgo con un valor evaluativo particular está probabilísticamente relacionado con la ocurrencia de otros rasgos con valores evaluativos similares.³⁷

Doris considera que los resultados experimentales suponen un rechazo de (1) y (3), pero permiten una variante de (2). Por el momento, dejaré de lado el punto (3).

Cabe remarcar que esta posición difiere parcialmente de la de Harman. En un principio, Harman consideró que de los resultados experimentales cabía inferir que los rasgos de carácter no existen y que, por lo tanto, debemos abandonar toda referencia a ellos en nuestro lenguaje y pensamiento. Mientras que Doris mantiene que los datos arruinan el *globalismo*, pero no cierran la posibilidad de que haya rasgos “locales”, particulares a situaciones y temporalmente estables, que estén asociados a importantes diferencias conductuales individuales.³⁸ No obstante, la diferencia entre ambos no es tan grande; pues el mismo Harman ha reconocido posteriormente la existencia de regularidades conductuales limitadas a situaciones similares.³⁹

³⁷ Doris (2002), pp. 22-3.

³⁸ Doris (2002), pp. 24-5 y 62-66.

³⁹ Compárese Harman (1999), (2000) y (2003).

En el siguiente apartado, trataré de desautorizar la conclusión de que los resultados experimentales nos obligan a inferir que no hay rasgos de carácter globales.

6.2.4. Resistiendo el eliminativismo

La situación es la siguiente. Por un lado, tenemos los resultados experimentales anteriores, que nos muestran el enorme peso de las circunstancias, incluso de factores de lo más trivial, en el origen de nuestra conducta, así como el peligro de las urgencias, las situaciones embarazosas, inesperadas, etc., que en conjunto parece que nos abocan a cierto escepticismo. Pero, por otro lado, la renuncia a los rasgos de carácter debería estar basada en una evidencia significativamente concluyente, dado su destacado papel en la psicología popular.

Empecemos suponiendo que los datos experimentales demuestran correctamente que los factores situacionales vencen o dominan ampliamente sobre los factores personales en la producción de la acción —y que la distinción radical entre los factores situacionales y personales tiene sentido. La cuestión sería: ¿por que este hecho socava exclusivamente los rasgos de carácter, y no otros factores personales? Parece que, para los eliminativistas, los únicos factores personales que los experimentos hacen *desaparecer* son los rasgos de carácter, pero ¿por qué? Por ejemplo, en el experimento de Milgram, ante las exigencias del experimentador, la mayoría de los sujetos actuaron del modo contrario a como previamente pensaban que lo harían. En este sentido, la situación prevalece sobre los factores personales. Pero los rasgos de carácter no son los únicos factores personales, sino que hay otras disposiciones, como es el caso de creencias, asunciones, deseos, preferencias, intenciones, planes o valores. Parece que lo que, en todo caso, se seguiría de los resultados experimentales

es la no funcionalidad en la producción de la conducta de todo este conjunto de factores personales. De hecho, el resto de factores personales parecen haber fallado en la misma medida a la hora de producir la conducta relevante. Piénsese, por ejemplo, en el valor o principio, que muchos de ellos seguramente compartirían, de no dañar a nadie sin una poderosa razón.

Pero si éste es el caso, la interpretación radical de los experimentos tendrá como consecuencia que la conducta entera depende por completo de la situación. El resultado sería una especie de nihilismo conductual (o acrasia generalizada): en cada instante de nuestras vidas estaríamos a merced de variaciones mínimas de los factores situacionales —incluso a merced de encontrar o no una moneda—, las cuales serían las únicas o verdaderas causas de nuestra conducta. Sin embargo, esta conclusión radical parece abiertamente absurda, o por lo menos nadie ha interpretado los resultados experimentales de este modo. No obstante, esta es la conclusión que se seguiría de la interpretación radical de los resultados necesaria para la eliminación de los rasgos de carácter.

Por otro lado, la misma distinción dicotómica entre factores personales y situacionales no parece muy prometedora.⁴⁰ En primer lugar, cabe notar que no es la situación en sí misma (esto es, los estímulos objetivos en sí) lo que en cualquier caso determinaría la conducta, sino la aprehensión subjetiva de la situación por cada persona. De hecho, esto es reconocido por Harman,⁴¹ así como por Ross y Nisbett cuando rechazan abiertamente una interpretación conductista estricta de los experimentos situacionistas.⁴² Además, los experimentos de Asch muestran la compleja

⁴⁰ Para intentos de socavar esta distinción: Malle (2005), Sabini *et al.* (2001) y Kupperman (1991), ap. A.

⁴¹ Véase Harman (1999), p. 318.

⁴² Véase Ross y Nisbett (1991), cap. 3.

función *persona-grupo* en la aprehensión de la situación. Por ello, el contraste en los experimentos *no* es sencillamente entre el poder de la situación y el individuo, cuya acción sería una especie de respuesta automática a estímulos objetivos; sino la respuesta del individuo a la situación, dependiente de la percepción subjetiva de los estímulos situacionales —esto es, en su construcción subjetiva de la situación. Estas consideraciones podrían bloquear aquel nihilismo, pero por supuesto no aseguran el control próximo del sujeto sobre sus propias acciones. Y no hay duda de que el situationista podría simplemente replicar: esto no es una objeción, dado el hecho de que un alto porcentaje de los sujetos actuó del mismo modo. En efecto, la situación particular dada funcionó igualmente para la mayoría en cada experimento.

Sin embargo, mi argumento contra el eliminativismo acerca de los rasgos de carácter es más bien este: si el eliminativista interpreta los resultados experimentales de un modo radical, como susceptibles de una completa generalización, tan radical como para socavar los rasgos de carácter; tiene asimismo que aceptar por la misma razón que el resto de los factores personales quedarán radicalmente socavados también. Por ello, me parece dudoso identificar el situacionismo como un problema *sólo* para los rasgos de carácter y no para otras disposiciones y factores personales. La cuestión puede formularse como un dilema: si se acepta que los resultados experimentales pueden generalizarse completamente —el único modo de negar los rasgos de carácter—, entonces se tendría que negar también el papel en la producción de la conducta del resto de los factores personales —lo que significaría aceptar el nihilismo conductual. Por otro lado, si se acepta que este nihilismo es insostenible, y no se lo considera como una conclusión plausible de los datos experimentales, entonces no se puede negar tampoco la existencia de los rasgos de carácter.

La única manera de no caer en este dilema sería ofrecer un principio por el cual los resultados experimentales deberían aplicarse sólo a los rasgos de carácter. Pero, que yo sepa, nadie lo ha ofrecido, ni es plausible que se encuentre uno. Por ello, el dilema me parece insalvable. Por otro lado, una vez estamos dentro del dilema, es claramente preferible optar por el rechazo del nihilismo.⁴³

Rasgos globales y rasgos locales

Ciertamente, cuando eliminativistas como Harman y Doris afirman que los resultados experimentales socavan la existencia de los rasgos de carácter, subsiguientemente deben explicar nuestra experiencia cotidiana de su atribución. Estas palabras de Ziva Kunda, una psicóloga social situacionista —que Harman cita con aprobación—, me parecen ilustrativas de la posición en cuestión:

Nuestra noción de rasgos como disposiciones amplias y estables que se manifiestan en el mismo grado en una variedad de situaciones no puede sostenerse. Pero esto no significa que no haya ningún tipo de diferencias duraderas y sistemáticas entre los individuos. Mi intuición de que soy una persona muy diferente a mi hermano o de que mi hijo tiene diferentes patrones de conducta predictibles no son necesariamente falsas. Estas intuiciones pueden basarse en diferencias significativas y estables entre individuos pero no en la clase de diferencias que conlleva la concepción tradicional [...].⁴⁴

En conclusión, parece que realmente somos bastante consistentes en nuestra conducta dentro de cada situación, y es bastante apropiado esperar esa consistencia en los demás. Pero tropezamos con problemas cuando esperamos que esta consistencia se extienda

⁴³ Más abajo ofrezco razones contra absurdo de generalizar la acrasia a todas nuestras acciones.

⁴⁴ Kunda (1999), p. 443. Citada en Harman (2003).

también a otras situaciones. Incluso mínimas variaciones en las características de la situación pueden conducir a dramáticos giros en la conducta de la gente.⁴⁵

En realidad, los eliminativistas acerca de los rasgos de carácter no niegan que la gente posea otras disposiciones no *robustas* (“robustas” en el sentido de rasgos de carácter globales); ni que no haya regularidades conductuales constatables regidas por esas otras disposiciones que las personas poseen. Lo que, en su interpretación, cabe inferir a partir de los resultados experimentales es la inexistencia de rasgos *globales* —tal y como fueron caracterizados: como disposiciones de carácter consistentes a lo largo de situaciones diversas y temporalmente estables—, pero no la inexistencia de rasgos *locales*, que no presuponen semejante consistencia ni estabilidad general. De hecho, la existencia de estos rasgos locales, cuyo alcance se restringe a situaciones similares, parece que estaría respaldada por un amplio conjunto de estudios empíricos, según sus defensores.⁴⁶ La fiabilidad conductual de los rasgos locales sería *altamente específica*, esto es, altamente relativa al contexto —una predictibilidad contextualizada insuficiente para fundar atribuciones de rasgos más amplios.

Por mi parte, no tengo ningún reparo en aceptar las palabras anteriores de Kunda. Sin embargo, la distinción entre rasgos globales y locales me parece mucho más problemática de lo que sus defensores reconocen. Mi sospecha es que se trata de una distinción excesivamente *ad hoc*, motivada por la necesidad de evitar, por su parte, una confrontación demasiado radical con nuestras intuiciones cotidianas. Tanto Doris como Harman quieren evitar un revisionismo *completamente antiintuitivo* de

⁴⁵ Kunda (1999), p. 499.

⁴⁶ Por ejemplo, por los experimentos de Harthorne y May (1928) sobre la honestidad; véase su descripción más arriba. Para la defensa de los rasgos locales, véase Ross y Nisbett (1991), Harman (2003) y Doris (2002), pp. 22-3.

nuestras creencias y prácticas al respecto.⁴⁷ En todo caso, mi ataque principal es que se trata de una distinción *demasiado débil*. Consideremos la historia de Carol que, en el contexto de la cita anterior, nos propone Kunda:

Carol es extremadamente extrovertida cuando se halla en situaciones uno a uno, es sólo moderadamente extrovertida en pequeños grupos y no es en absoluto extrovertida en grandes grupos. Se mostrará muy confortable y sociable cuando estás con ella a solas, pero cerrará el pico y se mostrará muy tímida e incómoda cuando te la encuentres en un gran grupo.⁴⁸

Pero, de hecho, Carol puede ser también extremadamente extrovertida sólo en situaciones uno a uno cuando la otra persona es un conocido, o un compañero de profesión, o otra mujer, o alguien más joven, o cuando la otra persona es un compañero de profesión y una chica, o un conocido más joven que ella; o en relación al lugar en que se hallan, lo serio que sea el encuentro, cómo esté dispuesto el mobiliario, etc., etc. Y lo mismo respecto a ser moderadamente extrovertida en pequeños grupos; ella puede que lo sea sólo cuando el grupo está compuesto por determinada gente y no otra, si están en un lugar público o privado, en un lugar de ocio y no en una comisaría de policía, etc. ¿Y qué hay de su estado de ánimo? La extroversión puede depender también de si se ha encontrado previamente una moneda o no.

Al final, parece que lo único que tendremos son descripciones de situaciones particulares. Doris reconoce que “quizás [la conducta] debería

⁴⁷ Esto es evidente a lo largo de todo el libro de Doris. Véase Doris (2002), *passim*; por ejemplo, p. 145. Harman (1999), p. 316, está especialmente interesado en admitir cosas como la esquizofrenia, la manía, la depresión, la timidez o la tristeza. Sobre esta cuestión, véase Miller (2003), nota 6.

⁴⁸ Kunda (1999), p. 444.

explicarse meramente por referencia a la continuidad situacional, en lugar de apelar a rasgos locales.” Sin embargo, él está indudablemente comprometido con la idea de que “algunas tendencias conductuales son lo suficientemente fiables como para garantizar la postulación de disposiciones duraderas”, y con que esas disposiciones se fijan en virtud de “características determinativas de las situaciones”.⁴⁹ Pero hemos visto lo difícil que resulta identificar estas *características determinativas de las situaciones*, en particular si, como afirma Kunda, “incluso las variaciones más ligereas en las características de la situación pueden conducir a giros dramáticos en la conducta de las personas.”⁵⁰

Vemos, pues, que la identificación de rasgos locales no es menos problemática que la identificación de rasgos globales, a la vista de los altos estándares impuestos por los eliminativistas para que algo cuente como un rasgo. Y creo que este problema constituye una seria objeción al proyecto antiglobalista, en tanto que el recurso a rasgos locales — “estrechamente especificados”, que “deberían satisfacer una estándar condicional”—⁵¹ le es esencial para acomodar muchas de nuestras intuiciones cotidianas.

Doris recomienda que en nuestras evaluaciones cotidianas, en vez de afirmar cosas como “Es una persona excelente”, “Es un buen tipo”, o incluso meramente decir que alguien es “compasivo” o “valiente”; deberíamos usar atribuciones locales del tipo “compasivo con alguien a quien le caen unos papeles cuando se ha encontrado una moneda” o “valiente con respecto a navegar con un amigo cuando el tiempo es malo”, juicios

⁴⁹ Doris (2002), p. 145.

⁵⁰ Kunda (1999), p. 499.

⁵¹ Doris (2002), cap. 4.

evaluativos estrechamente circunscritos a rasgos locales.⁵² En general, la recomendación de no simplificar en exceso las evaluaciones de una persona es importante y vale la pena tenerla en cuenta. Sin embargo, creo que —con todo lo problemático que puedan ser— no puede renunciarse por principio a la tendencia hacia cierta generalización de nuestros juicios evaluativos. Parece que si las atribuciones de rasgos tienen sentido es porque van más allá de la mera conducta observada, porque presuponen una regularidad —yo diría que *evaluativa* en sí misma; esto es, no meramente estadística— que fundamenta la atribución. Asimismo, el uso de palabras sencillas en las atribuciones de rasgos de carácter es un modo resumido de describir estas regularidades duraderas, cuya correcta interpretación depende, sin duda, del contexto y del contraste con otros sujetos. Por ejemplo, decir que alguien es valiente sólo adquiere un significado más o menos claro en un contexto determinado, que provea unos estándares evaluativos, y fije qué contaría como no valiente o qué *cantidad* de valentía se está atribuyendo al sujeto, que posibilite cierta comparación entre sujetos y contextos. Por otro lado, el uso, en el lenguaje cotidiano, de frases como la que propone Doris parece pragmáticamente inviable.⁵³

Con esto último, sin embargo, no pretendo decir que no haya razones para oponernos a tesis de la *integración evaluativa* de los rasgos de carácter. Recordemos que Doris afirmaba que esta era otra de las características de la Concepción Tradicional (o globalismo) que los resultados situacionistas arruinaban. Esta idea está relacionada con la tesis tradicional de la *unidad de la virtud*, cuyo pedigrí se remonta a Sócrates, Platón y Aristóteles; si bien es cierto que, desde siempre, no ha faltado la contro-

⁵² Véase Doris (2002), cap. 6. Recordemos que el eliminativismo acerca de los rasgos de carácter (globales) conlleva la eliminación de toda referencia a estos rasgos en nuestro pensamiento y lenguaje.

⁵³ Agradezco a Eduardo Ortiz que llamara mi atención sobre esto último.

versia en torno a ella. En parte, la cuestión se dirime en relación a cómo de unitaria sea caracterizada esta unidad. La definición de Doris, como una relación probabilística en cuanto a la concurrencia de rasgos de carácter evaluativamente similares es bastante liberal, y puede que, si mi objeción a su eliminativismo funciona, no esté autorizado a negarla. En cualquier caso, sí que hay una lectura de la idea de *fragmentación del carácter* que destacaré más adelante.⁵⁴ Pero primero quiero resumir las objeciones que he planteado en este apartado y decir algo más sobre la naturaleza de los rasgos de carácter.

La naturaleza plural de los rasgos de carácter

Mi rechazo del eliminativismo acerca de los rasgos de carácter ha constado de dos movimientos. Por un lado, si interpretamos radicalmente los resultados experimentales, de modo que estemos justificados a *eliminar* los rasgos de carácter; deberíamos hacer lo propio con los restantes factores personales. Si el nihilismo conductual a que esto nos conduciría no parece en absoluto atractivo, entonces no podemos eliminar nada de lo considerado —lo cual no excluye, más bien al contrario, que todos estos factores personales no queden *tocados* por los resultados situacionistas. Por otro lado, la individuación e identificación de situaciones *similares* en las que un agente actúa y a partir de las cuales podríamos inferir rasgos locales no es menos compleja, dados los estándares altamente exigentes que el eliminativista propugna —en particular, para los rasgos globales, pero que sería irrazonable no extender también a los locales—; teniendo en cuenta, en especial, que aspiramos a identificar situaciones similares,

⁵⁴ Adams ha propuesto lo que llama concepción *modular* de los rasgos de carácter y virtudes que puede compatibilizar la naturaleza fragmentaria del carácter y de los mismos rasgos y virtudes individuales con la aspiración a su integración evaluativa. Véase Adams (2006), caps. 8, 9 y 10.

no respecto a estímulos objetivos, sino respecto a construcciones subjetivas de ellos, sensibles a las más ligeras variaciones. Creo que con esto, el eliminativismo queda desacreditado, aunque no cierto deflacionismo moderado acerca de todos los factores personales con una función directa en la regulación de la conducta.⁵⁵ En la siguiente sección, precisaré algunas consecuencias particulares de esta conclusión.

Pero hay un problema con la misma noción de rasgo de carácter que está funcionando en el argumento eliminativista que no me resisto a apuntar. Con independencia de su caracterización, la concepción asumida por el eliminativista es más estrecha de lo que confiesa. En particular, para que el argumento funcione, es necesario presuponer que los rasgos de carácter son *disposiciones conductuales directas*.⁵⁶ Sin embargo, los rasgos de carácter no sólo constituyen patrones de conducta, sino también de pensamiento y sentimiento. Recordemos que Aristóteles los consideraba disposiciones a actuar, *desear* y *sentir*.⁵⁷ En este sentido, si nos fijamos concretamente en el grupo de experimentos sobre influencia social, podemos ver que si bien un gran porcentaje de personas actuaron de igual

⁵⁵ Otras posiciones que incluyo en el campo deflacionista (no-eliminativista) son Flanagan (1991), Goldie (2000), Adams (2006) y probablemente Morton (1980). Por ejemplo, interpretando los resultados del experimento de Milgram, Flanagan afirma: “Los miembros de ambos grupos [los que obedecieron y los que no] tienen todo tipo de disposiciones psicológicas que interactuaron de un modo complejo en la situación ideada por Milgram. Qué rasgos tienen, y exactamente cómo caracterizarlos y ponerlos juntos —tanto individual como colectivamente— defiere dramáticamente de persona a persona. Las personalidades de los miembros de ambos grupos son sensibles a la situación. Aunque lo son de diferente manera. No debemos ser tan ingenuos como para pensar que la principal variable que diferencia a las almas obedientes de las no obedientes es un único rasgo inflexible, y sin duda no uno hecho exclusivamente de fibra moral.” (1991, p. 295.)

⁵⁶ Véase Adams (2006), cap. 8, sobre esta cuestión.

⁵⁷ Quiero resaltar que mi réplica al eliminativismo ha sido interna a la misma concepción *directamente conductual* de los rasgos de carácter asumida en el debate. Por otro lado, parece que para poder finalmente eliminar los rasgos de carácter, el escéptico debería mostrar además que todos los rasgos de carácter son esencialmente disposiciones directas a actuar o, alternativamente, aportar evidencia empírica independiente contra los no directamente conductuales.

manera, parece que lo hicieron mostrando más o menos desazón y contradicción en relación a su propia conducta; es decir, se resistieron más o menos a actuar como actuaron, dado que es plausible pensar que percibieron su conducta como más o menos adecuada. Podría decirse que sus respuestas fueron más o menos *naturales* en relación a su carácter. Creo que esto marca una diferencia significativa entre ellos que remite a diferencias de carácter —diferencias que, de hecho, los experimentos no estaban diseñados para registrar en sus resultados, pues desatienden las conductas particulares a la hora de seguir uno u otro curso de acción, así como las creencias, deseos, emociones, etc. expresadas por los participantes con respecto a las diferentes alternativas. Eso sí, todas estas diferencias de carácter fallaron en convertirse en acciones efectivas.

En todo caso, los resultados no dejan por ello de ser altamente relevantes. De hecho, tiene razón Doris al afirmar que “mayoritariamente se considera que los rasgos de carácter, y las virtudes en particular, comportan disposiciones a la conducta.”⁵⁸ Aunque no tengan porqué ser todas disposiciones conductuales *directas*, sí que han de serlo por lo menos indirectamente. Su papel en la regulación de la conducta no deja de ser fundamental; si bien la conducta es algo bastante más complejo que los grandes cursos de acción registrados en los resultados oficiales de los experimentos considerados.

En la siguiente sección, trataré de explicitar importantes consecuencias de los resultados experimentales situacionistas para la determinación de la responsabilidad moral y cómo esto afecta, en particular, a la suerte moral circunstancial.

⁵⁸ Doris (2002), p. 15.

6.3. Circunstancias y evaluación

La exposición y discusión anterior de los experimentos situacionistas y su repercusión para nuestra concepción de los rasgos de carácter nos sirve, si no para rechazar su existencia, sí para precisar su naturaleza. Y recordemos que la investigación de la naturaleza de los rasgos de carácter era una condición previa para evaluar la tesis de que el locus último de la evaluación moral debían constituirlo los rasgos de carácter. Si algo ha quedado claro es que la que podemos llamar *dialéctica carácter-acción*, esto es, la relación entre poseer un rasgo de carácter determinado y llevar a cabo una acción determinada relacionada (en las situaciones apropiadas para la manifestación de ese rasgo) es harto compleja. En particular, parece que la misma realización de una acción meritoria, más allá de la posesión del rasgo virtuoso asociado, o la abstención de llevar a acto un rasgo de carácter vicioso, son méritos o deméritos extra que deben ser considerados para la asignación de responsabilidad moral.

A lo largo de esta sección, trataré de rechazar directamente la idea de que debemos juzgar a las personas exclusivamente en virtud de sus rasgos de carácter, o incluso motivos o intenciones.

6.3.1. Intenciones, planes e indeterminación. Significado de la suerte circunstancial

Tras nuestro recorrido por la bibliografía experimental situacionista, resulta patente que la presión situacional es un fenómeno del que ningún agente puede escapar. De hecho, cómo uno responde a las situaciones particulares en las que se ve envuelto, incluso a las más exigentes, parece que forma parte de aquello por lo que es y debe ser evaluado. Sin embargo, como vimos más arriba, no todo el mundo tiene esto tan claro. Una manera de afrontar esta situación es esta: si realmente estamos tan *inde-*

fensos ante las situaciones, entonces esto refuerza la idea de que para evaluar con justicia al agente no deberíamos excluir sus acciones del juicio moral merecido. Nótese que esta posición no necesita apelar a los rasgos del carácter, sino que puede limitarse a los motivos y razones del agente; con lo que sería consistente con el eliminativismo acerca de los rasgos de carácter. La idea sería que, si resulta que los rasgos de carácter son más bien ficciones y la conducta en situaciones particulares es tan influenciada, mejor refugiarnos en los motivos e intenciones que el agente posee y que, presumiblemente, adquiere tras concienzudas deliberaciones, libres de la influencia situacional.

Consideremos el caso de una infidelidad *no querida*; esto es, el caso de alguien que no quiere ser infiel, pero que, a pesar suyo, acaba siéndolo, a causa de *intensas presiones situacionales*. Supongamos que esta descripción es correcta y no se da ningún tipo de autoengaño en el agente respecto a su motivación. La cuestión sería: ¿es responsable el infiel de su infidelidad? Si lo es, ¿en qué punto de su historia cabría situar la responsabilidad? Parece que no puede serlo la infidelidad misma, pues actuó bajo presiones situacionales que, según la idea anterior, eximirían de la responsabilidad. La responsabilidad se hallaría más bien en la *negligencia* de no haber evitado la situación que desencadenó en contacto sexual.⁵⁹ Por ejemplo, en haber aceptado ir a cenar con cierta persona,

⁵⁹ El caso y esta supuesta solución son de Doris; véase Doris (2002), p. 144. En esta línea, Doris nos ofrece, como el *gran* consejo que se sigue de los resultados experimentales situacionistas, la idea de que debemos *evitar* las “ocasiones próximas al pecado” o “circunstancias éticamente peligrosas” (p. 147). Pero este consejo no puede ser más vacío, pues si los resultados experimentales deben generalizarse a todas las situaciones de nuestras vidas, de modo que cualquier situación sea susceptible de ser “éticamente peligrosa”, entonces el consejo es, además de inútil, imposible de seguir. La inacción no es una alternativa. Lo que se requeriría, más bien, es la especificación de qué circunstancias son susceptibles de ser éticamente peligrosas y a causa de qué efectos, para que podamos tratar de evitarlas. Si, en último término, la única lección moral que el situacionismo puede ofrecer es recordarnos que “el mundo es un lugar moralmente peligroso”, enton-

cuando podía sospechar que ese resultado era posible o probable. Sin embargo, esto resulta contraintuitivo: normalmente diríamos que si el agente es culpable lo es por ser infiel, no por no haber evitado la situación en la que podía caer en la infidelidad. Lo que no debe es cometer la infidelidad, cómo llegue a ella (o la evite) es su problema. De hecho, en sí mismo, ir a la cena no es ninguna *negligencia*, lo es, podríamos decir, sólo si resulta en infidelidad. (En el capítulo siguiente, ahondaré en la relación entre negligencia y resultado negativo.)

Propongo que recuperemos, en este punto, el ejemplo de los dos simpatizantes nazis, caso paradigmático de suerte moral circunstancial. Hemos visto que el adversario de la suerte moral afirmaba que lo realmente importante para su evaluación moral no es cómo cada uno de ellos finalmente actúa, sino que si Adolf hubiese estado en las mismas circunstancias que Rudolf, se habría integrado en el nazismo en la misma medida y habría cometido los mismos horrendos crímenes. Siendo esto así, ambos son igualmente responsables. Pero, ¿realmente se puede decir que deberíamos juzgar a uno tan duramente como al otro? Y, ¿en virtud de qué merecen ser tratados igual?

Si ambos realmente merecen ser tratados igual, no puede ser por sus acciones, pues el emigrado no realizó jamás las acciones por las que condenamos al no emigrado; ni por sus decisiones, pues tampoco tomó las decisiones inmediatas de llevar a cabo esas acciones. Lo único por lo que parece que pueden ser igualmente culpables, es por compartir unas mismas intenciones generales o rasgos de carácter. Pero, si la compara-

ces hay poco que aprender de él. Kamtekar (2004) y Sabini y Silver (2005) han insistido en esto de un modo muy perspicaz. Por mi parte, en el apartado siguiente, trataré de detallar algunas repercusiones concretas del situacionismo.

ción ha de ser significativa, no puede negar la fundamental indeterminación del contenido de estas intenciones de tipo general; por no hablar, como hemos visto, de la misma indeterminación de los rasgos de carácter respecto a acciones concretas. Pero centrémonos en las intenciones. Propongo que llamemos a estas intenciones generales “planes”. Los planes nos facilitan el esfuerzo deliberativo, permitiéndonos deliberar con antelación, así como la coordinación de nuestras acciones, por medio del establecimiento de compromisos con ciertas acciones futuras.⁶⁰ Nuestros planes establecen una agenda que restringe futuras deliberaciones. A partir de esta agenda habrá que ir determinando intenciones particulares que concreten nuestro plan general. Por ejemplo, el plan de pasar las vacaciones en Londres tiende a cerrar la cuestión de dónde pasar las vacaciones, a menos que las circunstancias cambien significativamente. Asimismo, el plan nos lleva a deliberar acerca de los medios para ir hasta Londres, y a ir precisando cada pieza del plan en su momento.

En este sentido, en el caso de los pro-nazis, ambos, como mucho, pueden compartir el plan general e indeterminado —en el sentido de que, como todo plan, no requiere de la determinación exacta de cada una de sus partes— de colaborar con el nazismo, pero resultaría implausible estipular que ambos comparten las mismas intenciones *concretas*, ya que cada uno vive en un entorno y tiene unas posibilidades inmediatas de acción diferentes. Podemos categorizar, también, esta diferencia en términos de la distinción entre intenciones *determinables* e intenciones *determinadas*. Aunque ambos compartan la intención determinable de “colaborar con el régimen nazi”, hay una diferencia significativa entre su determinación en

⁶⁰ Para alguien que ha defendido destacadamente la articulación de intenciones en planes, véase Bratman (1987). Nada de lo que digo depende de las particularidades de este análisis.

términos de, digamos, “pregonar en Argentina las bondades del régimen nazi” —y actuar en consonancia— y en términos de “quemar una tienda judía en Berlín”, o “atacar y golpear judíos” o mil cosas peores —y actuar en consonancia. Más allá de compartir una misma intención general, estas intenciones determinadas distintas parece que marcan una diferencia en la evaluación que cada agente merece. Evidentemente, el nazi emigrado podía tener la intención determinada de “quemar una tienda judía en Buenos Aires”, y actuar en consonancia, pero esta seguiría siendo una intención determinada diferente, dado que la situación política de Argentina era distinta a la de Alemania en ese tiempo y no podía significar lo mismo quemar una tienda judía en un lugar u otro. Alternativamente, el agente también podría tener la intención determinada de “quemar una tienda judía en Berlín”, pero entonces ni podría actuar en consonancia, ni parece que esta intención pueda tener la misma significación que si uno la tiene en el lugar donde puede hacerla efectiva. Se podría todavía recurrir a otras intenciones todavía más determinadas que reforzarían más la diferencia, como por ejemplo “ordenar a un judío entrar en una cámara de gas”, “asesinar comunistas la Noche de los Cuchillos Largos”, etc. No parece que tenga sentido decir que el emigrado puede incluso tener *realmente* la intención determinada de “ordenar a un judío entrar en una cámara de gas”, en Buenos Aires, lejos del contexto de las cámaras de gas —puede que ignorando incluso la misma existencia de éstas. En definitiva, si es inverosímil decir que ambos agentes pueden compartir incluso las mismas intenciones determinadas, entonces hay una diferencia significativa entre ellos, en sus mismas intenciones, que les hace merecedores de juicios morales distintos. (En un caso de infidelidad involuntaria como el anterior, el agente no tiene ni el plan de ser infiel. Por lo único que se le puede juzgar, y es por lo que de hecho le juzgamos, es por haber sido infiel.)

El adversario de la suerte circunstancial, sin embargo, puede responder que los rasgos e intenciones de un agente se manifiestan de muy diversas maneras. Y que, en muchos casos, lo que importa no es la manifestación misma, sino qué nos indica determinada conducta acerca de la voluntad o carácter del agente. Aunque el emigrado no cometa los horrendos crímenes que comete su contraparte, *si el agente potencial es tan semejante al real, entonces habrá algo en su conducta que demandará la misma respuesta*, que nos llevará a pensar que, si hubiera estado en la situación del otro, habría cometido los mismos o parecidos crímenes. En concreto, manifestaciones orales y actos de apoyo al nazismo, como la proferencia de injurias contra los judíos, el elogio de la política que está poniendo en marcha Hitler, etc., denotarían que el nazi emigrado comparte el carácter del no emigrado. Pero, aún así, hay una diferencia moral clara entre ambos. Además, aunque se estipule que el emigrado comparte los mismos rasgos de carácter relevantes con su contraparte, esta estipulación no puede extenderse plausiblemente más allá de sus rasgos de carácter o de sus intenciones más generales. Es por esto que he insistido en el matiz de la verosimilitud y significado de la comparación.

De hecho, el negador podría aún empeñarse en defender la idea de que si el ejemplo tiene fuerza es debido a la ambigüedad con que es descrito. Es decir, si afináramos los detalles de los ejemplos, veríamos que o bien ambos contrapartes son tan semejantes que no habrá entre ellos ninguna diferencia moralmente relevante, o bien son tan diferentes que no podremos considerarlos contrapartes en el sentido relevante para el caso.⁶¹ Hay un sentido en que, de esta manera, *disolveríamos* la cuestión, pero sería al precio de pasar por alto el significado mismo de los casos de la suerte circunstancial. Esto es, cómo las diferentes circunstancias abren

⁶¹ Véase Pritchard (2006) para una defensa clara de esta posición.

diferentes oportunidades para la acción, así como también establecen diferentes exigencias sobre el agente. En particular, ciertas circunstancias constituyen pruebas, algunas de ellas altamente exigentes, a las que sólo algunos de nosotros nos vemos sometidos. En una variación del caso anterior, Nagel afirma:

Los ciudadanos corrientes de la Alemania nazi tuvieron la oportunidad de comportarse heroicamente oponiéndose al régimen. También tuvieron la oportunidad de comportarse mal, y la mayoría de ellos son culpables por no haber pasado esta prueba. Pero es una prueba a la que los ciudadanos de otros países no fueron sometidos, de lo que resulta que incluso si ellos, o algunos de ellos, se hubieran comportado tan mal como los alemanes en semejantes circunstancias, simplemente no lo hicieron y por lo tanto no son igualmente culpables.⁶²

El verse envuelto en tan exigentes circunstancias —lo que sin duda no depende de uno mismo— constituye una prueba moral cuya superación parece que sólo puede venir del heroísmo. De hecho, un futuro héroe, sin circunstancias que le brinden la acción heroica, nunca llegará a ser un héroe.

Pero las circunstancias no sólo repercuten en cómo uno actúa, sino también en qué decisiones toma, en qué intenciones y actitudes se forma, etc. En esta dirección, Hanna Arendt afirmaba que “en ciertas circunstancias la más decente persona podría convertirse en un criminal”, como fue el caso de las circunstancias sociales que se dieron durante el período nazi. Con esta idea, asociada a la misma *banalidad del mal*, Arendt pretendía rebatir las interpretaciones prevalecientes, según las cuales las atrocidades nazis emanan de una malvada voluntad de hacer el mal. No cabe duda de que la idea de que el mal es producido por gente intrínsecamente

⁶² Nagel (1979), p. 34.

malvada puede ser muy consoladora, sobretodo cuando va acompañada de la simplista distinción binaria entre buenas y malas personas —que habitualmente coincide con la distinción entre *nosotros* y *ellos*. Pero, en la mayoría de los casos, es manifiestamente falsa. Y es muy conveniente superarla, si lo que queremos es reflexionar sobre la psicología humana real. Debemos darnos cuenta de cuán intrincado suele ser el conjunto de situaciones que conducen a alguien hacia el mal. Por supuesto, la banalidad del mal lleva asociada la *banalidad del heroísmo*. En ciertas circunstancias, también cualquiera podría convertirse en un héroe.

En todo caso, esto no es óbice para reconocer la diferencia, por ejemplo, entre Eichmann y quienes no somos Eichmann. Aunque otros no hayan tenido que afrontar las mismas pruebas, “incluso si ellos, o algunos de ellos, se hubieran comportado tan malvadamente como [él] en semejantes circunstancias —afirmaba Nagel—, simplemente no lo hicieron y por lo tanto no son igualmente culpables.” Tampoco Arendt pretende arruinar esta distinción con su insistencia en la *banalidad del mal*, más bien al contrario. Ello resulta evidente tras leer su imaginario alegato final contra Eichmann, en el que afirma:

[...] has hablado de una responsabilidad por igual, en potencia, no en acto, de todos aquellos que vivieron en un Estado cuya principal finalidad política fue la comisión de inauditos delitos. Poco importan las accidentales circunstancias interiores o exteriores que te impulsaron a lo largo del camino a cuyo término te convertirías en un criminal, por cuanto *media un abismo entre la realidad de lo que tú hiciste y la potencialidad de lo que otros hubiesen podido hacer*.⁶³

⁶³ Arendt (2006), pp. 405-6. Mis cursivas.

Sin la menor duda, todas éstas son cosas que el fenómeno de la suerte moral circunstancial nos revela.

6.3.2. Más repercusiones: la opacidad de los propios motivos e intenciones

Por otro lado, me gustaría aún comentar determinadas repercusiones que los resultados experimentales situacionistas tienen sobre nuestra autoimagen. En general, dos son los efectos fundamentales que las presiones situacionales producen sobre la conducta de los sujetos objeto de experimentación. En primer lugar, los experimentos hacen patente que a menudo ignoramos los verdaderos motivos de nuestras acciones, lo que apunta al hecho de que *nuestros* motivos serían frecuentemente meras *racionalizaciones* de nuestra conducta. Esto, sin duda, tiene que ver con el autoconocimiento de nuestros motivos, intenciones, razones, etc. Como vimos en la descripción de los experimentos, muchos de ellos incluyen un cierto desconocimiento por parte del agente de sus propios motivos. Esto es especialmente chocante en los estudios sobre los efectos en la conducta del estado de ánimo en que uno se encuentra; por ejemplo, en el experimento de Isen y Levin en el que se manipulaba el estado de ánimo de los sujetos mediante el hallazgo o no de una moneda. El segundo efecto (que se sigue del primero) es la merma de nuestra autonomía: si no actuamos movidos por las razones que queremos que nos muevan, o lo hacemos en menor medida de lo que pensamos, no hay duda de que nuestra autonomía, cuanto menos, se resiente.⁶⁴ Esto puede deberse a que desconocemos lo que verdaderamente nos mueve a actuar, como en el caso anterior; o puede deberse, alternativamente, a que no conseguimos convertir nuestros motivos o intenciones en acciones. En el experimento de Milgram, en el

⁶⁴ Nelkin (2005) y Nahmias (2007) han insistido sobre estos dos puntos.

que los sujetos pueden reflexionar, lo dramático es que muchos de ellos, aunque deciden no continuar, no pueden transformar su decisión en la acción correspondiente. Por supuesto, la falta de (auto)conocimiento está igualmente presente, más aun cuando hasta un nutrido grupo de expertos en la materia juzgaron que era imposible que se diese un nivel tan alto de obediencia.

En todo caso, no hay duda de que, en conjunto, los experimentos descritos arruinan definitivamente una concepción excesivamente autosuficiente —es innegable que idealizada— del agente y su carácter, esto es, altamente independiente de los factores externos. Y, por supuesto, esto apunta en la misma dirección que la idea de la *banalidad del mal* (y el heroísmo), según la cual, que uno se convierta en criminal (o en héroe), es algo que depende de manera muy destacada de las circunstancias en que se halle. Por otro lado, los datos experimentales nos llevan a remarcar la gran significación moral de la conducta misma, pues la mera posesión de un buen carácter o unas buenas intenciones no asegura mucho.

No obstante, tampoco debemos dejarnos atrapar por el pesimismo. Los efectos registrados por los experimentos son graduales y admiten mayor o menor resistencia. Esta resistencia puede aumentarse con el aprendizaje moral que su mismo estudio individualizado nos aporta. Además, la propia experiencia de cómo funcionan determinadas fuerzas psicológicas nos puede proveer con la habilidad de resistir situaciones futuras del mismo tipo. Esto viene corroborado por el testimonio de una participante en los experimentos de Milgram, que describió como determinante en su resolución y temprana resistencia a aplicar las descargas eléctricas su experiencia vital en la Alemania nazi.⁶⁵ En general, los resultados situacionistas deberían ser tenidos muy en cuenta tanto por aquellos

⁶⁵ En Milgram (1974), cap. 4.

que aspiren a caracterizar adecuadamente la conducta humana en ciertas situaciones, como por aquellos que pretendan proponer soluciones *realistas* a problemas morales.⁶⁶

Pero sí que hay algo más que vale la pena destacar, en conexión con el escepticismo acerca del autoconocimiento a que apuntan los resultados experimentales. En general, la realización de una acción suele conllevar un refuerzo importante del compromiso con la intención de que es producto. Pero, además, puede suponer el mismo descubrimiento de cierta volición o rasgo. Piénsese, por ejemplo, en alguien que odia sinceramente a los inmigrantes y sostiene que todos ellos deberían ser expulsados del país. No será muy difícil que cada cual le ponga cara a esta persona. Conocemos a personas —familiares, amigos, conocidos— que tienen opiniones extremistas respecto a ciertos temas sociales y políticos; opiniones que defienden sinceramente, e incluso están convencidos de que si estuvieran en los lugares adecuados para ello actuarían de acuerdo a estas creencias sin dudarlos. Escojamos a un hombre, que rechaza tan frontalmente a los inmigrantes que no soporta ni cruzarse con ellos. Pero que un día tiene que hacer cola en una oficina de desempleo y se ve rodeado, por todos lados, de inmigrantes. Sin embargo, al entrar en contacto directo con alguno de ellos, se muestra cordial e incluso se preocupa por uno al que le ha dado un pisotón. En un caso como éste, el hecho de encontrarse con inmigrantes en la cola del paro puede marcar una diferencia en el jui-

⁶⁶ A este respecto, es curioso oír que alguien se pregunte, en relación a la violencia de género, cómo es posible que todavía exista gente que sea testigo de malos tratos en su vecindad y no lo denuncie. Si atendemos a los estudios psicológicos sobre la difusión de la responsabilidad, los casos de este tipo no son difíciles de explicar. Por otro lado, las propuestas de solución deberían tomar en cuenta las sutiles fuerzas psicológicas actuan-

cio que el agente merece.⁶⁷ Parece que respecto a muchos rasgos, si no en todos, debemos esperar a la manifestación conductual para poder juzgar correctamente a su poseedor. Es simplemente presuntuoso, en general, suponer que podemos saberlo todo acerca de la situación concreta de una persona, incluido acerca de uno mismo. De hecho, normalmente sólo tenemos contacto con una representación de nuestro yo, que difiere en gran medida de nuestro yo real. Esto se manifiesta, por ejemplo, en el hecho de que a menudo dejamos de actuar del modo en que deseábamos y pensábamos que íbamos a actuar. Nunca somos totalmente transparentes, ni a los demás, ni a nosotros mismos.

Sin embargo, el negador puede mantenerse firme replicando que seguimos estando ante un problema meramente epistémico: de nuevo, la suerte puede afectar aquí a nuestra comprensión de cuál sea el carácter del agente, pero no a la maldad o bondad misma de su carácter. Podría replicarse que lo que muestra la conducta anterior es que aquel hombre no odiaba realmente a los inmigrantes, sino que simplemente estaba equivocado acerca de su actitud para con ellos. No obstante, aquí la cuestión es, a mi entender, que centrarnos exclusivamente en la intención o volición que el agente se autoatribuye internamente, o que incluso podríamos registrar si tuviéramos acceso directo a su mente, puede llevarnos a error. Nuestra conducta en muchas situaciones nos hace *reinterpretar* las intenciones o rasgos internos previos, lo cual sólo es posible a partir de la

⁶⁷ Un caso paralelo, o que al menos apunta en la misma dirección, es el de la conducta de Huckleberry Finn cuando ayuda a escapar al esclavo Jim, contra sus mismas convicciones morales. Es el hecho de afrontar unas circunstancias determinadas —la disyuntiva de denunciarlo o permitir que se escape— lo que hace que descubramos el mérito moral de Huck, que va más allá de sus mismas creencias acerca de sus propios valores morales en relación a la esclavitud —los imperantes en su sociedad, que es incapaz de cuestionar o revisar. Una importante discusión de este caso, que ha dado origen a una extensa bibliografía, se encuentra en Bennett (1974). Arpaly y Schroeder (1999) llaman a esto “acrasia inversa”. Arpaly (2003), cap. 3, vuelve a hacer uso de este ejemplo. Todos concuerdan con la comprensión general del caso aquí presentada.

prueba que supone estar en unas determinadas circunstancias y actuar de una u otra manera. Pero esta reinterpretación no es realmente un descubrimiento, sino que es en sí misma parte constitutiva de la intención o rasgos internos. No es que antes de aquella acción el hombre interpretase erróneamente su actitud y ahora haya llegado a la correcta interpretación. Sino que muchos rasgos están sencillamente indeterminados hasta que son fijados por comportamientos determinados. En este sentido, el carácter depende, parcial pero inevitablemente, de las acciones del agente.

En general, consideraciones como las anteriores ponen de relieve el que puede ser considerado como el problema central del juicio moral: normalmente, nuestro conocimiento del estatus moral de una persona, incluyéndonos nosotros mismos, es indirecto e incierto. Parece que, habitualmente, es más oscuro que en otros contextos cognitivos, y eso resulta frustrante. Si además le sumamos la gran exigencia de justicia que demanda el juicio moral referido a personas, se vuelve realmente difícil *emitir un veredicto* sobre el estatus moral de una persona, incluso en el caso de uno mismo. Las personas son, en un sentido, radicalmente incognoscibles, lo cual nos deja perplejos y nos fascina a un tiempo. Oscar Wilde, igualmente atraído por esta cuestión, ideó un relato en el que el retrato del protagonista, Dorian Gray, tiene el remarcable poder de revelar inmediata e infaliblemente la condición moral del retratado, para lo cual “habría de poder seguir su pensamiento [de Dorian] en sus lugares más recónditos”. Cuanto más feo aparece en el retrato, más depravado es el protagonista, sin ambigüedades posibles. El ideal del *alma visible* parece que debería ser la meta del juicio moral. Sin embargo, esa opacidad no es meramente epistémica, sino epistémica y metafísica a la vez.⁶⁸ La indeterminación

⁶⁸ Esto va en la misma línea de los resultados del capítulo 3. La incoherencia de la noción de merecimiento último, cuando es llevada al extremo, tornaban borrosa la distin-

resulta ser un rasgo de la misma alma humana, si se me permite la expresión.

6.3.4. Sobre los dilemas morales

Hay otra forma en que las circunstancias pueden marcar una diferencia moral. Me refiero a las situaciones en las que el agente se ve enfrentado, sin haber cometido falta alguna, al hecho de tener que elegir entre dos cursos de acción alternativos que le plantean un conflicto de obligaciones o valores, bien sea porque no puede cumplir con ambos o bien sea porque debería evitarlos igualmente. En todo caso, el resultado es que no puede evitar hacer algo incorrecto. Estoy hablando, por supuesto, de los dilemas morales. Por ejemplo, en el conocido caso del estudiante de Sartre, la devoción personal de éste hacia su madre choca con su compromiso con la Resistencia antinazi. El estudiante tiene dos obligaciones de tipo moral, igualmente poderosas y mutuamente incompatibles, entre las que debe decidir. Decidirse por una, supone dejar de satisfacer la otra. De igual modo, a Agamenón se le presentó en Áulide un conflicto ineludible entre el amor por su hija y la devoción por el bienestar de sus hombres, viéndose forzado a sacrificar literalmente uno de estos dos bienes. E incluso hay ocasiones en las que la elección es entre dos instancias alternativas de un mismo valor, como en el caso de la protagonista de *La decisión de Sophie*, que se ve obligada a elegir entre uno de sus dos hijos.

Hay un sentido obvio en el que estar atrapado en las circunstancias de un dilema moral es cuestión de (mala) suerte. Esto es, dado que estas circunstancias hacen más difícil para el agente el actuar correctamente, los dilemas morales constituyen casos en los que la suerte influye en la

ción entre cuestiones estrictamente epistémicas y cuestiones estrictamente metafísicas en esta esfera.

evaluación moral. En otras palabras, son casos de suerte moral circunstancial, al menos, en el sentido restringido de que son situaciones moralmente *adversas*, o difíciles, que sólo tiene que afrontar quien se encuentra en tal situación. Pero además, hay otro sentido en el que los dilemas morales serían casos especiales de suerte circunstancial.⁶⁹ La diferencia de los dilemas morales *genuinos* con los casos de suerte circunstancial considerados hasta el momento es que en los primeros el agente no tiene más remedio que actuar incorrectamente, sin que haya modo alguno de evitarlo. En los casos más usuales de suerte circunstancial, en cambio, el agente afronta una especie de prueba moral. Ciertamente, en el caso de los ciudadanos de la Alemania nazi, se trató de una prueba moral durísima y terrible; pero, aún con todo, uno podía pasar la prueba. Mientras que un dilema moral genuino se caracteriza por ser una situación en la que (1) una persona P tiene categóricamente (absolutamente, *habiendo sido todo considerado*) obligación de hacer A y puede hacer A, (2) P tiene categóricamente (etc.) la obligación de hacer B y puede hacer B, y (3) P no puede hacer tanto A como B. En este tipo de situaciones, si son posibles, uno no tendrá más remedio que actuar incorrectamente.

Sin embargo, para muchos la misma existencia de conflictos genuinos de deberes y obligaciones es inconcebible. Paradigma de esta posición es Kant, quien mantuvo que:

los conceptos de deber y obligación como tales expresan la *necesidad* práctica objetiva de ciertas acciones, y dos normas en conflicto no pueden ser necesarias a la vez: si es nuestro deber actuar según una de esas normas, entonces actuar de acuerdo con la opuesta no es nuestro deber y es incluso contrario a deber. [...] la

⁶⁹ Véase Nagel (1979), p. 70, nota 9.

razón obligante más fuerte prevalece (*fortior obligando ratio vincit*).⁷⁰

La idea es que cuando se nos presenta un conflicto de deberes, uno de los deberes siempre prevalece sobre el otro. En realidad sólo estamos ante un conflicto de deberes *prima facie*. En concreto, para Kant los deberes morales no son obligaciones de actuación, sino de fines. La necesidad de los fines obligatorios no constriñe la acción, sino la voluntad. Así, no hay conflicto moral alguno porque las razones de la obligación no obligan; la experiencia del conflicto moral establece, más bien, la tarea de deliberar, pero sólo la resolución revelará la obligación.⁷¹ Si uno elige el mejor curso de acción disponible, desde un punto de vista moral actuará correctamente, sin que pueda ser culpado de nada. De este modo, quienes niegan la existencia de dilemas morales genuinos, rechazan que éstos constituyan un tipo particular de casos de suerte moral circunstancial. Para esta concepción, los dilemas morales, con mucho, sólo podrían constituir casos de suerte circunstancial del tipo usual o estándar, en los que uno afronta situaciones altamente exigentes desde el punto de vista moral, pero en los que hay una opción correcta y otra incorrecta.⁷² Así, si bien es cierto que Agamenón se enfrentó a una prueba cruel, y tuvo que tomar una decisión terrible; no obstante, una sola de las opciones era la correcta.⁷³

Pero, por otro lado, el mismo fenómeno de la suerte moral ha sido utilizado para defender la existencia de dilemas morales genuinos. Como

⁷⁰ Kant, *Metaphysik der Sitten* (“Doctrina de la Virtud”), *Ak.* 6: 224.

⁷¹ Esta es, al menos, la lectura que Barbara Herman realiza de la posición de Kant. Véase Herman (1993), cap. 8.

⁷² Sobre el tipo de suerte moral circunstancial que constituirían los dilemas morales, véase Statman (1993), pp. 20-1.

⁷³ En relación a este caso, Martha Nussbaum ha afirmado que acabar con la vida de su hija Ifigenia fue lo más racional para Agamenón —por fuerte que fuera la tentación (moral) de decantarse por la opción de salvarla a su hija—, pues de otro modo habrían muerto todos, incluida Ifigenia. Nussbaum (1986), p. 34.

veremos en el capítulo siguiente, Bernard Williams ha insistido en el hecho de que la existencia de sentimientos como el pesar-del-agente refuerza la *realidad* de la experiencia de la tragedia moral. Hay hechos básicos, ligados a los dilemas morales, como son el sentimiento de culpa y el remordimiento, el anhelo de *restitución* ante el daño cometido o los restos morales (*remainders*) debidos a la obligación incumplida, que una teoría moral satisfactoria no puede dejar de atender. Williams enfatiza esta idea, en su clásico ensayo *Ethical Consistency*:

Me parece que una crítica fundamental a muchas teorías éticas es que sus explicaciones del conflicto moral y sus resoluciones no hacen justicia a hechos como el pesar y otras consideraciones relacionadas: básicamente porque borran de escena el *deber* en virtud del cual uno no actúa.⁷⁴

Además, afirmar la posibilidad de los dilemas morales genuinos es ver el mundo como moralmente trágico; esto es, como un mundo en el que los sujetos no pueden evitar, en ocasiones, la culpa moral.⁷⁵

En el siguiente capítulo profundizaré en la cuestión de las decisiones dilemáticas y el conflicto de valores.

⁷⁴ Williams (1973d), p. 175. Teorías éticas como las de Kant y Herman.

⁷⁵ La conexión entre las nociones de suerte moral y dilemas morales genuinos es estrecha, y a menudo llevan también unida la defensa de la inconmensurabilidad de valores. Williams, Nagel y Nussbaum comparten estas tres ideas. Véase Williams (1973d); Nagel (1979b), cap. 9; y Nussbaum (1986), pp. 47-50.

Dónde estamos y adónde vamos

En este capítulo, he defendido la realidad de la suerte moral circunstancial. Para ello, he presentado y atacado la tesis según la cual el locus de la atribución de responsabilidad moral no deben ser las acciones particulares del agente, sino sus rasgos de carácter o intenciones. He tratado de mostrar que circunscribir la evaluación exclusivamente a los rasgos internos del agente resulta injustificado, tanto porque desprecia elementos externos moralmente relevantes, como porque se basa en una concepción inadecuada del carácter.

En concreto, y respecto a este último punto, he presentado datos empíricos, en forma de experimentos psicológicos, que muestran que factores situacionales aparentemente triviales pueden ejercer una influencia sorprendente en cómo actuamos; por ejemplo, en si ayudamos a alguien o no en una situación determinada. Aunque rechacé que estos datos lleguen a arruinar la misma existencia de los rasgos de carácter, sí que acepté que suponen un importante reto para la acción por razones y el conocimiento de los propios motivos, y con ello para la evaluación moral, aunque moderadamente.

En definitiva, los casos de suerte moral circunstancial, inquietantes y fascinantes a la vez, nos muestran importantes elementos de nuestra experiencia moral, que el negador inevitablemente se pierde. En el capítulo siguiente, seguiré insistiendo en este hecho en relación a la suerte moral resultante.

7. Suerte resultante

Justificación retrospectiva y emociones morales

- 7.1. Contra la suerte moral resultante
 - 7.1.1. El primado de la voluntad
 - 7.1.2. Negligencia, riesgo y previsibilidad
- 7.3. Justificación y retrospectividad. La decisión de Gauguin
 - 7.2.1. Justificación racional y justificación moral. Conflicto de valores y necesidad
 - 7.2.2. Deliberación y resultado
 - 7.2.3. El lugar de la suerte moral: ¿qué pasó con la familia Gauguin?
- 7.3. Juicios, actitudes y emociones
 - 7.3.1. El pesar-del-agente, un sentimiento moral
 - 7.3.2. Actitudes reactivas y juicio moral
 - 7.3.3. Agencia y daño

...si antes de cada acción pudiésemos prever todas sus consecuencias, nos pusiésemos a pensar en ellas seriamente, primero en las consecuencias inmediatas, después, las probables, más tarde las imaginables, no llegaríamos siquiera a movernos de donde el primer pensamiento nos hubiera hecho detenernos.

José Saramago, *Ensayo sobre la ceguera*

La crueldad viciosa que rellenaba los confines de su boca apareció, sin duda, en el preciso instante en que la chica bebía el veneno, cuandoquiera que fuese. ¿O, quizás, era indiferente a los resultados? ¿Tomaría tan solo en consideración lo que pasaba en el interior de su alma?

Oscar Wilde, *El retrato de Dorian Gray*

El último tipo de suerte moral que nos queda por considerar es la suerte resultante o consecuencial; esto es, la clase de suerte implicada en cómo resulten las cosas a partir de las acciones de un agente y su influencia en

el juicio moral merecido. Si lo que he defendido en los capítulos anteriores es correcto, éste es el único tipo al que uno aún puede resistirse.

Hay que reconocer que las intuiciones más poderosas contra la suerte moral lo son, en particular, contra la suerte moral consecuencial. Por ejemplo, Adam Smith proponía, en su *Teoría del sentimiento moral*, que todo acto sea juzgado exclusivamente en virtud del motivo y la intención del agente.¹ También Kant abogó por este tipo de posición en su *Fundamentación*:

La buena voluntad no es buena por lo que efectúe o realice, no es buena por su adecuación para alcanzar algún fin que nos hayamos propuesto; es buena sólo por el querer, es decir, es buena en sí misma.

[...] esa buena voluntad como una joya brillante por sí misma, como algo que en sí mismo posee su pleno valor. *La utilidad o la esterilidad no pueden ni añadir ni quitar nada a ese valor.*²

Así, el valor o merecimiento del agente debería determinarse, exclusivamente, en función de su voluntad, de la calidad de su voluntad —según la conocida expresión.

Vimos en el capítulo anterior que la idea de la existencia de una voluntad única y transparente sobre la que fundar las atribuciones de responsabilidad moral era discutible. No obstante, dejaré aquí de lado esta cuestión. En este capítulo, responderé específicamente a los argumentos de los negadores de la suerte moral consecuencial. Empezaré precisando las respuestas de éstos ante supuestos casos de este fenómeno. A continuación, trataré de replicar tomando en consideración los argumentos de

¹ Véase Adam Smith (1790/1976), libro II, § iii. Su posición final es más completa de lo que esta breve referencia puede mostrar. Para una interpretación de su posición en relación a la suerte moral, véase Russell (1999).

² *Fundamentación*, secc. I; Ak. IV: 411. Traducción de García Morente; mis cursivas.

Williams a favor de la suerte moral; en particular, el del carácter retrospectivo de la justificación y el del pesar-del-agente. En la última sección, me abordaré la cuestión del papel del daño efectivamente producido.

7.1. Contra la suerte moral resultante

Como vimos en el capítulo segundo (sección 2.2), dentro del tipo de la suerte consecucional, había tres subtipos básicos: el caso de quienes tratan de producir intencionalmente cierto resultado, el de quienes producen un resultado que podían prever pero que no pretendían y el de los casos de decisión bajo incertidumbre. Ejemplos paradigmáticos de los dos primeros subtipos son los casos de los asesinos y de los conductores ebrios, respectivamente. En 7.1.1 me ocupo del primer subtipo; mientras que 7.1.2 trata del segundo y el tercero.

7.1.1. El primado de la voluntad. Intención e intento

En el primer caso, dos sujetos tratan de asesinar a alguien. Ambos llevan a cabo todos los preparativos para ello e intencional y deliberadamente tratan de acabar con la vida de sus respectivas víctimas. Finalmente, uno logra su objetivo y mata a su víctima, pero el otro falla el tiro, a causa, por ejemplo, del vuelo de un pájaro —como en el capcioso caso de Nagel— o simplemente porque el disparo resulta no ser mortal. El negador de la suerte moral, el adversario del hecho de que el carácter certero o mortal, o no, del disparo pueda marcar una diferencial moralmente relevante, situará el locus de la atribución de responsabilidad moral en la volición e intención de los agentes. Lo mismo sucedería en el caso de que, en el último instante, uno de ellos no hubiera podido disparar a causa, por ejemplo, de la presencia de un obstáculo —que, como vimos en el capítu-

lo anterior, algunos tenían por un caso ejemplar de suerte circunstancial.³ En ambas variantes del caso, los contrapartes serán igualmente responsables y culpables, en tanto que ambos intentaron acabar con la vida de una persona. La diferencia legal entre homicidio e intento de homicidio es moralmente irrelevante.

No obstante, uno es responsable de matar a alguien, mientras que el otro no puede ser responsable *de* matar a nadie, pues no ha matado a nadie. Uno es responsable por tratar de matar a alguien y matarlo, mientras que el otro lo es sólo por el intento. Así, parece que uno sería responsable por más cosas que el otro. Sin embargo, el adversario de la suerte moral puede responder de este modo:

No niego en absoluto que [uno] sea *más responsable* que [el otro]. Debemos distinguir el *grado* [*degree*] en la responsabilidad de alguien, de su *alcance* [*scope*]. [...] Mi tesis es que [uno] y [otro] cargan con el mismo grado de responsabilidad, más allá del hecho de que la responsabilidad de [uno] tenga un alcance mayor.⁴

Ambos son igualmente responsables de su intención e intento de acabar con la vida de una persona. “El hecho de que el intento de [uno] no tuviese éxito, mientras que el del [otro] sí, es irrelevante en relación a cuán censurables son.”⁵

³ Véase el capítulo 6, sección 6.1.1; y Zimmerman (2002), pp. 563-5.

⁴ Zimmerman (2002), pp. 560-1. Otra manera formular la distinción es esta: “P es censurable por más sucesos que P*” contra “P es más censurable que P*.” El desafortunado debe ser censurado por más cosas que el afortunado, más cosas que nos indican que debe ser evaluado negativamente en esta ocasión, pero esto no implica que deba ser evaluado negativamente “en mayor medida” que el desafortunado. Ver Zimmerman (1987), p. 227. En la misma línea, el *argumento epistémico* afirma que la muerte indica la necesidad de una evaluación negativa sólo indirectamente, pero una indicación indirecta es igualmente una indicación.

⁵ Zimmerman (2002), p. 561.

No obstante, hay algo moralmente horrendo, causar la muerte de alguien, que uno hizo y el otro no, y que es precisamente por lo que el “intento de” es moralmente censurable. Negando la relevancia moral de esta diferencia, ¿no se estará trivializando la (causación de la) muerte misma? Nagel afirmaba, parece que con razón, que hay algo “moralmente significativo” en torno a la diferencia entre ambos, aunque para el adversario de la suerte consecuencial esto no puede “marcar una diferencia” respecto a la responsabilidad moral.⁶ Que ambos sean igualmente responsables es una consecuencia de aceptar que no podemos ser moralmente responsables de lo que no está bajo nuestro control; ninguno de ellos tenía más control sobre las consecuencias de sus actos:

Lo mismo que con la responsabilidad, ocurre con el control: debemos distinguir entre el grado y el alcance. [Uno] tuvo *control de más cosas* que [el otro] (su control tuvo un alcance mayor), pero no tuvo *más control* de lo que sucedió del que tuvo [el otro] (tuvo el mismo grado de control). En tanto que el grado de responsabilidad rastrea el grado de control, [ambos] deben ser declarados igual de moralmente responsables.⁷

Ciertamente, hay un sentido por el que la censurabilidad de ambos es muy semejante, aunque uno no matara a nadie. En particular, son igualmente culpables de tener la misma intención de matar a sus respectivas víctimas —de ser iguales en carácter— y de haber llevado esta intención a acto. En

⁶ Para Zimmerman, esto sería objeto de un juicio *deóntico*; el tipo de juicio que se basa en el resultado o consecuencias de hecho de las acciones de un agente. Pero en cuanto al juicio de responsabilidad moral —de *grado* de responsabilidad, en particular— ambos están a la par. Véase Zimmerman (1987), p. 227; y (2002), p. 562.

⁷ Zimmerman (2002), p. 562. Afirma Zimmerman que, en el sentido pertinente de “control”, un agente puede tener control sobre una cosa aunque suceda por causa de alguna otra cosa que el agente no controla. Reconoce que la suerte es en gran medida ineliminable de nuestras vidas, pero eso no implica que marque una diferencia moral, ni que no podamos tener algún control. Zimmerman (2002), p. 562, nota 27.

este sentido, son igual de viles. No se trata sólo del hecho de tener la intención, e intentarlo, sino del grado de elaboración y realización de intención e intento. Su intención firme les llevó a emprender los cursos de acción necesarios para alcanzar su meta: consiguieron el arma, la cargaron y prepararon, investigaron a la víctima, fueron a encontrarla, se le aproximaron, buscaron el momento idóneo, le dispararon, etc.

Podríamos decir, pues, que lo que cuenta es la intención y el intento. ¿Pero son éstos realmente tan centrales? Comparemos los casos anteriores con un nuevo ejemplo. Pensemos en una madre que cruza la calle por un paso de peatones con su hija pequeña cogida de la mano. Un coche que circula a toda velocidad no respeta la señalización y no frena, atropellando a la niña pero no a la madre. El coche se detiene unos metros más allá. La mujer, que ve a su hija tendida en medio de la calle llena de sangre, es invadida por el odio más furibundo. Se aproxima el conductor del coche sin soltar palabra. El joven, queriendo ofrecerle consuelo, al tiempo que expresando sinceramente su pesar, trata de ayudar. Pero ella se alanza sobre él y trata de clavarle unas tijeras en el estómago. La historia tiene dos finales alternativos, ambos más allá del control de la mujer. En los dos la niña se salva sin grandes secuelas. Pero en R1 el joven resulta herido de muerte, mientras que en R2 el intento se ve frustrado por los reflejos del joven, que puede echarse atrás y bloquear la mano de la mujer. Lo que me interesa es el juicio que merecería la mujer en cada uno de los resultados. La intención es manifiesta sea cual sea el resultado. Pero parece que en R1 la mujer merecería una condena, que de ningún modo merecería en R2. La mujer tiene el derecho de reaccionar con rabia y odio ante la conducta del joven, así como de expresar ostensiblemente su reacción. Pero asesinarlo no es una respuesta a la que esté moralmente autorizada.

Sin duda, se puede replicar que aquello que tiene de condenable la actitud de la mujer no es matar o no al hombre, sino *querer* matarlo —la intención (y el intento) de matarlo—, que es igual en ambos casos y, en principio, no está subordinado a la suerte (consecuencial).⁸ Pero así, la mujer merecerá igual condena en ambos casos, cosa que choca con la presumible diferencia de reacciones que nos provoca uno u otro resultado. Por un lado, el caso parece paralelo al de los asesinos: si entre el asesino “exitoso” y el asesino frustrado sólo hubo la diferencia de la intromisión o no del pájaro en la trayectoria de la bala, entre R1 y R2 sólo hay la reacción más o menos rápida del joven. De todos modos, la reacción que nos provoca uno y otro caso no es la misma. En este caso, la diferencia en los resultados presumiblemente cambiará de forma más significativa la consideración moral que el agente merece. Y creo que la diferencia apunta a cómo la realización o no de intención e intento expresan y redefinen su carácter. En R1, la mujer se convertirá en una *asesina*, mientras que eso no sucederá en R2, donde la cuestión será más pronto liquidada. (Me ocuparé de esta última idea en la sección final; y de la tesis de que nuestra reacción ante los diferentes resultados es, o puede ser, una buena guía para determinar el juicio merecido, en la sección 7.3.)

Hay que remarcar que esta posición no niega que las intenciones o voluntad del agente jueguen un papel importante en la asignación de responsabilidad moral; sólo que las intenciones y la voluntad no son ni el único elemento a considerar, ni siempre el más importante. Esto es más evidente respecto al subgrupo de casos de suerte consecuencial ejemplifi-

⁸ Podría replicarse que la acción es producto de una pérdida momentánea de lucidez: la mujer estaría fuera de sí, no controlaría lo que hace, y por lo tanto no sería moralmente responsable de ninguno de los dos resultados posibles, pues no lo sería ni de su misma acción. Pero, entonces, puedo modificar el caso. Por ejemplo, la mujer maquinó la venganza fríamente mientras su hija estaba en coma en el hospital debatiéndose entre la vida y la muerte. Aquí, no podríamos hablar de pérdida momentánea del control volitivo.

cado por los conductores temerarios, significativamente distinto del anterior. Los conductores no tenían la *intención* de, ni *intentaron*, matar a nadie —al menos, no en el sentido ordinario de “intención” e “intentar”, que sí reconocemos en los asesinos. En tales casos, la asignación de responsabilidad no puede depender de su intención e intento, sino de otro factor.

7.1.2. Previsibilidad, riesgo y negligencia

Recordemos el caso de los conductores ebrios. Ambos conductores se emborrachan y deciden conducir. Supongamos que su decisión de conducir tras haber bebido, aún a pesar de los efectos del alcohol, es libre. Con ella ambos cometen una negligencia; esto es, asumen un riesgo moralmente injustificado con respecto a la vida de otras personas. El adversario de la suerte moral no necesita más, aquí termina la historia en lo que a la evaluación moral se refiere. El alcance de la responsabilidad moral dependerá de la *previsibilidad o probabilidad de las consecuencias*, dada la información relevante accesible al agente en el momento de la acción, o *ex ante*.⁹ Ambos conductores ebrios son igualmente negligentes, ambos someten a los demás a un riesgo moralmente injustificado. Podemos definir la negligencia como la asunción injustificada de un riesgo —esta asunción está injustificada cuando el daño potencial de la acción pesa más que sus beneficios potenciales, dadas las probabilidades.¹⁰ Conducir ebrio es, claramente, una acción en la que los daños potenciales son mucho mayores que los posibles beneficios. Además, el atropello o accidente automovilístico es una consecuencia previsible del hecho de conducir bajo la influencia del alcohol, y esto es algo que normalmente la gente sabe.

⁹ Ver Rosebury (1995), pp. 518–519; y también Enoch y Marmor (2007).

¹⁰ Ver Domski (2005), p. 533.

Sin embargo, una parte importante de nuestra asignación cotidiana de culpabilidad parece surgir del hecho de haber matado a una persona, más que exclusivamente de la mera temeridad del conductor —lo cual no sucede, o no sucede de la misma manera, en el caso de los asesinos. El *argumento epistémico* nos brinda una explicación desenmascaradora de este fenómeno: si juzgamos a un conductor de modo diferente al otro no es porque ambos merezcan un juicio diferente, sino porque su conducta no nos muestra con la misma claridad que lo que merecen sea igual. No siempre está claro cuáles son las intenciones de una persona, ni su compromiso para con un curso de acción determinado. Por ello, el atropello es el mejor indicador que tenemos para tratar a uno de los conductores más duramente. A este respecto, Richards afirma que si un conductor no causa ningún daño puede ser porque estaba suficientemente alerta o porque su velocidad era segura; o quizá no, y no estaba alerta o su velocidad era excesiva, o ambas cosas. Pero, en cualquier caso, si no ha habido ningún daño real la cuestión permanece incierta. Es sólo cuando este conductor atropella a alguien que se torna claro que su atención no era suficiente o que su velocidad era excesiva; en definitiva, que no sopesó correctamente el riesgo de daño a otras personas.¹¹ Sin embargo, este tipo de argumento hace depender muy estrechamente el merecimiento del agente de las situaciones que podemos llamar (supuestas) “situaciones epistémicamente transparentes”, esto es las situaciones en las que el agente produce de hecho un daño: el conductor ebrio que provoca un accidente, el asesino que consigue acabar con la vida de su víctima, etc. Lo que se afirma es que la producción real de determinado resultado —un daño (o un bien o beneficio, en el caso del elogio)— es un indicador fiable del compromiso del agente con dicho resultado. Sin duda, y como ya insistí, las intencio-

¹¹ Ver, en particular, Richards (1986), p. 171-2.

nes del agente no están a menudo claras (incluso para el mismo agente), como no lo está su compromiso con el curso de acción seguido, y normalmente estamos en una situación epistémica mejor para adscribir una temeridad cuando hay un accidente que cuando no; pero producir un resultado dañino puede ser tan accidental con respecto a las intenciones del agente como lo puede ser intentarlo y no conseguirlo. El atropello no tiene por qué mostrar que el conductor tomara una actitud temeraria. Quizá con una evidencia más clara, como sería el caso si se hubiera podido contar con un registro de velocidad que nos indicara que el coche no sobrepasó la velocidad recomendada, podría hacerse patente que el merecimiento era bien otro.¹² No obstante, Richards reconoce que hay ocasiones en las que, se produzca o no finalmente el accidente de tráfico, el sentido común respecto al tiempo de reacción y visión humana y la naturaleza de unas condiciones determinadas es suficiente para concluir que cualquiera que condujese de esta manera, en este lugar, en este tiempo, conduciría temerariamente. En tal caso, el accidente mismo no juega ningún papel, pues no aumenta la base evidencial para creer que la conducción era temeraria.

De todos modos, lo fundamental de esta posición (con independencia de nuestro acceso epistémico) es que el merecimiento depende exclusivamente de la negligencia o temeridad cometida. La acusación moral debería ser, en el caso de los conductores, sólo por conducción temeraria, sin contar el homicidio por imprudencia. Sin embargo, en contra de esta posición puede citarse la intuición opuesta de que una parte importante de nuestra asignación de merecimiento parece surgir del hecho mismo de haber efectivamente causado o no daño —en forma de un accidente, un atropello y un homicidio—, más allá de la temeridad misma. Y dando una vuelta de tuerca más a esta intuición, podría decirse incluso

¹² Para otra réplica de este tipo, véase Adler (1987), pp. 247-8.

que en algunos casos parece que la misma negligencia sólo existe cuando el resultado es efectivamente dañino. Consideremos esta última idea con la ayuda del siguiente ejemplo.

Una madre que está bañando a su bebé.¹³ Entonces, alguien llama a la puerta. Ella está esperando a su anciano padre. Está sola en casa, y decide dejar al bebé por un momento para ir a abrir la puerta. Al abrir la puerta se encuentra con un vecino que quiere informarla de una avería en la puerta del garaje. La mujer trata de deshacerse educadamente del vecino, pero éste insiste en repetir los pormenores de la avería. La mujer se entretiene más de la cuenta. Finalmente, logra deshacerse del vecino, cierra la puerta y vuelve corriendo al baño. Cuando llega, dos son los resultados posibles. En el primero, su bebé sigue chapoteando alegremente en el agua. En el segundo —el trágico—, el bebé se ha hundido bajo el agua. Podría decirse que en un caso como éste la negligencia o temeridad sólo *existe* si el resultado es malo. En el escenario en el que la madre vuelve al baño y su bebé sigue chapoteando en el agua, no parece razonable hablar de ningún tipo de negligencia por parte de ésta. Es sólo en el caso en el que el bebé resulta herido o incluso muere, cuando se puede hablar de una negligencia cometida por la madre. La explicación de que la negligencia o temeridad sólo *exista* cuando el resultado es malo es que el resultado dañino eleva el estándar de exigencia, de anticipación o precaución debida, para esa situación determinada. La base lógica de esta diferencia podría ser la siguiente: comúnmente, todos tomamos riesgos, riesgos morales incluso, pero supondría una concepción demasiado rigurosa de la moralidad y la vida en general, y con toda seguridad impracticable, igualar la censura moral que la madre merece en ambas situaciones, con independencia del resultado de hecho. El estándar de exigencia no puede ser coti-

¹³ Me baso en un ejemplo de Nagel (1978), p. 30.

dianamente tan alto. No podemos juzgar a la madre con la misma dureza en ambos casos. Tampoco podemos tratarla con la misma indulgencia.

Pero veamos otro caso, creo que aún más claro, que apunta en la misma dirección. Se trata de un caso de falta de atención o distracción. Un joven acaba el turno de noche en el hospital y conduce hacia su casa. Mientras circula por la avenida, entre decenas de coches, enciende la radio y trata de hallar una cadena en la que den las noticias. Le cuesta encontrar una, centra su atención en ello, mira durante unos segundos la pantalla del dial, y entonces oye un golpe contra su coche. Cerca de donde frena, hay una moto en el suelo y junto a ella un motorista que parece malherido. En realidad, ha fallecido.¹⁴ No puede decirse que cometiera ninguna gran temeridad, pero se descuidó. Este descuido podría haber sido de lo más intrascendente, pero no lo ha sido: su coche se desvió un poco de su trayectoria y embistió al motorista que —¡vaya mala suerte!— circulaba a su lado.

Por otro lado, la misma idea de riesgo (moral) es muy resbaladiza. Imaginemos que alguien acude a una fiesta y se emborracha. Él sabe que cuando está bajo la influencia del alcohol se desenfrena y puede realizar trastadas muy diversas. ¿Cuál es el riesgo que está infligiendo sobre los demás asistentes a la fiesta, meramente en virtud de su receptividad al alcohol? Pongamos que se acerca a alguien y le echa encima su cubata. Alternativamente, imaginemos que, para gastar una broma, coge a su mejor amigo y trata de tirarlo a la piscina desde el balcón de la casa; el amigo es arrojado por el balcón y, por supuesto, no llega hasta la piscina. Supuestamente, nuestro protagonista debe ser censurado igualmente en el caso de que cualquiera de estos resultados se dé, pues lo que importa es el riesgo a que injustificadamente somete a los demás, y en ambas situacio-

¹⁴ Tomo el caso de Corbí (2003), p. 38.

nes es el mismo. No hay duda (y esta es mi posición) de que, en tanto que responsable de emborracharse, debe cargar con las consecuencias de sus acciones subsiguientes, pero no estrictamente en virtud del riesgo moral, sino de las consecuencias *de hecho*. ¿Cómo justificar si no la diferencia en culpabilidad de una situación a otra (verter el cubata sobre alguien o arrojar a alguien por el balcón), si el locus de la responsabilidad es el riesgo mismo? ¿O es que estamos dispuestos a afirmar que, en ambos casos, la culpabilidad es la misma? Parece que no. Además, para ello se debería poder definir con exactitud y *ex ante* la idea de riesgo, sin reinterpretarse en virtud de alguna consecuencia determinada de la acción evaluada. Con esto no quiero negar, por supuesto, que haya casos en los que el riesgo es claro y determinable con independencia de los resultados —la defensa de la suerte moral consecuencial sólo necesita comprometerse con la tesis de que *en ocasiones* las consecuencias marcan una diferencia en la evaluación moral.

Pero hay aún un tercer subtipo de suerte consecuencial, estrechamente ligado al que acabamos de considerar: el propio de las decisiones bajo incertidumbre. Por ejemplo, los líderes de un grupo opositor clandestino deben decidir cuándo atacar al gobierno del cruel régimen dictatorial que subyuga su país. La decisión de lanzar finalmente la insurrección no puede basarse en el conocimiento cierto de cómo acabará todo; pues no existe este conocimiento. Inevitablemente, estamos ante una decisión que implica un gran riesgo. Los factores que pueden influir en el resultado son muy variados. Si la insurrección tiene éxito, se obtendrá la libertad; pero, por el contrario, también puede suceder que el intento quede frustrado y los dirigentes del régimen tiránico respondan con una represión sangrienta que provoque un coste amplio en vidas y sufrimiento. Así, en muchos

casos, la elección se torna muy difícil, a causa del hecho de que el resultado no puede ser previsto con certeza. O imaginemos aún el caso de un experto en negociación que, arriesgando la vida de los rehenes, se niega a satisfacer las demandas del secuestrador, *convencido* de que con ello logrará que éste se derrumbe. Puede que crea que tiene buenas razones para hacer esto, pero el riesgo es muy grande. Si la estrategia sale bien y el secuestrador se acaba derrumbando y se entrega, parece que el negociador tendrá poco o nada que reprocharse; pero si resulta que el secuestrador era psicológicamente más fuerte de lo que el negociador intuía y responde matando a un primer rehén, la actitud del negociador hacia su propio comportamiento cambiará, o debería cambiar.

No hay duda de que este tercer tipo está muy estrechamente conectado con el anterior. Tanto en los casos de decisión bajo incertidumbre como en los de negligencia, el negador de la suerte moral circunscribe la asignación de responsabilidad moral a la decisión misma de actuar, a la perspectiva *ex ante* —la cual toma exclusivamente en consideración la previsibilidad o probabilidad de las consecuencias en virtud de la información relevante deliberativamente accesible al sujeto en el momento de decidir. Por el contrario, mis objeciones al negador apuntan al papel de la perspectiva retrospectiva en la asignación de responsabilidad moral. Esto conecta con la idea principal de Bernard Williams al respecto: el carácter fundamentalmente retrospectivo de la justificación; así como su principal argumento a favor de ella: el pesar-del-agente. Es por ello que dedicaré las secciones siguientes a estas cuestiones.

7.2. Justificación y retrospectividad. La decisión de Gauguin

Como empezamos a ver en el capítulo segundo, la defensa de Williams de la suerte moral es indirecta y nada sencilla. Ello se debe, en gran medida, al carácter polifacético de sus ejemplos. La meta principal de su argumento es refutar la idea, especialmente vinculada a la concepción kantiana, de que la moralidad es inmune a la suerte. Para esta concepción, hay un conjunto de nociones, como son las de moralidad, racionalidad, justificación y valor supremo, que están estrechamente conectadas, lo cual “tiene un gran número de consecuencias para la evaluación reflexiva por parte del agente de sus propias acciones”. Dos elementos especialmente destacados de esta concepción son, por un lado, que en último término no puede ser cosa de suerte que alguien esté justificado a hacer lo que hizo y, por el otro, que el valor moral es siempre el valor supremo. El argumento de Williams tiene dos objetivos centrales. El primero es demostrar que la justificación racional puede, en ocasiones, chocar con la justificación moral, hasta el extremo de acabar imponiéndose sobre ésta última. Si esto es posible, si la justificación moral no siempre prevalece, entonces el valor moral no será siempre el valor supremo. El segundo objetivo es mostrar que, así, la suerte puede introducirse incluso en la esfera de la moralidad, pues si resulta que la justificación racional puede quedar sometida a la suerte y, dado que ésta podía imponerse sobre la justificación moral, entonces el valor moral mismo puede quedar sometido a la suerte. (Finalmente, como movimiento defensivo, Williams plantea su conclusión en forma de dilema, como veremos más abajo, en la sección 7.2.3.) En particular, la justificación racional está sometida a la suerte a causa de su pretendido carácter “esencialmente retrospectivo”.

Tras refrescar las características fundamentales del caso de Gauguin, empezaré abordando la cuestión del choque entre justificación ra-

cional y justificación moral, y valor moral y no-moral; para, a continuación, pasar a ocuparme del papel de la perspectiva retrospectiva. Terminaré volviendo a la cuestión misma de la suerte moral. Mi crítica del argumento de Williams no será tanto destructiva como (re)constructiva; esto es, no trataré de rechazarlo o invalidarlo, sino de separar lo que tiene de poderoso y convincente de lo que resulta ambiguo e innecesario, variando y completando algunos de sus aspectos, y respondiendo a ciertas críticas razonables. Respetaré sus principales conclusiones, pues comparto decididamente con Williams tanto la idea de que en ocasiones el valor no-moral puede prevalecer sobre el valor moral, como la idea de que lo que suceda después de la decisión y las acciones del agente es relevante para la justificación y la evaluación moral —aunque variaré el énfasis en algunos aspectos.

Recordemos que Williams nos proponía el siguiente caso, basado en la vida del pintor Paul Gauguin —si bien hay que distinguirlo del Gauguin real, en tanto que se trata de un experimento mental. El Gauguin de Williams debe decidir entre estas dos opciones:

- (1) quedarse con su familia, de la que se siente responsable; o
- (2) trasladarse a una isla de los Mares del Sur, donde piensa que podrá desarrollar su capacidad artística y convertirse en un gran pintor.

Se trata de dos opciones excluyentes —llevar a cabo una hace imposible llevar a cabo la otra—, y ambas se le presentan como valiosas. Finalmente, Gauguin se decidirá por (2), esto es, emigrará al Pacífico para perseguir su pasión por la pintura, abandonando a su mujer e hijos. Dado este escenario, la cuestión que se plantea es si podemos decir que la decisión que tomó Gauguin estaba o no justificada. Es decir, ¿fue racional para él,

dados sus intereses, actuar como actuó? Ciertamente, creo que no resulta del todo claro a qué se refiere exactamente Williams aquí. Como dije, muchos de sus ejemplos son tan polifacéticos que llegan a dificultar su correcta comprensión. Propongo que deslindemos la cuestión misma de la decisión entre (1) y (2), de su justificación para creer en (2). Comenzaré ocupándome del conflicto entre justificación racional y justificación moral, para pasar a continuación a tratar la cuestión del carácter retrospectivo de la justificación.

7.2.1. Justificación racional y justificación moral. Conflicto de valores y necesidad

La decisión de Gauguin guarda una semejanza formal importante con un dilema moral. Pensemos de nuevo en el famoso caso, expuesto por Sartre, del estudiante que debe decidir entre militar en la resistencia y ocuparse de su madre. El estudiante desea vengar a su hermano, asesinado en la ofensiva alemana de 1940, y combatir a los nazis; pero su madre vive con él y es su único consuelo en la vida. De este modo, el estudiante se ve a sí mismo ante un conflicto de obligaciones, que Sartre describe como la duda entre dos obligaciones morales de diferente tipo. Una de alcance limitado pero eficaz: la devoción personal a su madre; la otra de alcance mucho mayor, pero de incierta eficacia: intentar contribuir a la derrota del injusto agresor.¹⁵ En todo caso, lo relevante es que existen dos obligaciones de tipo moral, igualmente poderosas y mutuamente incompatibles, entre las que el agente ha de decidir. Sin embargo, en el caso de Gauguin no estamos ante una decisión entre dos alternativas morales, esto es, ante dos obligaciones de tipo estrictamente moral. Por ello, no podemos hablar de dilema *moral*. Gauguin tiene ante sí una opción que apare-

¹⁵ Ver Sartre (1946).

ce como la moralmente correcta (hacerse cargo de su familia) y otra que responde a lo que, provisionalmente, podemos llamar sus *intereses personales*. Claramente existe un conflicto, pero un conflicto entre tipos diversos de valores, uno de los cuales no es moral. No obstante, la intención de Williams no es presentarnos un dilema *moral*. En realidad, lo que pretende el caso es mostrar un conflicto *práctico* —un dilema práctico, podríamos llamarlo— entre dos valores que pertenecen a órdenes distintos. Se trata del choque entre:

- a) un valor moral: las obligaciones de Gauguin para con su familia; y
- b) un valor no-moral, relacionado con la autoperfección o florecimiento individual: su vocación artística, el proyecto sobre el que construye su vida.

Así, si consideramos que Gauguin hizo bien en actuar como actuó, aún a pesar de haber abandonado a su familia, ello mostrará que la obligación o el valor *moral* no siempre posee el valor más alto en la vida humana; esto es, que siempre que se dé un conflicto entre un valor moral y otro de cualquier otro tipo, el primero tenga que necesariamente terminar imponiéndose.

El defensor del carácter supremo del valor moral no concederá que Gauguin hizo bien sacrificando a su familia por su vocación. Gauguin debería haberse quedado con su familia, cumpliendo su obligación moral para con ellos; abandonarlos, sea cual fuere el motivo, es sencillamente inmoral. Más aún, la cuestión de cómo resulte el proyecto es irrelevante. El triunfo del proyecto sólo hará que el agente, *de hecho*, se preocupe emocionalmente menos de su decisión inmoral, cuando en realidad —*por derecho*— debería preocuparse en igual medida. Pero, en un caso como el

presente, parece que responder a la exigencia moral —someterse a ella— supone un sacrificio personal demasiado grande. Ello supondría sacrificar el propio proyecto vital, a aquello esencial a su desarrollo personal, que da sentido a su existencia: su vocación por el arte de la pintura y su cultivo en un entorno que prevé proclive para tratar de desarrollar al máximo sus potencialidades y conseguir la reforma del arte pictórico mismo.

Williams considera que es absurdo, o irrazonable, que la moralidad insista en el hecho de que una persona haga o deje de hacer X incluso ante el coste de desprenderse de aquello que, para esta persona, da sentido a las demandas morales mismas. El empeño de Gauguin por perseguir su vocación artística podría constituir incluso una *necesidad práctica* para él. Es fundamental en esta noción el hecho de que las “conclusiones de necesidad práctica alcanzadas con seriedad” constituyen “en un mayor o menor grado, *descubrimientos acerca de uno mismo*”.¹⁶ Williams afirma en otro lugar lo siguiente acerca de las necesidades prácticas de Ajax:

Las necesidades son internas, fundadas en el *ethos*, los proyectos, la naturaleza individual del agente, y en el modo que él concibe la relación de su vida con las demás personas.¹⁷

Y una vocación tiene todas las propiedades de las necesidades prácticas:

la comprensión de una vez de los propios poderes e incapacidades, y de lo que el mundo permite, y el reconocimiento de un límite que no es ni simplemente externo al yo, ni un producto de su voluntad.¹⁸

¹⁶ Williams (1981), p. 130; cursivas añadidas. Para una idea semejante teorizada por Harry Frankfurt (la de “necesidades volitivas”), véase la nota 29 del capítulo 5.

¹⁷ Williams (1993), p. 103; cf. p. 75ss.

¹⁸ Williams (1981), p. 130-1. Sigo aquí la exposición de Mulligan (2008).

Estaríamos, así, ante el conflicto entre el *deber* (moral) y el *tener que* (práctico). El deber moral de ocuparse de la propia familia y el tener que de la propia vocación.¹⁹

Por otro lado, el descubrimiento de la propia vocación puede ser negativo; es decir, puede consistir en el descubrimiento de lo que *no es para mí*. De hecho, en ciertas personas los descubrimientos de este tipo son exclusivamente negativos; como, por ejemplo, en el caso de Ulrich, el protagonista de *El hombre sin atributos* de Musil.²⁰ Además, suele ser gradual. En este sentido, el caso de Gauguin, tal y como lo expone Williams, puede resultar algo artificioso por cuanto parece que todo depende de una decisión determinada y puntual. En un momento *t* de su vida, Gauguin se lo juega todo a una carta: debe decidir entre marcharse o quedarse, y esta decisión parece ser definitiva. De hecho, este planteamiento contrasta con la historia real del verdadero Gauguin, que fue yendo y viniendo al Pacífico, hasta que convencido de su valía, acabó estableciéndose allí. El alejamiento de su familia fue progresivo y motivado, también, por las desavenencias de todo tipo con su esposa. La historia de Rembrandt puede resultar más *creíble* a este respecto. Rembrandt disfrutaba de una posición social elevada cuando pintaba *al gusto de la época*. Pero cuando comenzó a pintar a su estilo personal, fue perdiendo progresivamente el favor del público y, con ello, su estatus social como artista, así como sus bienes materiales. Saika, su esposa, le pidió que volviera a pin-

¹⁹ Puede resultar interesante resaltar la crítica que, en su estudio de la vocación de Goethe, realiza José Ortega y Gasset de la no distinción entre el *deber* moral y el *tener que* vocacional: “Es terrible pero innegable: el hombre que *tiene que ser* un ladrón y, gracias al esfuerzo virtuoso de su voluntad, consigue no ser un ladrón, falsifica su vida. Uno no debería, pues, confundir el *deber ser* de la moral, que reside en la región intelectual del hombre, con el *tener que ser* de una vocación personal, que está localizado en lo más hondo y primario de su ser.” Ortega y Gasset (1947), p. 406; tomado de Mulligan (2008).

²⁰ Véase Mulligan (2008) si se desea continuar con esta cuestión.

tar al gusto del público, como hacía Flinck, un mal ex alumno suyo que comenzaba a gozar de un gran prestigio. Rembrandt tenía que decidir entre: a) pintar al gusto de sus clientes, o b) seguir cultivando su propio estilo artístico. La fuerza de la opción a) procede, en parte, de su deber (moral) para con el bienestar económico de su familia; mientras que la de b) surge de su necesidad de ser fiel a su instinto artístico. De nuevo, el *deber* contra el *tener que*. El paralelismo con la decisión de Gauguin es significativo. Pero en este caso vemos como la decisión es paulatina. Rembrandt va convencándose (o descubriendo) progresivamente que su voluntad es seguir con su estilo propio y original. Pero la opción contraria sigue estando a su disposición, aunque por supuesto conforme va perdiendo más y más el favor del público, le resultaría cada vez más difícil volver a su estatus anterior.

En todo caso, el resultado es que la esfera misma de lo moral está de alguna manera subordinada al *tener que*, a la necesidad práctica, que vertebra nuestra vida; esto es, a nuestros proyectos vitales, vocaciones, deseos categóricos, que conforman una esfera más amplia que posee una relevancia fundamental. Sin ellos la moralidad misma no tendría sentido, pues el significado de la moralidad depende de estas cosas.²¹ Creo que esta idea es correcta, si bien el moralista podría contraatacar incidiendo en el consiguiente problema que su aceptación plantea: si realmente es absurdo que la moralidad demande de una persona el sacrificio de aquello que es de fundamental importancia para él, entonces la moralidad perderá toda su fuerza, pues cualquier cosa habría de ser permitida si es suficientemente significativa para el agente. Así, si, por ejemplo, un pederasta afirma, y cree, que lo que da significado a su vida es abusar sexualmente

²¹ Además de en Williams (1981), la tesis también es defendida explícitamente en la contribución de Williams a Smart y Williams (1983).

de menores, esto le habrá de ser permitido, o al menos estará justificado a actuar así. Obviamente, ésta es una consecuencia indeseable del punto anterior que cabría tratar de excluir. Se requeriría una teoría positiva que limite el rango de cosas que pueden jugar este papel. Puede que, a este respecto, resulte útil la siguiente distinción, propuesta por Susan Wolf, entre *lo que da sentido a la vida de las personas*, por un lado, y *el auto-interés, la felicidad o la satisfacción de preferencias*, por el otro. Afirma Wolf:

El significado en la vida, eso que nos da tanto una razón para vivir como una razón para estar concernidos por el mundo en el que vivimos, requiere una clase de conexión valorativa —*una relación en la que subjetivamente valoramos cosas o metas o actividades que son objetivamente dignas de valoración*— que nos engancha motivacionalmente al mundo y a las razones.²²

No puede valer cualquier proyecto, sino sólo aquellos que valoramos (subjetivamente) porque, en alguna medida, son *objetivamente valiosos*. Esta condición objetivista puede concebirse de un modo más o menos fuerte, si bien, con toda seguridad, sería rechazada por Williams. En todo caso, y aun con el problema de limitar el rango de cosas que pudiera constituir de manera justa lo que vertebra o confiere sentido a la vida de una persona, podemos reconocer que, en el caso de Gauguin, su vocación artística es un candidato apto para esta calificación. Perseguir una vocación por el arte es, sin duda, algo objetivamente valioso. Así, dado el sacrificio excesivo que le supondría permanecer con su familia y no poder desarrollar convenientemente su aptitud artística, Gauguin estuvo justificado a tomar la decisión que tomó. Y, de este modo, un valor no moral se impo-

²² Wolf (1997), p. 306. Mis cursivas.

ne sobre un valor moral; con lo que el valor moral es así tratado como un valor entre otros, y no como el valor incuestionablemente supremo.

Adicionalmente, cabe notar lo siguiente acerca de la naturaleza de la justificación racional que Williams invoca. Imaginemos que Gauguin tomara la decisión contraria, y decidiera permanecer con su familia en lugar de perseguir su vocación pictórica por las islas del Pacífico. En esta nueva situación, parece que efectivamente su decisión estaría moralmente justificada, pero ¿no lo estaría racionalmente? Una respuesta plausible podría consistir en decir: si el cultivo de su pasión por la pintura era algo fundamental en su proyecto vital, renunciar a ello estará racionalmente injustificado. En esta lectura, la decisión sería “irracional”, aunque moral. No obstante, este uso de “irracional” plantea serios problemas, ya que excluye una decisión moralmente correcta, lo cual parece contraintuitivo. Más adecuado sería abogar por un uso más amplio de “racional”, que permitiera la posibilidad de que dos decisiones contrarias pudieran ser igualmente racionales (igualmente permisibles racionalmente). Pero, por su lado, eso conllevaría que sería igualmente racional optar por cualquiera de las dos alternativas, en virtud de la misma evidencia disponible, pero no igualmente moral, ya que sólo una de las opciones está moralmente justificada. Posiblemente, la mejor manera de enunciar la cuestión es que Gauguin estaba racionalmente justificado a perseguir su vocación, aunque estaba moralmente injustificado a abandonar a su familia. En todo caso, no puede negarse que los propios deseos, proyectos y/o valores individuales e idiosincrásicos de Gauguin necesariamente entran en, y mediatizan, su deliberación práctica.

Con esto, si, dado el sacrificio excesivo que hubiera supuesto para Gauguin permanecer en Europa con su familia y no poder intentar des-

arrollar convenientemente su aptitud artística, reconocemos que estuvo (racionalmente) justificado a optar por perseguir su vocación; estamos entonces ante un caso en el que la justificación racional prevalece por encima de la moral, lo que muestra que el valor moral no es siempre el valor supremo.

No obstante, el choque entre la justificación racional y la justificación moral es una cuestión diferente de la del carácter retrospectivo de la justificación. Podríamos reconocer que la justificación racional puede en ocasiones imponerse sobre la justificación moral, pero que para ello el hecho de que se convierta finalmente o no en un gran pintor no tiene ninguna relevancia —por mucho que su situación devenga trágica si fracasa. Paso a ocuparme de esta cuestión.

7.2.2. Deliberación y resultado

Es obvio que Gauguin no puede estar completamente seguro respecto a si se convertirá en un gran pintor, o no, hasta que lleve adelante su intento para conseguirlo y acabe teniendo éxito, o no; aunque, por supuesto, puede disponer de datos que le indiquen si esta predicción es o no plausible. Resulta evidente que si no ha experimentado nada que le indique que sus capacidades artísticas son prometedoras —como la obtención actual de buenos resultados, el reconocimiento y encomio por parte de profesores y expertos, etc.— junto con una verdadera vocación por pintar, la decisión parecerá del todo descabellada. Pero, suponiendo que sí que ha tenido tales indicios razonables, y que éstos parecen apuntar a un futuro prometedor, ¿en virtud de qué podremos afirmar que la decisión de Gauguin estuvo justificada?

También podemos preguntarnos: ¿por qué los Mares del Sur? ¿Por qué una apuesta tan arriesgada? Parece que el Gauguin real creyó que sólo

abandonando el mundo occidental y refugiándose en el paraíso virginal de la Polinesia francesa podría desarrollar adecuadamente sus dotes artísticas, pues pensaba que la pintura necesitaba un cambio radical que la apartase del camino de la artificiosidad. El primitivismo y exotismo de las islas del Pacífico se le presentaba como el contexto ideal —que en parte rememoraba sus años de infancia transcurridos en Perú.

La tesis de Williams es que, dada esta situación, la única cosa que justificará (racionalmente) la decisión de Gauguin será el éxito mismo de su proyecto:

Si falla... habrá hecho lo incorrecto... en el sentido de que el haber hecho lo incorrecto en aquellas circunstancias no le da un fundamento para pensar que estaba justificado en actuar como actuó. Si tiene éxito, sí que tendrá el fundamento para este pensamiento.²³

Sólo si tiene éxito podrá pensar que estaba justificado en actuar como actuó. Pero como es imposible prever con antelación si la empresa tendrá éxito, no hay más remedio que reconocer que “[l]a justificación, si es que la hay, será esencialmente retrospectiva”.²⁴ Con ello, y en tanto que el éxito o fracaso del proyecto está fuera del control del agente, la justificación racional quedará irremediabilmente abierta a la intervención de la suerte. Si bien no a cualquier tipo de suerte. Lo que probaría que Gauguin “no está en lo correcto en su proyecto no sería sólo que éste [el proyecto] ha fallado, sino que él [el agente] ha fallado”.²⁵ La justificación de Gauguin no será cuestión de cualquier tipo de suerte. Hay aquí una distinción relevante entre suerte intrínseca y extrínseca. Si bien ambas son necesas-

²³ Williams (1981), p. 23.

²⁴ Williams (1981), p. 24.

²⁵ Williams (1981), p. 25.

rias para el éxito de su proyecto vital, y así para la justificación; sólo la suerte intrínseca está conectada con la falta de justificación.

En el caso particular de Gauguin, la suerte extrínseca es la suerte en relación a aquello que afecta a su proyecto pero que no está internamente relacionado con éste. Por ejemplo, que el barco que le lleve a Tahití naufrague. Mientras que la suerte intrínseca se identifica con el tipo de suerte originado a partir de sus decisiones y acciones, con la suerte en relación a su mismo proyecto artístico, a su triunfo o fracaso como artista.²⁶ La decisión de Gauguin podrá estar *injustificada* sólo debido a la intervención de la mala suerte intrínseca, pero no por ningún factor de mala suerte extrínseca. *Él* fallará, por ejemplo, si llega a Tahití y se entrega completamente a la pintura, pero no obtiene los resultados esperados.

Cabe destacar que la suerte intrínseca no tiene por qué depender exclusivamente del agente, y de hecho en la mayoría de casos no lo hace. En este sentido, el ejemplo de Gauguin puede llevar a confusión. En el caso de Ana Karenina la suerte intrínseca al proyecto está parcialmente fuera de ella. Su decisión de dejar a su marido por Vronski podría haber sido equivocada si, por ejemplo, éste se hubiera suicidado. Se trataría de un factor de suerte intrínseca —interna al proyecto— aunque no directamente dependiente de las decisiones y acciones de la misma Ana. Otro caso: imaginemos que una joven se encuentra lejos de su ciudad habitual de residencia. De repente sufre una dolencia y tiene que decidirse entre un listado de médicos del lugar en el que se halla. Allí no tiene a nadie de confianza a quien pedir consejo. Elige uno de los médicos y acude a su consulta. La diagnóstica y le informa que le han de realizar una pequeña operación. Ella se lo piensa, observa los títulos médicos del doctor, sus

²⁶ Williams (1981), p. 26-7. Como vimos en el capítulo 2, esto no tiene porqué ser siempre así.

calificaciones, y acepta. Hay un tipo de suerte intrínseca a este proyecto que claramente no depende de la joven: que el doctor sea bueno.

En este punto, puede parecer que lo que está en el trasfondo del problema aquí planteado es la cuestión de la *suerte epistémica* relativa a las consideraciones accesibles al sujeto en el tiempo *t*. De este modo, la intuición de que lo que *realmente* justifica una decisión es el resultado exitoso vendría motivada por el hecho de que no importa cuán grande sea la evidencia reflexivamente accesible al agente en el momento de la decisión, pues ésta siempre podrá ser retrospectivamente revisada. Duncan Pritchard ha afirmado, en esta dirección, que el problema identificado por Williams coincide con la sospecha escéptica tradicional.²⁷ Pero este problema depende de las demandas exageradas (exageradamente internistas) del escéptico.

Por otro lado, ¿por qué tenemos que aceptar —aun reconociendo que, ciertamente, el éxito o fracaso del proyecto de Gauguin sólo puede conocerse retrospectivamente— que la justificación racional misma es “esencialmente retrospectiva”? Es decir, ¿por qué identificar justificación y resultado? De hecho, resulta *prima facie* razonable afirmar que lo que la justificación racional demanda del sujeto es sólo que esté justificado en su decisión *en el tiempo t en que toma su decisión*. Esto es, A estará justificado en decidir *X* si sigue una ruta deliberativa válida, la cual incluye una correcta apreciación de los hechos implicados, del conocimiento relevante, así como un despliegue correcto del razonamiento, en relación, todo ello, al momento en que se toma la decisión. Así, y aplicado a los casos anteriores, o bien el agente deliberó correctamente, con lo que estaba justificado en el momento de tomar la decisión (con independencia de lo que

²⁷ Pritchard (2006), p. 18.

pase después), y así no puede ser considerado culpable; o bien cometió un fallo racional y no estaba en realidad epistémicamente justificado. Si el agente cumple con las condiciones de la buena deliberación, parecerá excesivo querer culpar —*epistémicamente*; esto es, considerarlo no justificado— a quien actuó como lo hizo de acuerdo con la evidencia disponible y el razonamiento correcto, por mucho que ulteriores evidencias le hubieran conducido a una visión diferente de lo que debería haber hecho.²⁸

Sin embargo, el razonamiento anterior no aprecia adecuadamente *la fuerza que tiene en nuestras vidas la manera en que resultan las cosas*. Cuando afrontamos elecciones difíciles nos encontramos en una situación muy precaria; aquello que realmente nos preocupa, *la meta misma de la decisión, es tener éxito*.²⁹ Y, más fundamentalmente, lo que la tesis del carácter retrospectivo de la justificación pretende enfatizar es que la suerte (intrínseca) está conectada con las acciones del agente de un modo que sobrepasa su relación con sus propias acciones como deliberador racional *ex ante*. Aquello por lo que Williams se muestra preocupado es:

en qué medida [...] la concepción que el agente toma retrospectivamente de sí mismo puede estar afectada por los resultados y no simplemente dirigida a la manera en que deliberó, o pudo haber deliberado, antes del suceso. Podríamos decir que es bastante natural estar disgustado si las cosas salen mal [...] a causa de las acciones de uno mismo [...] pero esta *autocrítica* se aplica racionalmente sólo en la medida en que uno podría haber evitado el resultado siendo más reflexivo o más cuidadoso de antemano. La reflexión apuntará naturalmente a la cuestión de cuándo es verdadero que uno podría haber evitado el resultado; y esta reflexión puede finalmente conducirnos al escepticismo. Pero aquello en lo que quise incidir precede a esta reflexión. Pone en cuestión [...] el supuesto de dividir de esta manera aquello por lo que nos vemos

²⁸ Para esta posición, véase Rosebury (1995), Latus (2001) y Pritchard (2006).

²⁹ Cf. Moore (1997), p. 232-3.

concernidos. El presupuesto puede expresarse así: *como agentes, buscamos ser racionales; en la medida en que somos racionales, nos vemos concernidos por nuestra agencia y sus resultados en tanto que pueden ser moldeados por nuestro pensamiento racional; en la medida en que los resultados de nuestra agencia no podrían verse afectados por una mayor racionalidad, deberíamos verlos como resultados de la agencia de otro o como un suceso natural.*

Esta idea me parece crucialmente equivocada, y con los ejemplos que di en *Moral Luck* [...] traté de incidir en esta cuestión de que [...] *mi implicación en mi acción y sus resultados trasciende la relación que tengo con ella en tanto que deliberador racional ex ante.*³⁰

El propósito de distinguir entre dos aspectos diversos de la justificación, a saber, la justificación en el momento en que se toma una decisión (o justificación *ex ante*) y la justificación retrospectiva o circunscrita a las consecuencias (o justificación *ex post*), choca con la importancia vital que, en tanto que agentes constituidos por nuestras mismas acciones y sus consecuencias, otorgamos a los cambios que nosotros mismos producimos en el mundo. Tratar de establecer este tipo de distinción *técnica* resulta gratuito, en tanto que inerte ante la consideración anterior. (Por otro lado, una noción de justificación exclusivamente *ex post* podría carecer de sentido.)

Williams apela, por otro lado, a una cierta actitud que, según cree, mostramos en relación a decisiones como la tomada por Gauguin. A saber, mostramos nuestra *gratitud* ante Gauguin por su empeño en llevar adelante su proyecto, en perseguir su vocación. Y la razón de que mostremos esta gratitud para con su proyecto es, según Williams, que *acabó teniendo éxito*. Pero, ¿qué significa esto? ¿Qué quiere decir Williams con que mostramos nuestra gratitud hacia Gauguin, y que esta gratitud depen-

³⁰ Williams (1993), p. 245; subrayado añadido. Craig Taylor (manuscrito) incide en este punto, contra Pritchard.

de de su triunfo artístico? Parece que la idea es que le estamos agradecidos por haber producido las obras de arte que produjo, que sin duda hacen nuestro mundo mejor; lo que, no hay duda, depende del éxito de su proyecto.³¹ Para Williams, la base, el alcance y el significado de esta gratitud apuntan a las limitaciones de la idea de querer atender sólo a aquello que el agente controla, en el momento de tomar la decisión. Sin embargo, parece difícil defender que esta gratitud tenga algo que ver con la justificación de la decisión misma; a no ser que esta muestra de gratitud comporte el reconocimiento de todas aquellas decisiones y acciones que le condujeron a la realización de las obras de arte por las que estamos agradecidos. Esto último es particularmente difícil de aceptar. Imaginemos que para poder pagar su pasaje a Tahití, Gauguin aceptara un soborno para encubrir, pongamos por caso, un asesinato. ¿Implica nuestro agradecimiento por su obra pictórica que le reconozcamos ese hecho (si, supongamos, es condición necesaria de su éxito futuro)? Creo que no.³² Por otro lado, la

³¹ Como él mismo ha reconocido, Williams se aprovecha en este ejemplo de una cierta concepción “romántica” del artista (que el artista se debe a su arte por encima de todo), que podría llevarnos a malinterpretar el caso. Véase Williams (1993), p. 255. Sería inadecuado pensar que de lo que se trata es de que *algunas personas* estarían exentas de cumplir ciertas obligaciones morales, dada su naturaleza extraordinaria; esto es, porque poseen alguna o algunas características especialmente sobresalientes, el cultivo de las cuales reportaría un gran beneficio a la humanidad —como en la teoría defendida por Raskólnikov en *Crimen y castigo*, libro II. De hecho, la apelación a la gratitud que supuestamente sentimos hacia Gauguin parece reforzar esta interpretación. No obstante, entender la cuestión de este modo haría mucho menos poderosa la conclusión. Se trata más bien de que en ocasiones el valor moral puede verse superado por un valor no moral, y ello, en principio, para cualquier tipo de persona.

³² Un problema más general al que remite esta dificultad puede enunciarse en términos que lo que se ha dado en llamar la *paradoja de la disculpa* (aunque mejor sería llamarla la *paradoja del pesar*): creemos que debemos disculparnos por los agravios que nuestros antecesores infligieron a ciertos colectivos, indígenas, negros, judíos, etc., pues lamentamos que cometieran las injusticias que cometieron, lo que implica que desearíamos que no lo hubieran actuado como actuaron, pero entonces la historia del mundo habría sido diferente y puede que nosotros mismos no hubiéramos nacido. Sea ese el caso, mi obligación de pedir disculpas por, pongamos por caso, la expulsión de España de los judíos, podría comportar que yo no existiera. Pues bien, no creo incompatible pensar, a la vez, que debo pedir disculpas por ello (por la razón que sea) y seguir siendo feliz con mi

decisión de Gauguin de perseguir su proyecto podría igualmente haber tenido éxito aun sin estar fundada en una deliberación (*ex ante*) adecuada, sin tener ningún indicio razonable para ello. En este caso, podríamos seguir mostrando nuestra gratitud por su arte, pero parece que no podríamos decir que su proyecto estaba justificado —Williams no dice, ni debería decir, que la deliberación sea superflua para la justificación, sino que la justificación depende “esencialmente” del resultado. Por ello, creo que si la gratitud puede jugar aquí algún papel, no es en el sentido de apoyar el carácter retrospectivo de la justificación, sino en tanto que supone el reconocimiento (desde una cierta perspectiva) de la elección de su vocación sobre la dedicación a su familia —esto es, en conexión con la cuestión que tratamos en el apartado anterior.

La conclusión de que la justificación es esencialmente retrospectiva no ha quedado todavía suficientemente establecida. Para ello resultará fundamental la noción del pesar-del-agente, que abordaré en la sección 7.3. Pero antes, hay algo más que quiero decir acerca de la decisión de Gauguin.

7.2.3. El lugar de la suerte moral: ¿qué pasó con la familia Gauguin?

Si aceptamos lo dicho hasta ahora, esto es, que alguien puede, en ocasiones, escoger justificadamente ser fiel a un valor no-moral antes que a un valor moral, y que la justificación incluye una instancia retrospectiva, entonces Gauguin puede estar justificado a marcharse a los Mares de Sur, donde piensa que podrá desarrollar sus dotes artísticas, abandonando

existencia, aunque lo primero implique que el mundo debería haber sido un mundo en el que yo no existiera. Por supuesto, con ello no pretendo resolver la paradoja. Compárese con las cuestiones que tratamos en el capítulo 4. Véase J. Thompson (2000), para un tratamiento elaborado de esta cuestión.

a su suerte a su mujer e hijos, pero sólo si su proyecto no fracasa. (Recordemos que esto último todavía ha de ser defendido.) Pues bien, dado este escenario y el interés por mostrar que los casos de suerte moral son posibles, resulta extraña la desatención a un punto central de la cuestión, más cuando es algo que significativamente puede dar una fuerza mayor al argumento; a saber, el destino que aguarde a su familia tras el abandono. De hecho, sobre la familia sabemos bien poco. Williams sólo nos dice que (su) Gauguin siente algún tipo de responsabilidad hacia su mujer e hijo y que es razonablemente feliz viviendo con ellos, pero aun así se marcha y los abandona. Parece que se muestra preocupado por su bienestar y en su deliberación ha tomado seriamente en consideración su obligación para con ellos. Aunque la historia podría también interpretarse como muestra de la falta de arraigo real en Gauguin de su obligación moral para con su familia, dado que finalmente les da la espalda para perseguir su vocación.³³

Una manera sencilla de sacar más jugo al caso de Gauguin, con la meta de mostrar el papel de la suerte en la asignación de responsabilidad moral —más en consonancia con el modo en que hemos venido abordando previamente la cuestión de la suerte moral—, consistiría en subrayar la importancia del destino de la familia Gauguin para el juicio moral que finalmente Paul merezca. Propongo, pues, que nos fijemos en este otro extremo del dilema, y nos preguntemos: ¿qué papel juegan *retrospectivamente* en todo este embrollo los familiares abandonados por Gauguin para irse al Pacífico? Es obvio que Gauguin no podía estar seguro respecto al éxito o fracaso de su empresa, al margen de su plausibilidad actual;

³³ Williams rechaza explícitamente que se trate del caso de un amoral. Gauguin se siente realmente preocupado por el bienestar de su familia, pero el peso de su vocación es superior. Véase Williams (1981), p. 22-3. Véase también Mulhal (2007) sobre este punto.

como tampoco podía estarlo del destino de su familia, al margen de las previsiones actuales más o menos plausibles. Ambas son cuestiones cuya resolución sólo le serán accesibles retrospectivamente. Por su parte, Williams ponía el énfasis en la justificación o no de tomar el camino de Tahití, esto es, ¿estaba Gauguin justificado a pensar que escogiendo esta opción incrementaría sus opciones para convertirse en un gran pintor? O, ¿con qué base podremos decir que Gauguin estaba justificado a optar por esta alternativa? Mi atención se sitúa ahora, más bien, en si Gauguin estaba justificado a pensar que su familia seguiría viviendo de una manera suficientemente digna —tanto económica como emocionalmente— tras su abandono como para poder optar justificadamente por marcharse.

Para apreciar adecuadamente el papel que juega este extremo de la decisión conflictiva hemos de caracterizar las situaciones futuras posibles de la familia de Gauguin tras el abandono. Es posible que tras ser abandonados sus condiciones de vida se resientan; puede incluso que entren en unas condiciones de vida muy precarias; o, por el contrario, que sigan viviendo como hasta la fecha. También hay que tener en cuenta la situación emocional futura: puede ser que su esposa no resista su ausencia, o, en el otro extremo, que se sienta infinitamente aliviada, porque no le soportaba. También los hijos pueden sentirse mejor o peor, o su formación puede verse mermada por la falta de su padre, etc., etc. Simplificando la cuestión, propongo caracterizar en conjunto las condiciones posibles de vida futura de la familia (también en conjunto, sin distinguir entre mujer e hijos), en términos de la noción de *vida digna de ser vivida*,³⁴ con valor positivo o negativo. Así, la familia, tras el abandono, o tendrá una vida digna de ser vivida o tendrá una vida indigna de ser vivida. Conjuntemos ahora los dos posibles resultados del proyecto de Gauguin (éxito o fracaso).

³⁴ De esta noción ha hecho un uso amplio Derek Parfit, en Parfit (1986).

so) con los dos posibles destinos de su familia (vida digna o vida indigna). El resultado son las cuatro posibilidades básicas siguientes:

- a) éxito del proyecto y vida digna de la familia,
- b) éxito del proyecto y vida indigna de la familia,
- c) fracaso del proyecto y vida digna de la familia,
- d) fracaso del proyecto y vida indigna de la familia.

Por supuesto, también en relación al abandono tendríamos que aplicar la distinción entre suerte intrínseca y suerte extrínseca. Así, y centrándonos en el aspecto económico, podría ser que entraran en la miseria porque les falta el apoyo de Paul, un factor de suerte intrínseca; o por padecer un robo importante, fruto de la suerte extrínseca. Por otro lado, su esposa podría, por ejemplo, haberse suicidado a causa del impacto emocional de ser abandonada por su esposo. Este caso sería semejante al de Ana Karenina, si su amante Vronski se hubiese suicidado. En todo caso, la suerte intrínseca con respecto al futuro de su familia, aunque interna al proyecto, estará parcialmente fuera del agente.

Con el movimiento de considerar el destino de la familia, en tanto que resultado de la misma decisión de Gauguin, la historia se convierte en un caso de suerte moral directa, mucho más incisivo. Con ello, el foco de atención recae sobre el elemento que nos aparece como el más significativo para la evaluación moral de Gauguin, más que el éxito o fracaso de su proyecto: el bienestar futuro de la familia.

Así, el significado de la decisión de Gauguin dependerá no sólo de su éxito o fracaso artístico, sino también del destino de su familia. Si tiene éxito y su familia continúa viviendo una vida digna —la posibilidad (a)— habrá poco de lo que lamentarse. En el otro extremo, si fracasa y encima su familia no vive una vida digna —la posibilidad (d)—, no habrá nada

que atenúe su inmoralidad. Si falla, el hecho de que su familia no caiga en la penuria —posibilidad (c)— podrá ser un atenuante. La posibilidad (b) es el caso más claro de *mala* suerte moral resultante: el abandono de la familia puede ser compensado por el éxito del proyecto, sólo si la familia no cae en desgracia a causa de este abandono. Ciertamente, hay algo significativo en su vida (y no cancelado) —a saber: el triunfo de su proyecto— que poner en contra de su “inmoralidad”, pero es difícil que sus mismos familiares reconozcan este hecho.

Esta última consideración conecta con un elemento central del planteamiento de Williams: la repercusión del resultado para *la evaluación reflexiva por parte del agente de sus propias acciones*. En el caso de la decisión de Gauguin, la justificación tiene que ver con si éste puede justificársela ante sí mismo, o si puede justificarla ante otras personas. Si tiene éxito en su empresa, Gauguin podrá justificar su decisión ante sí mismo, y ante aquellos que le reconozcan su decisión, pero puede que nunca lo consiga ante su familia. Como he dicho, parece que no lo conseguirá, en particular, si ésta cae en desgracia tras su abandono.³⁵

En mi evaluación del argumento de Williams he distinguido hasta el momento tres aspectos principales, que he defendido, con matices, de posibles ataques. La primera cuestión era la de que en la decisión de Gauguin se da un conflicto entre un valor moral y un valor no-moral. El caso de Gauguin, correctamente interpretado, parece un ejemplo de una situación en la que es razonable no decidirse por la alternativa moralmente más correcta, dado el sacrificio personal que supondría.³⁶ En segundo lu-

³⁵ En la sección siguiente se retoma la cuestión de las asimetrías entre la autoevaluación y la evaluación de otras personas.

³⁶ Quien dude de esto, que piense en el caso de Rembrandt, propuesto más arriba.

gar, respecto al carácter retrospectivo de la justificación, destaqué la idea de que la meta misma de la deliberación es el éxito, por lo que no se puede rechazar sin más la importancia del resultado para la justificación de la decisión. En todo caso, debe aún defenderse el papel del pesar-del-agente para establecer este punto. Finalmente, propuse una presentación alternativa del caso de Gauguin que, a mi modo de ver, incide más adecuadamente en el punto moral del conflicto —el bienestar futuro de la familia.

En la sección siguiente, me ocuparé directamente del papel de sentimientos y actitudes reactivas —elementos de carácter retrospectivo— en la evaluación del agente. Empezaré con el llamado pesar-del-agente.

7.3. Juicios, actitudes y sentimientos

Como destaqué en la primera sección, hacia el conductor que atropella a una persona experimentamos una reacción adversa mucho mayor que la que experimentamos hacia el que no lo hace, aunque este hecho esté más allá de su control y ambos hayan sido igualmente temerarios o imprudentes. Lo mismo sucede en el caso de la mujer que acaba matando al conductor que atropella a su hija, o a la que se le ahoga su hijo en la bañera. Sin embargo, el defensor del principio de control rechaza que la distinta actitud reactiva con que respondemos a uno u otro resultado sea parte integrante del juicio moral, o deba ser considerada en el juicio moral reflexivo. El hecho de que se dé uno u otro resultado, por mucho que marque una diferencia en la actitud emocional o reactiva que nos provocan, no se justificaría en el grado de responsabilidad o culpabilidad moral, sino en alguna otra cosa; y, en todo caso, la diferencia en la actitud emo-

cional es un fenómeno que no debería marcar una distinción en el merecimiento del agente.³⁷

Sin embargo, en casos como éstos, nuestra tendencia a reaccionar de manera distinta no puede ser considerada simplemente como un mero impulso vacío de todo contenido cognitivo. No me parece correcto despachar las diferentes actitudes reactivas de esta manera. En lo que sigue, trataré de dilucidar la cuestión de cuál es el significado para la evaluación moral de las diferentes actitudes reactivas que experimentamos hacia resultados distintos. Empezaré con el fenómeno conocido como el *pesar-del-agente*.

7.3.1. El pesar-del-agente, un sentimiento moral

Recordemos que Williams introdujo la cuestión del sentimiento o actitud que se experimenta ante el daño cometido por uno mismo mediante la noción del “pesar-del-agente”.³⁸ Se trata de un sentimiento particular que sólo experimenta un agente en relación a sí mismo, en tanto que causalmente conectado con el daño producido a otra persona. El “pensamiento constitutivo” de la lamentación, en general, es que “sería mejor que las cosas hubiesen sido de otro modo”,³⁹ en contraposición a como de hecho fueron. En particular, el pesar-del-agente es un tipo de lamentación o pesar que una persona siente con respecto a sus propias acciones, en tanto que *autor* o *partícipe*. La conciencia relevante de haber causado un daño

³⁷ Véase Zimmerman (2002), p. 562. Para otros argumentos en contra del significado moral de este sentimiento, ver Jensen (1984), Richards (1986), Thomson (1989) y Rosebury (1995).

³⁸ Williams (1981), pp. 27-30.

³⁹ Williams (1981), p. 27. En lo que sigue hablaré más bien de pensamiento *asociado*, dejando de lado el problema de la conexión entre pensamientos y emociones, esto es, la cuestión de si los pensamientos son constitutivos de, causan o son causados por las emociones. Para problemas con la idea del pensamiento constitutivo de una emoción, véase D’Arms y Jacobson (2003).

es básicamente la de algo que ha sucedido como consecuencia de mis propios actos. El agente A se lamenta de haber hecho algo, que causó un estado de cosas S (cree que ha contribuido a la ocurrencia de S) y, característicamente, juzga que S es dañino, malo o indeseable. Por otro lado, no es preciso que el agente sea enteramente responsable de la ocurrencia de S, puede también ser una causa intermedia; además, el sentimiento debe estar adecuadamente asociado con la historia causal que lo conecta a su verdadero objeto intencional, para que cuente como pesar.⁴⁰

A. El pesar-del-agente y sus asimetrías

El ejemplo típico es el del camionero que atropella a un niño. Al conocer el suceso, el conductor sentirá una clase de pesar por la muerte del niño que nadie más sentirá, pues, al fin y al cabo, fue él mismo quien causó la muerte del niño. Afirma Williams:

sentimos pena por el conductor, pero este sentimiento coexiste con, y de hecho presupone, que hay algo especial respecto a su relación con este suceso, algo que no puede ser eliminado meramente por la consideración de que no fue culpa suya.⁴¹

En particular, la acción del camionero no es una acción *voluntaria*, ni fue antecedida por falta alguna, pero de todas maneras siente y *debe sentir* el pesar-del-agente.

Este es un caso en el que vemos claramente cómo las perspectivas de primera y tercera persona nos dan acceso a hechos básicos distintos e irreductibles. Por un lado, el agente implicado se lamenta (y debe lamentarse) del daño que ha causado. Por otro, los espectadores pueden tratar de

⁴⁰ Véase Rorty (1980), para interesantes consideraciones sobre la noción de pesar-del-agente.

⁴¹ Williams (1981), p. 43.

aminorar el lamento del agente, de consolarle, insistiendo en que no fue su culpa. Pero éste no puede simplemente adoptar la perspectiva de la tercera persona y decirse a sí mismo que no tiene por qué lamentarse, dado que no cometió falta alguna. De hecho, consideraríamos como negativo, e incluso inmoral, que el camionero no reaccionara con una profunda tristeza y pesar en relación al producto de su propia acción. Lo racional para él es pensar que hubo cosas que pudo hacer que habrían evitado la muerte del niño, como salir con más cuidado todavía, conducir más hacia el centro de la calzada, incluso haber perdido unos minutos más mientras repostaba o, simple y llanamente, haber visto al niño. Será racional para él preguntarse si podría haber hecho algo más por evitar la tragedia.

Estamos ante un caso en el que la evaluación, que parece sencilla aplicada a los demás, se torna inestable e incluso objetable cuando es dirigida a uno mismo. De hecho, el camionero tampoco debe desatender el hecho cierto de que no cometió ninguna falta del tipo de las que le hubieran hecho culpable por el atropello, como no mirar al salir, circular por el arcén, no tener los frenos correctamente reglados, etc. Sin duda, esto también forma parte de la evidencia total que debe considerar, y no tenerlo en cuenta es desatender injustificadamente a una parte de los hechos. No se trata de caer en el *victimismo*, especialmente desagradable ante la mirada de la víctima real. Pero el espectador *ve* los hechos de otro modo. Su mirada no está comprometida, como lo está la del participante, tanto debido a que no está directamente implicado en la producción del daño, como en tanto que sus pensamientos acerca de la cuestión no forman parte de los hechos mismos, como sí lo hacen los del agente. La perspectiva de la primera persona contiene exigencias propias y particulares; en concreto, con respecto a la responsabilidad por la propia vida mental, que no se aplican a la aprehensión de la vida mental de los demás. El agente autor

de los hechos no puede sencillamente concentrarse en una comprensión impersonal de los hechos, propia de la perspectiva de la tercera persona, pues su mismo pensamiento acerca de sí marca una diferencia constitutiva en relación a los hechos morales y psicológicos mismos. En cierto sentido, no puede ni considerar *seriamente* esta perspectiva de los hechos —al menos durante el tiempo en que el sentimiento de pesar-del-agente en cuestión (el *token* de ese sentimiento) resulte apropiado.

Consideremos un conocido caso de Richard Moran, basado en la historia del protagonista de la novela *That Uncertain Feeling* de Kingsley Amis.⁴² Un hombre casado y padre de familia acude una noche a una discoteca con otra mujer compañera de trabajo. (No sé si la historia incluye la práctica efectiva de relaciones sexuales; pero, para evitar una excesiva mojigatería, añadámoselas.) De vuelta a casa, sus pensamientos se ven envueltos por un profundo sentimiento de culpa. Se siente un “libertino”, un mal marido; se desagrada a sí mismo. Pero, a su vez, siente que es un buen tipo por no gustarse a sí mismo a causa de esta conducta. No obstante, vuelve a no gustarse a sí mismo por sentirse un buen tipo por no gustarse a sí mismo. Y, así, sucesivamente. Se da aquí una cadena de giros de 180 grados, que sitúa sus reflexiones en un punto ciertamente inestable y paradójico. Desde la perspectiva externa, no habría ningún problema en evaluarle negativamente por su conducta inicial y, a continuación, reconocer que el rechazo de su propia conducta le hace mejor persona (no tan mal tío). Sin embargo, este pensamiento, aprobado desde la primera persona, se torna contraproducente. Algo parecido le sucedería a nuestro camionero si, tomando seriamente en consideración la evaluación de su ac-

⁴² Moran (2001), pp. 174-5. Para la elaboración de este párrafo y el anterior, me he apoyado en su idea de la irreductibilidad de las perspectivas de primera y tercera persona y su repercusión en la comprensión de ciertos fenómenos psicológico-morales; véase, en particular, el capítulo 5.

ción desde la perspectiva externa, empezara a pensar que no debería sentir como siente, pues no fue su falta, y nada pudo hacer (intencionalmente) para evitar el atropello del niño, y acabara pensando que, en relación a su agencia, nada tiene que lamentar.⁴³ El paralelismo no consiste en que el camionero también entre en un bucle de pensamientos como en el caso anterior, sino en el simple hecho de que aquello que el agente piensa o siente acerca de lo que piensa o siente, *cambia* lo que piensa o siente; es decir, cambia el conjunto total de sus pensamientos y sentimientos. En el caso de la evaluación en primera persona, de la autorreflexión, los pensamientos y sentimientos reflexivos forman también parte del conjunto total, mientras que no lo hacen en el caso de la tercera persona o evaluación de otras personas. Su atención excesiva al hecho de que no fue su falta, supone inmediatamente un abandono o modificación de su sentimiento de pesar-del-agente, que en todo caso arruina el valor de este mismo sentimiento. Tanto en el caso del libertino como en éste último, apoyarse más de la cuenta en la mirada impersonal de los hechos supone una forma de *evasión moral*.

B. ¿Un sentimiento moral?

En todo caso, la cuestión fundamental que se suscita aquí en relación a nuestro tema es la relación de este sentimiento con la moralidad. Es decir, ¿qué tiene que ver este sentimiento con la evaluación de las obligaciones morales del agente? Una réplica tajante consistiría simplemente en afirmar que sentimientos y emociones nada tienen que ver con la morali-

⁴³ Obviamente, el agente no puede controlar de manera voluntaria y directa el sentir pesar u otras emociones, ni el dejar de hacerlo; aunque sí que dispone de ciertas estrategias indirectas, como la modificación de sus hábitos de atención y focalización. El valor moral del lamento se halla primariamente en sus efectos sobre los hábitos y disposiciones del agente.

dad. Como mucho, constituirían epifenómenos, o apoyos motivacionales para el cumplimiento del deber moral; pero sin que exista, en ningún caso, una conexión fundamental entre sentimientos y principios o deberes morales. En esta línea, Brian Rosebury remarca que, dado que nuestra vida emocional está menos estructurada analíticamente y diferenciada que nuestros juicios racionales, el mismo concepto de responsabilidad moral “perdería su coherencia si incluyéramos esta insatisfacción residual [el sentimiento de pesar] entre los sentimientos morales”.⁴⁴ De hecho, continúa el argumento, tanto las consideraciones intuitivas como la utilidad semántica van en contra de llamar a este sentimiento “moral”.

Sin embargo, el pesar-del-agente parece que es un sentimiento moral, al menos, en el sentido mínimo e innegable de que su pensamiento asociado es la evaluación *moral* negativa de la propia acción por parte del agente, que resulta del choque de su acción con un valor moral sostenido por él.⁴⁵ De hecho, el negador de la suerte moral resultante puede reconocer esto sin problemas, es decir, no tiene porqué ser un racionalista moral *à la Kant*. Puede reconocer la naturaleza moral de este tipo de sentimientos, pero negar que ello signifique que tengan algo que ver con la determinación de la responsabilidad moral.

Susan Wolf sostiene que hay “una virtud sin nombre”, consistente en *asumir la responsabilidad* por las propias acciones y sus consecuencias, especialmente las dañinas, aunque uno no sea moralmente responsable de ellas.⁴⁶ Esta virtud sensibiliza a quien la posee y la cultiva con la conciencia de la conexión causal entre cómo resultan las cosas y los propios actos, así como con su *asunción de responsabilidad*, y se manifiesta,

⁴⁴ Rosebury (1995); p. 514.

⁴⁵ Véase Tannenbaum (2007).

⁴⁶ Wolf (2001), p. 13.

ciertamente, en un sentimiento especial de pesar, como es el pesar-del-agente. Al ser un sentimiento doloroso que el agente tratará de evitar, el lamento propiamente centrado incentiva la conducta responsable del agente, sensibilizándole a las medidas preventivas y de remedio.⁴⁷ En este sentido, cabe reconocer la propiedad de que el agente sienta este tipo especial de pesar por la muerte del niño, aun cuando aquella acción suya que produjo este resultado no fue voluntaria. Es más, el agente tiene la obligación *moral* de experimentar este sentimiento, si ha de contar entre las personas moralmente *sanas*. Pero reconocer este hecho no impide negar que esto repercuta en, o sea una guía de, la culpabilidad del agente. Wolf remarca que esta virtud sin nombre se asemeja a la virtud de la generosidad, una virtud que sobrepasa el requisito estricto de justicia.⁴⁸

Así, esta última posición acepta que este tipo de sentimiento es moral en un sentido amplio, pero niega que sea una guía, al menos fiable, para determinar el merecimiento del agente. De manera un tanto irónica, puede utilizarse el mismo caso de Williams para mostrar este hecho. El pesar-del-agente puede ser experimentado ante las consecuencias trágicas tanto de una acción que es fruto de una falta del agente como de una que no lo es. De hecho, el mismo ejemplo paradigmático del camionero es un caso en el que el atropello del niño está más allá de la comisión de ninguna falta moral por su parte. A esto cabe responder de dos maneras.

C. La conclusión de Williams

En primer lugar, Williams no presenta el fenómeno del pesar-del-agente como una prueba o indicio de la culpabilidad del agente. De hecho, incluso destaca que el agente debe experimentar este sentimiento con in-

⁴⁷ Ver Rorty (1980).

⁴⁸ Wolf (2001), p. 14. Véase también Latus (2001), § 1.

dependencia de su culpabilidad. Lo que pretende mostrar con la apelación a este fenómeno es la fundamental importancia que tiene para nosotros la perspectiva retrospectiva de nuestras decisiones y acciones. La idea sería: ¿mira si importa la perspectiva retrospectiva, que hasta cuando no hay voluntariedad ni falta por nuestra parte nos sentimos fuertemente concernidos por las consecuencias de nuestras acciones y decisiones! Esto es, la perspectiva retrospectiva, tanto en relación a cómo las cosas resultan, como en relación a cómo el agente reacciona ante estos resultados, es fundamental en nuestra concepción de nosotros mismos y de los demás.⁴⁹

Ello explicará, además, un aspecto de la disensión de Nagel acerca de la naturaleza moral de este tipo de sentimiento o actitud. Para éste, era necesario que el camionero hubiera cometido una negligencia para hablar de un sentimiento moral, y así de suerte moral, esto es, de suerte en relación a la atribución de responsabilidad moral. Si el camionero fuera culpable de la más mínima negligencia y ésta contribuyera a la muerte del niño, entonces el camionero no sólo debería sentirse enormemente mal, sino también culpable por la muerte.⁵⁰ (Explotaré esta idea en la subsección siguiente.) Pero el papel que el pesar-del-agente juega en el argumento de Williams es principalmente mostrar la importancia en nuestra concepción de la moralidad, en general, de la perspectiva retrospectiva.

Por otro lado, recordemos que el argumento de Williams dependía de estas dos afirmaciones: 1) que la justificación racional puede chocar y llegar a imponerse a la justificación moral, y 2) que la misma justificación

⁴⁹ Recuérdese que ya Aristóteles insistió en la importancia, para que una acción contara como involuntaria, de que el agente se sienta apenado por su realización y consecuencias, de modo que le fuese excusada o mitigada su responsabilidad (dado que Aristóteles asociaba responsabilidad con acción voluntaria). Véase Aristóteles, EN III, 1.

⁵⁰ Nagel (1979), p. 29. “Mi desacuerdo con Williams es que su teoría no explica por qué tales actitudes retrospectivas pueden llamarse morales.” (p. 28.) También Rosebury (1995) y Latus (2001) insisten en esta idea.

racional resulta ser vulnerable a la suerte. Esta vulnerabilidad de la justificación racional a la suerte se debe a la incontrolabilidad de los resultados y la importancia para nuestras vidas de los resultados se ve refrendada por nuestras reacciones sentimentales. En realidad, la postura final de Williams, ante un posible contraataque, es más bien una disyunción: o bien el valor moral puede quedar sometido a la suerte, o bien no es lo que últimamente nos importa, pues no siempre es el valor supremo. En todo caso, el valor moral no podrá ya ser el tipo de cosa que pensamos (o que la concepción kantiana afirma) que es, a saber, supremo e inmune a la suerte —que es lo único que garantizaría su papel como “consuelo de un sentido de la injusticia del mundo”.⁵¹

En otras palabras, puede que en último término la suerte no afecte al núcleo más esencial de la moralidad, a la esfera del merecimiento (de la obligación y la culpabilidad más estrechas), pero esto sólo puede lograrse a fuer de estrechar tan sumamente el alcance de esta esfera que poco quedará en ella de lo que realmente nos importa. En tal caso, casi todo aquello que da significado en nuestras vidas a la moralidad, que la conecta con la cuestión de cómo hemos de vivir, quedará excluido de ésta. Este argumento mantiene cierto paralelo con otro que el mismo Williams presenta en *Verdad y veracidad* contra cierta tradición de pensamiento que ha insistido hasta la saciedad en la idea de que mentir es moralmente incorrecto siempre y sin excepción. La defensa de esta idea choca con ciertas situaciones en las que parece que uno debería mentir para evitar un mal mayor, como, por ejemplo, cuando un asesino te pregunta por el paradero de su pretendida víctima y tú lo conoces pero no quieres contribuir a su asesinato, o cuando decirle la verdad sobre su hija a una anciana podría causarle un gran daño emocional. Para superar estos posibles contraejem-

⁵¹ Williams (1981), p. 21.

plos sin renunciar a la idea de que la mentira es siempre condenable, se idearon circunloquios de diverso tipo, como la distinción entre mentir y llevar a engaño, la posibilidad de una reserva mental, etc.; movimiento con el cual ciertamente se salva la objeción, pero a razón de que la posición defendida pierda toda su fuerza intuitiva. De esto modo, las razones iniciales de que no se debe mentir quedan completamente ausentes del concepto estrecho final de mentira vindicado —que deja fuera cosas tan intuitivamente igual de rechazables como el llevar a engaño, esconder la verdad, etc. Del mismo modo, el negador de la suerte moral puede ganar el debate defendiendo que el núcleo esencial de la moralidad permanece inmune a la suerte, pero lo hace estrechando de tal manera lo que entiende por moralidad que deja fuera buena parte de lo que la hace que la moralidad sea significativa e importante. El resultado es, digámoslo así, una especie de *fetichismo de lo moral*.

De hecho, ante la insistencia, que vimos más arriba, en negar el carácter moral de ciertos sentimientos, Williams responde: “¿cuál es la razón para insistir en que cierta reacción o actitud o juicio es o no *moral*? ¿Qué se supone que esta categoría nos da?” Si previamente no tenemos bien definida y acotada la categoría de lo *moral* y, a su vez, una teoría sobre la importancia singular de lo moral en este sentido restringido, “no se conseguirá absolutamente nada”.⁵² No es extraño que Williams proponga que nos deshagamos de esta *concepción estrecha de la moralidad*, que la identifica con la mera esfera de la obligación moral, a favor de concepción más amplia que incluya todo *lo ético*.⁵³

⁵² Williams (1993), p. 245.

⁵³ Williams (1985), pp. 6 y 174-196. Véase también sobre esta distinción, Mackie (1977), pp. 106-7.

Considero que esta conclusión es sólida y *filosóficamente* relevante, es decir, con importantes consecuencias a gran escala. No obstante, aun con todo, el negador de la suerte moral puede desprenderse de toda aspiración a ofrecer una imagen amplia de la moralidad libre de suerte, y centrarse exclusivamente en la esfera más estrecha de atribución de culpabilidad (responsabilidad moral) y negar que la suerte (resultante) marque aquí ninguna diferencia. En cierta manera, con este movimiento se renuncia a una posición substantivamente significativa, pero hay que reconocer que es un movimiento argumentativamente legítimo. En la siguiente sección, trataré de defender, contra esta última salida, que la diferente actitud o sentimiento, que en los casos de culpabilidad experimentamos en relación a uno u otro de los agentes contrastados en los ejemplos, en particular, de suerte resultante, puede ser una guía fiable, dadas ciertas condiciones, para la mayor o menor censura de uno y otro. En el trasfondo de esta afirmación se encuentra la idea de que ciertas actitudes, sentimientos o emociones nos ayudan a descubrir valores.

7.3.2. Actitudes reactivas y juicio moral

La cuestión a discernir es si la experiencia de algún tipo de sentimiento puede ser una guía para la atribución de responsabilidad moral. En particular, los sentimientos de tipo moral que pueden desarrollar este papel son aquellos que sólo son apropiados cuando se dirigen a una acción o consecuencia que es producto de una acción voluntaria o fruto de una falta moral del agente. Más concretamente, los sentimientos sobre los que hemos de centrar nuestra atención son, desde la perspectiva del agente, el sentimiento de culpa y otros de su especie, así como el resentimiento y la indignación, propios de la perspectiva del espectador.

En cuanto a los sentimientos de primera persona, el sentimiento de culpa y el remordimiento contrastan con el pesar-del-agente en tanto que dependen de acciones voluntarias o, más bien, que son fruto de algún tipo de falta moral. Además, el pesar se distingue del sentimiento de culpa y el remordimiento, así como de la vergüenza o la autorrecreminación, no por el tipo de sensación sufrida o la intensidad de la emoción, sino en virtud de sus pensamientos asociados —aunque a menudo éstos resulten difíciles de precisar. Así, uno puede sentirse moderadamente culpable ante una falta pequeña, mientras que puede sentir un intenso pesar-del-agente, como en el caso del camionero. Por otro lado, las actitudes típicas de la tercera persona ante la comisión de un daño son el resentimiento y la indignación. Éstos pueden dirigirse también a acciones involuntarias o que no son fruto de ninguna falta por parte del agente; especialmente en el caso del resentimiento: podemos experimentar resentimiento ante el camionero que atropelló al niño. Pero cabe distinguir este resentimiento del que experimentamos ante el atropello causado por un conductor imprudente o, más aun, ante la acción criminal de un asesino. Puede decirse que, aunque en el primer caso el sentimiento sea típico y comprensible, no es racional; mientras que en el segundo caso es también racional. (Más sobre esta caracterización a continuación.)

La dialéctica del debate es como sigue. El defensor de la suerte moral resultante afirma que la diferencia en las actitudes reactivas y sentimientos que las diversas consecuencias de las acciones de los diferentes agentes provocan en nosotros son una buena guía, en general, para la atribución de una responsabilidad moral mayor y, con ello, de una mayor censura. Esto es, cuando vemos o nos informan de que alguien ha dañado injustificadamente a otra persona reaccionamos, en general, con indignación y resentimiento hacia aquél. Además, parece que reaccionamos más

adversamente o mostramos mayor indignación y resentimiento hacia el conductor que ha atropellado a una persona que hacia el que no ha atropellado a nadie. Por supuesto, ante la temeridad misma de conducir sin control reaccionamos con animadversión hacia el conductor, pero típicamente esta animadversión aumenta cuando lleva a un resultado fatal. Lo mismo sucede en el caso de la mujer que acababa matando al conductor que atropelló a su hija. De hecho, ella misma tendrá algo (mucho) por lo que sentirse culpable si se da este resultado, como lo tendrá la madre que no impide que su hijo se ahogue en la bañera. Esperamos también que el conductor desafortunado se culpe más a sí mismo que el afortunado, porque de hecho sólo él causó efectivamente daño en otra persona. Ante esto, el negador puede contraatacar con ejemplos en los que los sentimientos son claramente una mala guía. Por ejemplo, Richards rechaza la relevancia de los sentimientos de indignación y resentimiento en virtud de que éstos pueden dirigirse también a objetos inanimados, como la navaja de un violador, aunque es claramente absurdo pensar que objetos de este tipo sean moralmente responsables y merezcan tal reacción.⁵⁴ Sin embargo, la idea no es que cualquier actitud reactiva o sentimiento sea en general una buena guía para determinar la responsabilidad moral, ni que lo sean en exclusiva. Sino, más bien, que no sólo la mera razón está autorizada a ello, por un lado; y, por otro, que para ello tenemos que estar ante actitudes o sentimientos que cumplan ciertas condiciones de racionalidad.

No cabe duda de que los sentimientos juegan en ocasiones un papel epistémico en el desvelamiento del valor. Por ejemplo, uno puede descubrir que es hincha de un determinado equipo de fútbol cuando se descubre a sí mismo triste por la derrota de ese equipo, o contento de su victoria. Las emociones pueden revelarnos cosas por las que estamos con-

⁵⁴ Richards (1986), p. 179.

cernidos, así como confieren pesos determinados a la gran diversidad de razones posibles —suponen un anclaje particular de las razones en cada individuo. En particular, las emociones morales pueden traer a nuestra atención ciertos hechos acerca de la situación y necesidades de otras personas. Esto apunta a su papel central en la sensibilización del evaluador a las diferencias en el merecimiento del agente evaluado. Sin embargo, esto no significa que las emociones y sentimientos marquen siempre la diferencia a este respecto, ni que sean siempre fiables. De hecho, no tienen por qué ser consultadas en todos los casos, ni cuando lo son es siempre evidente a qué valor o evaluación apuntan. Pero tampoco nuestras creencias son siempre fiables y esto no es una objeción a su utilidad general. De lo que se trata es de hallar y superar las distorsiones y demás errores en todas partes, tanto en las creencias como en las emociones.

Por otro lado, emociones, actitudes y sentimientos presentan diferencias respecto a su receptividad a razones que nos permiten discriminar racionalmente entre ellos.⁵⁵ En particular, emociones y sentimientos pueden ser típicos, comprensibles y apropiados o racionales. Por ejemplo, el enfadarse cada vez que uno queda atrapado en un atasco de tráfico previsible puede ser típico, pero no es comprensible ni apropiado; o el enfurecimiento de una madre por un daño mínimo inflingido sobre su hijo es típico y comprensible, pero inapropiado; mientras que el sentimiento experimentado por el camionero no es sólo típico y comprensible, sino también apropiado o racional, ya que representa correctamente su acción co-

⁵⁵ Las emociones, como las sensaciones y los estados de ánimo son episodios u ocurrencias; mientras que los sentimientos son disposiciones profundamente enraizadas cuyas manifestaciones son emociones y otros tipos de estados afectivos. Además, los fenómenos afectivos y sentimientos se distinguen de las intenciones y voliciones. Las bases cognitivas de éstas últimas están proposicionalmente estructuradas, orientadas al futuro y son apoyo de acciones posibles. Mientras que las emociones no necesitan tener una base proposicional, a menudo se dirigen al pasado y no requieren una conexión directa con la conducta. Ver Mulligan (1998).

mo moralmente indeseable. Un sentimiento es racional sólo si el contenido del pensamiento asociado a él representa correctamente el mundo. Por ejemplo, el temor es racional sólo si el agente piensa que el objeto temido es peligroso y el objeto es realmente peligroso.⁵⁶ Tener miedo a volar o asustarse al ver un ratón, por típicos y comprensibles que en ciertas personas puedan ser, no son sentimientos racionales, pues sus objetos intencionales no son *realmente* peligrosos.

No dudo de que el adversario de la idea de que los sentimientos pueden ser una guía fiable para la determinación de la responsabilidad moral verá aquí una circularidad: los juicios morales racionales dependen de los sentimientos, pero los sentimientos apropiados son definidos en virtud de su racionalidad. Sin embargo, la defensa de esta idea no presupone necesariamente algo así como un emotivismo ingenuo, según el cual el juicio moral depende de las meras reacciones emocionales, sino más bien que las emociones incluyen y desvelan valores. Del mero hecho de que uno esté enfadado no se sigue que algo malo se le haya hecho; sino sólo que, en relación a cómo uno se toma las cosas, ha sufrido un mal. Se requiere además una justificación, que aquello que uno sufre como un mal sea un mal, como dije. Sin embargo, la posición que defiendo sólo se compromete con la tesis mínima de que tanto razones como sentimientos tienen algo que decir en el descubrimiento de deberes y valores, y que, en ciertas ocasiones, el sentimiento puede tener la última palabra.

Además, comprender cómo y por qué sentimos que algo es indignante o digno de resentimiento parece ser necesario para comprender adecuadamente esos conceptos evaluativos, por lo que la comprensión de al

⁵⁶ Véase Tannenbaum (2007) sobre esto. Una emoción es además racional sólo si tiene el objeto, la duración, la intensidad, etc., apropiados; Aristóteles, EN, libro III, epígrafes 6-8 y libro IV, epígrafe 5. Véase también D'Arms y Jacobson (2000a).

menos algunos conceptos morales y evaluativos requerirá de un tipo particular de comprensión de las emociones que invocan o están relacionadas con esos conceptos. Es más, nuestros sentimientos, nuestras respuestas emocionales, juegan un papel en la misma constitución de ciertos aspectos de nuestras vidas morales. La misma noción de censura parece contener un elemento emocional inextricable. Las emociones de admiración y gratitud o de ira y resentimiento proveen a menudo un fundamento para la elogiabilidad y la censurabilidad, para la misma atribución de responsabilidad moral, que no parece plausible querer erradicar, ni tampoco separar de la misma elogiabilidad o censurabilidad. De hecho, la conocida apelación a las “actitudes reactivas” de Peter Strawson supone que las condiciones bajo las cuales es apropiado juzgar a alguien como responsable no pueden fijarse independientemente de las condiciones bajo las cuales nuestras actitudes reactivas resultan apropiadas.⁵⁷ Emociones reactivas como el resentimiento, la indignación y el sentimiento de culpa constituirían la clase de las emociones esencialmente conectadas con las atribuciones de responsabilidad moral, por medio de nuestras expectativas acerca de la conducta moral de las personas. Pero no necesito comprometerme aquí con una concepción metaética determinada.⁵⁸ Lo único que estoy

⁵⁷ Strawson (2003); véase también Wallace (1994). Ver capítulo 1, sección 2.

⁵⁸ Si tuviera que pronunciarme sobre este tema, no dudaría en reconocer que mis simpatías se dirigen al “neosentimentalismo” y, en particular, a lo que últimamente se ha dado en llamar *sentimentalismo racional*. Para este análisis, un objeto particular posee una propiedad evaluativa Φ sólo si es apropiado tener la respuesta emocional R a Φ . (D’Arms & Jacobson 2000b, p. 729) “Lo que una persona hace es moralmente incorrecto si y sólo si es racional para él sentir culpa por hacerlo, y para los demás sentir resentimiento.” (Gibbard 1990, p. 42; ver también Blackburn 1998.) Por ejemplo, que la acción de un agente sea vergonzosa por, digamos, no respetar un minuto de silencio dedicado a una persona fallecida, es que sea apropiado para un agente apropiadamente situado sentir vergüenza por esa acción. Esta tesis constituye una evaluación de segundo orden de las emociones en términos de su carácter apropiado (*appropriateness*). La naturaleza de los conceptos evaluativos, en este caso “lo vergonzoso”, depende de las respuestas emocionales apropiadas debido a que la clase de cosas que en el mundo merecen una respuesta

defendiendo es que sentimientos morales como la indignación o la culpa pueden desempeñar un papel fundamental en la determinación o descubrimiento del merecimiento y, así, en la asignación de la responsabilidad moral.

Permítaseme apelar al siguiente caso. Raskólnikov, el protagonista de *Crimen y castigo*, sufre una fuerte afección emocional, que le lleva incluso a la enfermedad fisiológica, tras la comisión de su crimen. Raskólnikov creía que su crimen, el asesinato de la anciana usurera, podía estar justificado por ciertas consideraciones de utilidad y justicia social; pero después del crimen se hunde sin remedio: Raskólnikov descubre su culpabilidad por medio de su propia reacción. Asimismo, el Mayor Claude Eatherly, piloto de uno de los aviones participantes en el lanzamiento de la bomba atómica sobre Hiroshima, se halla en una situación semejante. Tras el bombardeo, Eatherly no reaccionó como *se esperaba*. Fue el único de todos los participantes en los dos bombardeos atómicos que se resistió a la tentación de ser homenajeado de forma pública y masiva en calidad de héroe de guerra. El acontecimiento supuso para él un fuerte estremecimiento. Se apartó de la esfera pública e intentó readaptarse a una

de “vergüenza” no puede ser determinada con independencia de la respuesta emocional misma. En general, el sentimentalismo considera que las emociones son aspectos ineliminables en la experiencia del valor, pues el valor depende de nuestra experiencia de él, aunque no sea reducible a esta experiencia.

Véase Ramírez (en preparación) para una defensa cualificada de esta posición y su uso como fundamento de las actitudes reactivas, núcleo de una teoría de la responsabilidad moral *à la Strawson*.

(El *neosentimentalismo* puede hacer frente a importantes objeciones dirigidas tradicionalmente contra el sentimentalismo: como las objeciones de la ausencia de emoción, del desacuerdo y del razonamiento moral. Los sentimientos son integrales al juicio moral, pero uno puede seguir pensando que sentir culpa por cierta acción es apropiado aunque haya perdido la disposición a sentirse culpable por esa acción. Por su parte, el desacuerdo moral puede explicarse como desacuerdo acerca de si es apropiado sentir culpa por realizar determinada acción. Finalmente, cuando razonamos sobre cuestiones morales, razonamos sobre la propiedad de sentir culpa en respuesta a la realización de determinadas acciones.)

vida familiar normal, pero ya nada volvería a ser como antes. Los recuerdos terroríficos comenzaron a atormentarlo: los rostros desencajados de los cuerpos ardientes en el fuego infernal de Hiroshima llenaban sus sueños. Inició una fuerte medicación con todo tipo de píldoras y calmantes. Decidió también mandar cartas de autoacusación y disculpa a Japón, así como enviar dinero para la reconstrucción o el mantenimiento de hospitales y residencias de afectados por la bomba. Intentó también suicidarse. Realizó acciones ilegales para perder el aura pública de héroe. Hasta que fue finalmente internado en un hospital psiquiátrico militar, donde volvería a intentar suicidarse.

No se trata de construir esas mismas reacciones como juicios de merecimiento, sino de incluirlas en el análisis del merecimiento. En este sentido, sentir que estamos justificados a experimentar tales sentimientos a pesar de que reconozcamos que no son merecidos no supone una objeción a la posición defendida. Por supuesto, estos sentimientos de culpa pueden ser exagerados e inapropiados y, en ocasiones, pueden ser manifiestamente desaconsejables. Pero esto puede determinarse mediante el análisis individualizado de cada caso. Cuando no responden a una falta moral cometida por el agente considerado o son desproporcionados deben, obviamente, ser revisados, al menos en tanto que guía para la asignación de responsabilidad. Lo que no cabe duda es que en los casos anteriores, el sentimiento de culpa rastrea claramente el daño cometido. Esto es común a la indignación y el resentimiento experimentado en tercera persona. Estos sentimientos, que sin duda deben también pasar por el tamiz de la reflexión y crítica racionales, apuntan primariamente al daño injustificado (y, con ello, moralmente no permisible) inflingido sobre una o más personas. Y este hecho es fundamental si convenimos en que la razón de ser de la moralidad es el respeto a las personas.

7.3.3. Agencia y daño

En las secciones anteriores, he valorado la función de los sentimientos en el juicio moral, en tres sentidos: i) en tanto que un sentimiento que siente el agente, con independencia de la culpabilidad de su acto (el pesar-del-agente, en concreto); ii) en tanto que sentimiento en primera persona en relación a una acción culpable (el sentimiento de culpa); y iii) como sentimiento experimentado en tercera persona (indignación o resentimiento). Aunque la conexión de cada tipo con la culpabilidad y falta moral del agente, así como su fiabilidad racional, sea distinto, es característico que todos ellos apuntan al daño causado por una acción del agente en cuestión.

Todos estos sentimientos (el pesar, sentimiento de culpa, remordimiento, indignación o resentimiento) conectan la acción de un agente con la causación de daño. Y pueden focalizar más intensamente en uno u otro elemento. De hecho, “A se lamenta por haber producido S” (o “B lamenta la producción de S por A”) es ambiguo entre “A lamenta S” (“B lamenta S”) y “A lamenta ser el causante de S” (“B lamenta que A sea el causante de S”). Esta ambigüedad, sin duda más evidente en el caso del sentimiento dirigido a uno mismo, es reflejo de los dos posibles focos del lamento: el estado de cosas o dolor en la persona dañada (S), o la acción del agente que causó (o contribuyó a causar) el estado de cosas, esto es, la propia agencia.⁵⁹

De este modo, el lamento o arrepentimiento puede focalizarse fuertemente en el mismo agente. Fue él, en último término, y nadie más, quien produjo el daño por el que se lamenta. Es la producción efectiva de daño, en particular si es especialmente significativo o reiterado, y en con-

⁵⁹ Véase Rorty (1980) en relación a esta ambigüedad o dualidad.

traste con la mera posibilidad de haber causado daño, la que marca una distinción en cuanto a la adquisición de una *obligación* de reparación o retribución. Además, la producción efectiva de daño comporta una conversión en la identidad moral del agente mismo. Su estatus e identidad moral depende, por lo menos en parte, de cómo realiza las intenciones, valores y proyectos que emprende. En los casos de acciones involuntarias, el agente ha fallado, con sus propias acciones y sus consecuencias, a su ideal moral. Y en el de las acciones voluntarias, el éxito de los propósitos del agente comporta una diferencia en relación a *en quién se ha convertido*. La conducción temeraria o el intento de homicidio no convierten a sus ejecutores en homicidas o asesinos, como sí que lo hace el homicidio y el asesinato cumplidos.⁶⁰

Pero el exceso de atención del agente a su propio sentimiento de culpa o pesar, a su propia respuesta afectiva, si bien a menudo comprensible, desvirtúa la función *moral* primaria de este tipo de sentimiento: la atención a las acciones y consecuencias dañinas que provocaron esta respuesta y que constituyen su objeto intencional. Piénsese, por ejemplo, en el caso de Fred Vincy, personaje de la novela *Middlemarch* de George Elliot, que ha perdido el dinero que le prestaron los Garth. Fred está realmente apenado, pero focaliza la evaluación del significado moral de lo que hizo en relación a qué revela esto de su carácter, más que en relación a las consecuencias de lo que hizo en los demás. Se da aquí una interferencia entre cuál sea la apreciación ajena de su conducta y el significado de lo que hizo, un tipo de evasión posible gracias a que su pensamiento

⁶⁰ Sobre esta cuestión ya insistió Peter Winch. Véase Winch (1972), cap. 7. Éste afirma que: “Al *hacer* algo malvado uno se *convierte* en alguien malvado [...]. No es que el hombre que falla no tenga nada por qué condenarse a sí mismo o ser condenado por los demás; sino que no tiene que condenarse por aquello por lo que se hubiera condenado si no hubiera tenido éxito.” (pp. 149-50.)

sobre sí mismo adopta la perspectiva imaginaria de otra persona —en particular, la que imagina es la mirada de Mary, su amada e hija de los Garth. Recordemos las consideraciones anteriores acerca de las asimetrías entre las perspectivas de primera y tercera persona. La insistencia de Fred en que Mary le perdona desespera a esta última: “¿Qué importa que yo te perdone?” le espeta Mary. “¿Mejoraría eso en algo que mi madre haya perdido el dinero que tardó en ganar cuatro años con sus lecciones...? ¿Verías esto con suficiente agrado si yo te perdonara?”⁶¹

Si bien, en ocasiones, focalizar la atención en la propia agencia puede resultar muy consolador; vemos que el reproche de Mary trata de hacer ver a Fred que su lamento debería centrarse, en primer lugar, en las consecuencias dañinas de sus acciones. Es la consciencia del daño causado por él lo que debería preocupar a Fred en primer lugar. Desde la perspectiva externa, podemos lamentar tanto el prejuicio que para la madre de Mary supone la pérdida de sus ahorros, como el pesar de Fred por éste prejuicio y por ser su causante. Es la consciencia tanto del daño producido como de ser el causante lo que induce al agente a buscar la reparación de sus actos, en la víctima y en su misma identidad. Se tratará, finalmente, de conseguir un cierto equilibrio entre ambos elementos.

En suma, el tipo de reacciones que he estado discutiendo enfatiza, por un lado, la importancia del daño mismo y, por otro, la relevancia de la contribución del agente a este resultado dañino. Constituyen reacciones afectivas que pueden ser consideradas como una suerte de juicio impreciso —siempre que rastreen hechos reales—, pero que nos descubren aspectos relevantes con respecto a la *verdad moral* del caso. Perderemos una cosa importante si no los incluimos en la caracterización de nuestra con-

⁶¹ George Elliot, *Middelmarch*, libro 3, cap. 25. Citado en Moran (2001), pp. 188ss.

sideración moral de los demás y de nosotros mismos. Su exclusión desvirtuaría fatalmente la comprensión de nuestras prácticas.

Asimismo, no cabe duda de que la historia causal que conecta al agente con el daño producido es fundamental, si bien la conexión misma es parcialmente normativa. Profundizaré en esta cuestión en el capítulo siguiente.

Dónde estamos y adónde vamos

Con la defensa de la realidad del último tipo suerte moral, la consecutiva, finaliza la segunda parte de este trabajo, dedicada a defender la existencia de la suerte moral en cada uno de sus tipos.

Los dos siguientes capítulos, que constituyen la tercera y última parte de esta investigación, están dedicados a analizar, explicitar y defender las principales repercusiones de la existencia de la suerte moral. En particular, el próximo capítulo aborda directamente la significación y consecuencias de la revisión del principio de control. En el último, me ocuparé principalmente de cuestiones metateóricas, o de segundo orden. Ambos son fundamentales para acabar de articular mi posición.

Parte III

8. El rechazo del control y sus consecuencias

Reconsideración del sistema de la responsabilidad

- 8.1. El principio de control
 - 8.1.1. Nociones de control. El problema escéptico
 - 8.1.2. El Principio de Equidad y el ideal de justicia última
- 8.2. Contra el principio de control
 - 8.2.1. Causación y asimetrías normativas
 - 8.2.2. Asunción de responsabilidades, agencia impura e integridad
- 8.3. La insostenible fragilidad del juicio moral
 - 8.3.1. Jurisdicción, límites borrosos y perspectivas
 - 8.3.2. Censura, castigo e ironía
- 8.4. Diferencias culturales en el concepto de responsabilidad moral

Es imposible... enumerar las condiciones suficientes y necesarias del merecimiento personal en abstracto...

Las exigencias de la justicia difícilmente se agotan en el vacío principio de que, *ceteris paribus*, cada cual debe obtener lo que merece.

Joel Feinberg, *Doing and Deserving*.

La idea común a todos los capítulos de la Parte II ha sido la de que el reconocimiento de la suerte moral resulta inevitable. Esta tercera, y última, parte estará dedicada a analizar y explicitar las principales repercusiones de este reconocimiento. En concreto, defenderé que esta conclusión no es preocupante; y si lo parece, es sólo porque tenemos una idea equivocada de la agencia o de la moralidad, o de ambas. En este capítulo abordé directamente la significación y consecuencias de la revisión del principio de control. En el capítulo 9 me ocuparé de ciertas cuestiones metateóricas

relacionadas. Todo esto me dará también la oportunidad de comparar y distinguir mi posición de la de otros defensores de la suerte moral.

Recuérdese que en el capítulo 1 ofrecí un esbozo, que pretendía ser intuitivamente plausible y teóricamente neutral, del entramado que constituye el *sistema de la responsabilidad moral* —esto es, el sistema de normas, prácticas, actitudes, juicios y conceptos asociados a la responsabilidad moral— que nos sirviera al menos de punto de partida. Sin embargo, tras todo el camino recorrido hasta aquí, parece que algunos elementos cruciales de ese sistema, tal y como fueron esbozados, deberían ser corregidos o matizados. En este capítulo reconsidero algunos de estos elementos, a la luz de la aceptación de la suerte moral.

8.1. El principio de control

Comencemos por lo obvio. Afirmar la existencia de la suerte moral conlleva el rechazo del principio de control. En otras palabras, reconocer la suerte moral como un fenómeno no ilusorio supone privilegiar las prácticas ordinarias de juicio moral por encima del principio de control. Sin embargo, esto no implica necesariamente el rechazo radical del papel del control en la asignación de responsabilidad moral; sino que hay amplio abanico de posiciones posibles.

En esta sección investigo diferentes ataques al principio de control, que están conectados con concepciones diversas de la noción de control; y, a continuación, presento y discuto el argumento directo más obvio a favor del principio de control.

8.1.1. Nociones de control. El problema escéptico

Una línea de ataque radical al principio de control consiste en denunciar la misma falta de coherencia del tipo de control exigido. Esta es,

por ejemplo, la posición de Brynmor Browne, quien afirma que el negador de la suerte moral demanda un tipo de control que requiere que la agencia esté libre de toda contingencia, tanto en la producción de la acción misma —esto es, que sea una acción incausada, más allá de la intervención del agente—, como en lo que sucede después de su ejecución. Se presupone, así, que el agente es una especie de *instigador completo*.¹ Pero este tipo de control es claramente imposible de alcanzar. Ciertamente, Thomas Nagel, en su artículo original,² apela a una noción muy exigente de control, que parece demandar el control completo por parte del agente de todos los factores que contribuyeron causalmente al resultado, incluida la cadena causal que sigue a la elección; de modo que controlar un resultado se convierte en controlar todos los factores necesarios para este resultado. De hecho, aprovecha esta noción de control para conferir más fuerza a los que, para él, es la paradoja de la suerte moral. Pero esta es, sin duda, una noción demasiado fuerte de control. Así, si la *paradoja* de Nagel necesita de esta concepción del control, parece que podrá ser fácilmente resuelta a favor de las prácticas ordinarias.

Recordemos que para Nagel el fenómeno de la suerte moral constituía una paradoja, ante la cual no cabía más que una actitud pesimista. Su idea, sucintamente, era que la misma práctica de juicio moral es internamente paradójica. Por un lado, es condición del juicio moral que lo que está más allá del control de una persona no puede ser el fundamento de lo que merece. Implícitamente nos comprometemos a ver las cosas que están más allá de nuestro control como fuera del alcance de nuestra responsabilidad. Pero, por otro lado, consecuencias, circunstancias y factores consti-

¹ Browne (1992), pp. 347-8. Por supuesto, este argumento está en la estela del de Galen Strawson y otros; véase, en particular, Strawson (1988) y (1994).

² Nagel (1979).

tutivos del carácter —que la reflexión nos dice que están más allá de nuestro control— son tratados como relevantes para determinar nuestra responsabilidad moral. Y lo que es más, la aplicación consistente del principio de control nos llevaría a socavar la mayor parte de las evaluaciones morales que consideramos normal hacer. Así, la tensión implícita en nuestras prácticas de juicio moral se vuelve una paradoja paralizante cuando una “comprensión más completa y precisa de los hechos” se introduce en nuestras deliberaciones reflexivas. Cabe remarcar que, para Nagel, este resultado es una muestra particular, en el ámbito de la moralidad, del problema escéptico general; y vivir en este estado paradójico es, a su modo de ver, parte de la condición humana, así como comprender que vivimos en este estado forma parte de nuestra autocomprensión.

Sin embargo, como sugerí en el capítulo 2, la cuestión de la suerte moral, en sentido estricto, puede distinguirse —aunque sólo sea metodológicamente— de esta paradoja acerca del control. La primera tendría que ver estrictamente con las *prácticas ordinarias de juicio moral*; en concreto, con cómo cosas que están más allá del control de los agentes pueden marcar una diferencia en el juicio merecido. Mientras que la paradoja del control apunta más bien al escepticismo radical, o a lo que podemos llamar el *problema escéptico del control*. Recordemos que Nagel llegaba a esta formulación de la paradoja del control:

Una persona puede ser moralmente responsable sólo de lo que hace; pero lo que hace resulta en gran medida de lo que no hace; por lo tanto, no es moralmente responsable de aquello por lo que es y no es responsable. (Esto no es una contradicción, sino una paradoja.)³

³ Nagel (1979), p. 34.

Convenimos que este argumento era un tanto oscuro y, tras algunas variaciones, me decanté por esta reformulación:

1. Una persona *P* es moralmente responsable de la ocurrencia de un suceso *s* sólo si la ocurrencia de *s* no estuvo más allá de su control.
2. Ningún suceso es tal que su ocurrencia no está nunca más allá del control del agente. Por lo tanto,
3. ningún suceso es tal que *P* sea moralmente responsable de su ocurrencia.⁴

Sin embargo, esta paradoja, tal y como está formulada, puede salvarse con un sencillo movimiento. Para ello, tendríamos sólo que distinguir, como ha propuesto Michael Zimmerman, entre dos maneras en las que una cosa puede estar bajo el control de alguien: bajo el control *restringido* o bajo el control *irrestringido*. Tendríamos, por un lado, el *control restringido* de un suceso, que tiene lugar cuando una persona puede producir, e impedir, que este suceso ocurra; y, por otro lado, el *control completo* o *irrestringido* de un suceso, que tiene lugar sólo en caso de que una persona disfrute del control restringido tanto del suceso como de todos aquellos sucesos respecto de los que es contingente su ocurrencia. Con esta distinción, caben ahora dos lecturas divergentes del argumento anterior, ninguna de las cuales resultará convincente:

Versión primera:

- 1a. *P* es moralmente responsable de la ocurrencia de *e* solo si *P* tenía *control restringido* de *s*.
 - 2a. Ningún suceso es tal que alguien tenga alguna vez *control restringido* de él.
- Por tanto:

⁴ Véase Zimmerman (1987), p. 217. He sustituido “cuestión de suerte” por “más allá de su control”.

3a. Ningún suceso es tal que *P* sea moralmente responsable de su ocurrencia.

Versión segunda:

1b. *P* es moralmente responsable de la ocurrencia de *s* solo si *P* tenía *control irrestricto* de *s*.

2b. Ningún suceso es tal que alguien tenga alguna vez *control irrestricto* de él.

Por lo tanto:

3b. Ningún suceso es tal que *P* sea moralmente responsable de su ocurrencia.

El problema desaparece, pues ninguna de las dos versiones alternativas en que ha quedado el argumento resulta persuasiva. Las premisas 2a y 1b son obviamente falsas.⁵

Lo que cabe concluir a partir de esto es que no podemos identificar “estar bajo nuestro control” con “estar *completamente* bajo nuestro control”. Es un error pensar que si la ocurrencia de *x* depende de cualquier factor fuera mi control, entonces no controlo (realmente) *x*. O como mi control sobre cualquier cosa no es invulnerable, entonces no tengo (realmente, verdaderamente, últimamente) control. Pero, ¿cómo podría alguien tener este control? Y, en todo caso, esta no es la noción ordinaria de control, que es la que realmente estaría implicada en la noción ordinaria de evaluación moral —por bien que normalmente tampoco pensamos que carecemos de control sobre los efectos de nuestras acciones.

No obstante, no pretendo negar que en el trasfondo de la cuestión de la suerte moral se halle el problema escéptico del control; que, además, puede ser el que le confiere (parte de) su fuerza. Pero lo que sí nos permite este movimiento es desvincular (como programáticamente ya hice en el Apéndice a la Parte I, ahora con una mayor justificación teórica) el pro-

⁵ Zimmerman (1987), pp. 219-21.

blema escéptico, objeto del debate sobre el libre albedrío y la posibilidad de la responsabilidad moral, de la cuestión de la suerte moral. La *paradoja* realmente importante para nuestro tema es la que presuntamente se produce entre el principio de control y nuestras prácticas de juicio moral. Más exactamente, la suerte moral pone ante nuestros ojos el hecho de que un agente pueda ser más censurado que otro en virtud de factores que están más allá de su control.⁶

En el extremo opuesto, puede también defenderse una versión deflacionaria del principio de control, que preserve la intuición de que se requiere el control del agente para poder asignar responsabilidad moral, pero con un compromiso mínimo. De hecho, es innegable que el principio de control posee un cierto magnetismo, que reside en el hecho de que parece rastrear algo importante respecto a la atribución de responsabilidad: algo así como que aquello por lo que uno es moralmente responsable *ha de estar conectado con uno mismo de una manera apropiada*. Sin embargo, no está claro que esta idea equivalga necesariamente al principio de control. De hecho, si bien puede haber teóricos que acepten esta expresión como una paráfrasis adecuada del tipo de control requerido, también otros pueden considerar que ésta constituye de un requisito más general y laxo que el principio de control.⁷ De este modo, en el primer caso estaríamos ante una versión deflacionaria del principio de control y, en el segundo, ante el rechazo del principio. No obstante, nada más lejos de mi intención que convertir la discusión en una mera disputa verbal, o de matices míni-

⁶ Por otro lado, y como ya dije, la cuestión de la suerte moral persiste al problema escéptico. Véase el Apéndice a la Parte I.

⁷ Véase, por ejemplo, Adams (1985), p. 26.

mos.⁸ De hecho, aquello que confiere un contenido sustantivo al debate es precisamente que el principio de control es por definición incompatible con la existencia de la suerte moral.

Sobre esto último caben dos consideraciones. Primero, que el carácter intuitivo de algo así como la idea de que una persona sólo puede ser moralmente responsable por cosas con las que esté apropiadamente conectada, no es en sí incompatible con la existencia de la suerte moral. Respecto a esto, defenderé en las secciones siguientes que algo cercano a esta paráfrasis constituye una condición correcta para la atribución de responsabilidad moral, pero que es compatible con el reconocimiento de la suerte moral y, por lo tanto, diferente del principio de control, hasta el punto de chocar con su misma razón de ser.

En segundo lugar, cabe remarcar que ninguno de los más destacados defensores del principio de control —Nagel, Richards, Zimmerman o Greco, entre otros— dan ningún argumento directo a favor del principio de control; sino que simplemente lo asumen como un principio que rige, o debe regir, nuestras prácticas e intentan desenmascarar los supuestos casos de suerte moral como apariencias o como casos de prácticas erróneas. Así, simplemente lo toman como evidente, sin más. Sin embargo, esto no deja de ser controvertido. Tras lo visto en los capítulos anteriores, no parece que sea en absoluto *autoevidente*; sino, más bien, el reflejo de una visión sustantiva acerca de las condiciones bajo las cuales *deberíamos* juzgar o atribuir responsabilidad moral, para cuya vindicación se requiere algún tipo de argumento. En el apartado siguiente trataré de caracterizar un argumento central en la defensa directa del principio de control.

⁸ El problema reside en la dificultad de precisar exactamente en qué consiste el control. Particularmente, no confío que pueda hallarse una teoría general realmente significativa de qué es que el agente controle algo, o que algo dependa del agente. Lo más prometedor parece ser el estudio de los casos particulares.

8.1.2. El principio de equidad y el ideal de justicia última

El motivo principal que, explícita o —la mayoría de las veces— implícitamente, mueve a los adversarios de la suerte moral a adoptar tal actitud es la supuesta injusticia que el reconocimiento de la suerte moral supondría. Esta idea motiva el que bautizamos como Corolario del Principio de Control:

(Corolario de PC) Dos o más personas no deben ser moralmente evaluadas de manera diferente si las únicas diferencias entre ellas se deben a factores fuera de su control.

Y no deben serlo, se dice, porque resulta injusto censurar o considerar responsable a alguien por cosas (acciones, omisiones, rasgos, etc.) moralmente negativas que están más allá de su control. A menudo esta afirmación adopta la forma de un reproche: juzgar a alguien por algo que está más allá de su control es ser moralmente injusto e incluso cruel.⁹ Actuar de ese modo supondría romper con el que podemos llamar Principio de Equidad:

(PE) Todas las personas merecen ser tratadas por igual en relación a lo que está bajo su control.

El argumento básico resultante sería este:

1. Es injusto considerar a las personas moralmente responsables por cosas más allá de su control.

⁹ Ver, en particular, Domski (2004), pp. 463-4. Cabe añadir que, presuntamente, no puede decirse que *elogiar* a alguien incumpliendo el Corolario de PC sea cruel para con esa persona, aunque sí que puede ser injusto para con otras personas, y se incumpliría igualmente el principio que viene a continuación.

2. X es algo que está más allá del control de A.
3. Es injusto considerar a A moralmente responsable de X.

No obstante, y aunque parezca sorprendente, el Principio de Equidad (como el Principio de Control) ha recibido muy poco apoyo argumentativo explícito. En particular, muchos pueden haber considerado que se trata de un principio tan obvio que es innecesario dar razones a su favor, o que es tan básico que es manifiestamente difícil encontrar premisas más básicas.¹⁰ No obstante, parece que algún tipo de justificación puede, y debe, exigirse. Especialmente, cuando éste principio entra en conflicto con la intuición contraria de que las personas pueden ser censuradas por sus rasgos básicos, de carácter, o que son más censurables aquellas personas que de hecho cometen actos malvados que quienes meramente lo habrían hecho si hubieran tenido la oportunidad, o que efectivamente producen daño a alguien.

Por otro lado, el principio de control conecta con el elemento más central de cierta concepción de la moralidad: este principio es expresión, en palabras de Williams, de uno de los ideales “más conmovedores: el ideal de que la existencia humana puede ser finalmente justa”.¹¹ Todo el cúmulo de diferencias naturales, biológicas y psicológicas, así como sociales y culturales, iniciales y posteriores —como hemos estado viendo en los capítulos precedentes— marcan desigualmente a cada cual. Pero la moralidad parece que debe protegernos de esas contingencias de la suerte y la infortuna. Y lo que la concepción de la moralidad típicamente moderna (y, en concreto, kantiana) viene a ofrecernos es “consuelo a un sentido de la injusticia del mundo”; pretender ofrecernos una justicia última a

¹⁰ Véase Sher (2005), pp. 180-1, sobre esta posibilidad y su rechazo.

¹¹ Williams (1985), p. 195.

la que asirnos.¹² En la moralidad, en su concepción *pura*, anida la esperanza de que los agentes humanos puedan, a través de sus esfuerzos, crear y sostener un mundo moral que trascienda la suerte.

Sin embargo, y en primer lugar, los resultados de la Parte II de esta investigación sugieren que este ideal o esperanza de una justicia última y perfecta no es más que una ilusión. Pero por otro lado, e independientemente, en tanto que el Principio de Equidad constituye una meta moral determinada, también podría entrar en conflicto con otras metas morales y resultar finalmente apropiado su sacrificio en beneficio de una meta moral superior. Es por ello que en la sección siguiente caracterizaré una meta moral que confronta con el Principio de Equidad, a saber, la *asunción de responsabilidades*, que trataré de defender por encima de aquélla. Para ello destacaré algunos rasgos de nuestra noción de responsabilidad moral —particularmente, la importancia de su núcleo causal y la idea de *dependibilidad*—¹³ sobre los que, en general, creo que no se ha puesto el suficiente énfasis.

8.2. Contra el principio de control

Como he dicho, el principio de equidad entraba en conflicto con la práctica opuesta de censurar más a aquellas personas que de hecho cometen actos malvados que a quienes meramente lo habrían hecho si hubieran tenido la oportunidad, o las que efectivamente dañan a alguien, contra

¹² Williams (1981), pp. 21-2.

¹³ Con el neologismo “dependibilidad” traduzco la palabra inglesa “*dependability*”. Reconozco que no es una buena traducción, pero no he encontrado otra mejor. En inglés existe el adjetivo “*dependable*”, que significa “fiable, digno de confianza”, y de ahí el sustantivo. Optar por “fiabilidad” podría llevar a error.

quienes solamente lo intentan. Y esto apunta, como veremos, a la idea de que parte de lo que está en juego son los derechos de las personas.

8.2.1. Causación y asimetrías normativas

Me gustaría empezar esta sección con una anécdota que cuenta Elias Canetti en *Fiesta bajo las bombas*. A principios de la II Guerra Mundial, Canetti se reencontró en Londres con el pintor checo Oskar Kokoschka. Apenas empezaron a hablar, éste le hizo una confesión: él era el verdadero responsable de la guerra. Su explicación fue que era culpable de que Hitler, que quería ser pintor, se acabara dedicando a la política, pues ambos se habían presentado a la misma beca de la Academia de Bellas Artes de Viena, con el resultado de que Kokoschka fue admitido y Hitler, rechazado. Si en lugar de admitirlo a él, hubiesen admitido a Hitler, éste nunca se habría dedicado a la política, no habría existido el nazismo y no habría estallado la guerra.¹⁴ Por supuesto, esta opinión de Kokoschka es exagerada, incluso ridícula, y nadie puede considerarlo seriamente el verdadero responsable de la guerra por ganarle aquella beca a Hitler. Además, si Kokoschka experimentara realmente un sentimiento de culpa intenso por ello, no podríamos más que juzgar que se trata de un sentimiento inapropiado, o incluso patológico. No obstante, hay algo que

¹⁴ Canetti (2005), pp. 156-7. Esta idea puede rastrearse incluso en relación a la suerte antecedente. El siguiente fragmento de la novela de Dennis Lehane *Mystic River* (2001) es significativo a este respecto. Dice Jimmy Marcus: “La cuestión es, Sean, que si nos hubiéramos subido a ese coche y si nos hubieran llevado quién sabe dónde, y hubiéramos tenido que aguantar durante cuatro días todo lo que aquellos jodidos lunáticos hubieran deseado hacernos cuando tan sólo teníamos... ¿qué, once años?, no creo que hubiera sido tan osado a los dieciséis. Creo que habría acabado como un caso desahuciado y me habría atiborrado de tranquilizantes. Sé que nunca habría tenido lo que hacía falta para pedir relaciones a una mujer tan bella y tan arrogante como Marita. Y por lo tanto, nunca habríamos tenido a Katie. Y entonces nunca la habrían asesinado. Pero lo han hecho. Todo porque no nos subimos a aquel coche, Sean. ¿Entiendes lo que te quiero decir?” (p. 185.)

da sentido a esta opinión, y que no lo haría en el caso de alguien que afirmase lo mismo, o algo parecido, sin haber nunca interferido en la vida de Hitler (o sus antepasados): esto es, la conexión causal.¹⁵ La historia tiene sentido porque la conexión causal importa para la atribución de responsabilidad moral, e importa mucho más de lo que se suele reconocer.¹⁶

No obstante, lo que importa no es meramente la conexión causal, pues la importancia misma para el juicio moral de una conexión causal determinada depende de otros factores, como el signo o la gravedad moral de la conducta y sus consecuencias de hecho, que son el resultado de los cambios que el agente produce en el mundo. La imagen intuitiva del mecanismo de atribución de responsabilidad moral sería como sigue: primero determinamos si la conducta en cuestión del agente guarda la necesaria relación con sus estados mentales y, por lo tanto, si es responsable o no por ella; independientemente, valoramos moralmente la acción; y, a continuación, atribuimos responsabilidad positiva o negativa al agente por esa acción. Sin embargo, ciertos experimentos recientes muestran que en nuestras atribuciones ordinarias de responsabilidad moral hay diferentes relaciones psicológicas que son o no relevantes *según la conducta sea buena o mala*. Es conocido el estudio de Joshua Knobe en el que un agente lleva a cabo una acción porque quiere causar un efecto —el efecto de-

¹⁵ Algo parecido cabe decir en relación a la angustia de Jimmy Marcus. El texto anterior de Lehane (2001) continúa así: “Jimmy miró a Sean como si esperara una confirmación, pero Sean no tenía ni idea del tipo de confirmación que quería oír. Parecía necesitar que le perdonaran, que le absolvieran por no haber subido al coche cuando era un niño y por haber engendrado a una criatura que había sido asesinada.” (p. 185.)

¹⁶ La importancia de la conexión causal es la que nos permite, asimismo, entender, por ejemplo, la experiencia de verse comprometido por los crímenes del propio país, en oposición a los de un país que nos es ajeno, aun en el caso en que nos resulte imposible hacer nada por impedirlos. Véase Nagel (1979), nota 10. Lo mismo cabe decir del fenómeno de la *mácula moral*, esto es, de “tener el expediente moral mancillado [o “desfigurada la personalidad moral”] por la conducta injusta de aquellos con quien uno está relacionado”. Véase sobre esto, Oshana (2006). Ambos fenómenos están conectados con la idea de dependibilidad que exploraré en el apartado siguiente.

seado— pero es consciente de que también causará otros efectos que específicamente no desea —el efecto colateral previsto.¹⁷ Se trataba de poner a prueba si los sujetos participantes en el experimento consideran que el agente es responsable por producir los efectos colaterales previstos o no. Se les presentaron las dos descripciones siguientes:

Condición de daño

El vicepresidente de una compañía se presentó ante el presidente de la junta directiva y dijo: “Estamos pensando en poner en marcha un nuevo programa. Nos ayudará a incrementar los beneficios, pero dañará el medio ambiente.”

El presidente respondió: “no me importa en absoluto dañar el medio ambiente. Sólo quiero aumentar los beneficios lo máximo posible. Pongamos en marcha el nuevo programa.”

Pusieron en marcha el nuevo programa. Con toda seguridad, el medio ambiente resultó dañado.

Condición de beneficio

El vicepresidente de una compañía se presentó ante el presidente de la junta directiva y dijo: “Estamos pensando en poner en marcha un nuevo programa. Nos ayudará a incrementar los beneficios, y beneficiará además al medio ambiente.”

El presidente respondió: “no me importa en absoluto beneficiar el medio ambiente. Sólo quiero aumentar los beneficios lo máximo posible. Pongamos en marcha el nuevo programa.”

Pusieron en marcha el nuevo programa. Con toda seguridad, el medio ambiente resultó beneficiado.

La mayoría de los sujetos dijo que el vicepresidente merecía ser censurado en la condición de daño, pero pocos dijeron que merecía ser elogiado en la condición de beneficio. Hacer cosas buenas (como efectos secundarios) no dio origen al elogio, mientras que hacer cosas malas (como efectos secundarios) sí que originó la censura. Este resultado, conocido como

¹⁷ Knobe (2003). Véase también Knobe (2006) y Knobe y Fraser (2008).

el *efecto Knobe*, apunta a una asimetría entre la atribución de censura por efectos colaterales negativos y la atribución de elogio por efectos colaterales positivos. Las condiciones para la elogiabilidad son, a este respecto, más exigentes. Así, si quisiéramos saber si un efecto colateral es el tipo de cosa por el que el agente merecería ser elogiado o censurado, no sería suficiente saber meramente cuál es la relación psicológica del agente con el efecto; necesitaríamos saber si el efecto mismo es bueno o malo. Es de destacar que estos resultados se dan también respecto a la misma atribución de acción intencional: mayoritariamente, los sujetos juzgaron que el vicepresidente actuó intencionalmente al dañar (colateralmente) el medio ambiente, pero no al beneficiarlo.

Esta asimetría entre las condiciones de elogiabilidad y censurabilidad ha sido también confirmada en relación a las atribuciones de responsabilidad por intenciones no satisfechas. Supongamos que un agente tiene la intención de ϕ pero nunca tiene la oportunidad para ϕ . ¿Asignará la gente elogio o censura meramente porque uno tenga la intención de ϕ ? Malle y Bennett realizaron una serie de experimentos en los que invitaron a los sujetos a juzgar cuán responsables consideraban a los agentes por las siguientes acciones e intenciones:

Par positivo:

[acción]	ayudar a un vecino a arreglar el tejado
[intención]	intentar ayudar a un vecino a arreglar el tejado

Par negativo:

[acción]	vender cocaína a un primo adolescente
[intención]	intentar vender cocaína un primo adolescente

En ambos casos, se preguntó a los sujetos participantes en el estudio cuánto elogio o censura merecían los agentes. Esta metodología permitió

a los investigadores medir la *diferencia* entre la cantidad de elogio o censura dada a una acción y la cantidad dada a la correspondiente intención. En general, las intenciones recibieron menos elogio y censura que sus correspondientes acciones. Respecto a las intenciones mismas, el elogio disminuyó *dos veces* más que la censura. De nuevo, la respuesta varió significativamente según la conducta en cuestión fuese buena o mala.¹⁸

Por su parte, Pizarro, Uhlmann y Salovey han obtenido unos resultados interesantes en relación a la asignación de responsabilidad en casos en los que la conducta del agente se debe a emociones fuertes. En concreto, un grupo de sujetos recibió esta descripción de una conducta moral positiva:

Debido a su abrumadora e incontrolable simpatía, Jack le dio impulsivamente al mendigo su única chaqueta, aun a pesar del frío intenso del exterior.

Otro grupo, en cambio, recibió la descripción de un mal comportamiento:

Debido a su abrumador e incontrolable odio, Jack golpeó impulsivamente la ventana del coche de enfrente porque estaba aparcado demasiado cerca del suyo.

En contraste, otros sujetos recibieron descripciones en las que Jack se comportaba bien o mal, pero ahora “calmada y deliberadamente.” Resultó que los sujetos juzgaron que los agentes eran considerablemente menos censurables cuando actuaron mal movidos por una emoción, que cuando lo hicieron de modo calmado y deliberado; mientras que la diferencia en los grados de elogiabilidad por buenas acciones hechas a causa de una emoción fuerte y las buenas acciones realizadas calmada y deliberada-

¹⁸ Malle y Bennett (2004).

mente fue despreciable.¹⁹ Así, actuar movido por una emoción fuerte mitiga la censura en el caso de mal comportamiento, pero no disminuye el nivel del elogio cuando se trata de buenas acciones.

Pero la variación en la atribución de responsabilidad registrada experimentalmente no se limita sólo al signo moral de la conducta y/o las consecuencias. En un importante estudio, que llevó a cabo S. Walster, quedó demostrado que es más probable que la gente considere responsables a los agentes que producen un daño severo (por ejemplo, el daño serio a un niño inocente) que a los que causan un daño ligero (por ejemplo, un guardafuego dañado), incluso aunque en ambos casos la acción del agente sea descrita como igualmente negligente (aparcar en un colina sin haber comprobado los frenos).²⁰ Hay que añadir que el efecto del estudio original de Walster ha sido replicado posteriormente en un número de estudios adicionales usando metodologías bastante diferentes.²¹ Estos resultados, que establecen una asimetría en relación al nivel de daño, causaron sorpresa, pues la asunción previa era que los juicios de responsabilidad dependerían sólo del nivel de negligencia y no del nivel de daño final que el agente terminó produciendo (de acuerdo con el Principio de Control). Sin embargo, los datos indican que la severidad del daño afecta a la exigencia con la que determinamos el umbral de la negligencia, incluso aunque la severidad del daño sea a menudo simplemente cuestión de suerte.

Cabe destacar que estas variaciones en la atribución de responsabilidad moral no sólo suponen un refrendo de la suerte moral, sino que además corroboran muchas de las tesis particulares que defendí en la Par-

¹⁹ Pizarro, Uhlmann y Salovey (2003).

²⁰ Walster (1966).

²¹ Lo que es más, un reciente metaanálisis de 75 estudios sobre el tema deja poca duda de que el efecto es real. Véase Robbennolt (2000).

te II de esta investigación. Por supuesto, el hecho de que la gente haga estos juicios no implica que deban hacerse así. Y, de hecho, el mismo Walster concluyó que el efecto hallado en su estudio se debía a un prejuicio o sesgo motivacional: la gente no quiere creer que su conducta podría conducir a daños severos, y por ello están motivados a creer que sólo lleva a daños severos cuando el agente es negligente de manera culpable.²² En una línea similar, Darren Domsy ha defendido que

subconscientemente creemos que culpar mucho más al moralmente desafortunado que al moralmente afortunado nos da un privilegio a expensas de los demás, haciéndonos menos culpables de lo que lo seríamos de otra manera y haciéndolos más a los demás.²³

No obstante, y en línea con lo que he estado defendiendo en los capítulos anteriores, puede también defenderse que el efecto refleja un principio moral legítimo: las atribuciones de responsabilidad moral incluyen el objetivo de restituir a la víctima, y cuanto más severo sea el daño, mayor será la necesidad de restitución. Por ello, tiene sentido que la gente sea menos rigurosa en sus condiciones para atribuir responsabilidad en los casos en los que el daño es severo que en los que el daño es leve, hasta el punto de que extraordinariamente deje de atribuirle en casos de daño muy leve.²⁴ Por otro lado, el rechazo global de estos resultados experimentales,

²² Royzman y Kumar (2004) diagnostican un prejuicio similar en relación a la atribución de control en general.

²³ Domsy (2004), p. 462. Domsy apela a una gran cantidad de experimentos en psicología social que muestran que “[e]stamos egoístamente dispuestos a seleccionar las teorías y creencias morales que nos benefician, y tendemos a creer que somos especialmente afortunados en comparación con nuestros iguales” (2004, p. 455); según él, esto explica (y desenmascara) el supuesto fenómeno de la suerte moral.

²⁴ Véanse, a este respecto, los resultados experimentales de Shaver (1970).

dada su extensión, comportaría un revisionismo radical muy difícil de sostener.²⁵

En todo caso, parece indiscutible que los resultados de estos estudios apuntan todos al hecho de que el esquema cotidiano real de atribución de responsabilidad moral es bien distinto del intuitivamente plausible. Como adelanté, intuitivamente parece que el mecanismo consta de dos operaciones distintas y en serie: primero determinamos los hechos y establecemos la responsabilidad y, a continuación, determinamos qué responsabilidad, elogio o censura, y en qué grado es apropiada. Sin embargo, los estudios de las prácticas sociales nos muestran que las dos operaciones aparecen profundamente mezcladas.²⁶ Y lo relevante de este hecho para nuestro tema no es tanto cómo funciona realmente nuestro mecanismo psicológico de atribución de responsabilidad moral (si bien no deja de ser un hallazgo importante en sí mismo), sino que muestra que esas atribuciones están condicionadas por consideraciones normativas que tienen que ver con el signo moral de la conducta efectiva y con las consecuencias mismas de las acciones. Esto afecta incluso a la determinación de los juicios de causación, cuyos criterios de atribución varían también dependiendo del estatus moral de la conducta y de sus consecuencias.

Doris, Knobe y Woolfolk afirman en relación a esta cuestión:

La observación general es ésta: la importancia de los antecedentes causales y psicológicos de la conducta parece variar con el estatus normativo de la conducta y el resultado. En lugar de que el estatus normativo de la conducta o el resultado determine la pertinencia del elogio o la censura *subsiguiente* a una atribución de responsabilidad basada en atribuciones causales y psicológicas neutrales,

²⁵ Dejo la evaluación final de este y otros tipos de revisionismo para el capítulo siguiente.

²⁶ Véase Doris *et al.* (2007), p. 193.

en cambio la atribución de responsabilidad está *desde el principio* profundamente imbuida de consideraciones normativas: es, podemos decir, normativa “hasta el final”. Y en tanto que las consideraciones normativas importantes en diferentes contextos pueden ser altamente variables, es de esperar que las atribuciones de responsabilidad exhiban una variación similar.²⁷

En resumen, podemos concluir que la implicación causal desempeña un papel central en la atribución de responsabilidad moral, pero el mismo establecimiento de la implicación causal viene determinado por criterios normativos estrechamente vinculados al signo de la conducta del agente y de las consecuencias. En particular, una meta fundamental de la atribución de responsabilidad moral es responder al daño efectivamente causado. Esta idea recibe un apoyo considerable del sentido común y de nuestras prácticas. Así, nuestra comprensión de la relevancia de la responsabilidad causal para la justificación de las atribuciones de censura está constreñida por la noción de los derechos de la gente, a través del daño que se les ha producido. En general, esto apunta al hecho de que, para la justificación moral de nuestras respuestas apropiadas a la interferencia en la conducta de personas que han causado daño, es muchos menos importante la evaluación de la voluntad del agente de lo que muchos han pensado.

8.2.2. Asunción de responsabilidades, agencia impura e integridad

La meta moral que, como avancé, principalmente quiero confrontar con el Principio de Equidad es lo que podemos llamar la *asunción de responsabilidades*, una cuestión que ya ha surgido de pasada en capítulos precedentes. La relevancia moral de estar causalmente conectados con la

²⁷ Doris *et al.* (2007), p. 197.

producción de resultados dañinos reside en el hecho de que demandan que el agente asuma su responsabilidad.

Fijémonos, en primer lugar, en la posesión de creencias o actitudes moralmente negativas, como los casos del joven racista y la chica ingrata, que vimos en el capítulo 5. Desde la perspectiva de primera persona, asumir la responsabilidad por sus estados mentales forma parte de la respuesta moralmente adecuada. De hecho, resultaría del todo inapropiado que el propio agente los viese como cosas que meramente le suceden y no como algo suyo. Rechazar estos estados mentales, como ajenos, conllevaría una forma de alienación, además de la posibilidad de acusación de *mala fe*, en el sentido sartreano. El agente debe reconocer su culpa por haber mantenido una creencia moralmente equivocada; y, en el caso de la actitud malvada, además se requiere el esfuerzo de resistirse.²⁸ A su vez, el arrepentimiento y la culpa deben llevar aparejado el hacer algo por reparar la situación. Ser responsable significa también comprender la propia posición y responder de manera apropiada. Esto mismo, por supuesto, se aplica en relación a acciones y consecuencias dañinas.

Por otro lado, la idea de asunción de responsabilidades no es sólo retrospectiva. Sino que es también eminentemente prospectiva: el agente que se planea actuar asumiendo sus responsabilidades y que considera que estas responsabilidades no se limitan al principio de control, es mucho más fiable. En este sentido, el reconocimiento de la suerte moral tendría la consecuencia normativa positiva de lo que, siguiendo a Margaret Urban Walter, podemos llamar la *dependibilidad* del agente.²⁹ Walker ha defendido que sólo la realidad de la suerte moral da sentido a un importante ámbito de evaluación: el de la fiabilidad o dependibilidad personal.

²⁸ Véase Adams (1985), p. 20, sobre este punto.

²⁹ Walker (1991).

Para ella, la cuestión de la suerte moral tiene que ver con la concepción de la agencia que uno tiene y el tipo de responsabilidad moral que se deriva. Su manera de afrontar la cuestión consiste en comparar dos imágenes polarizadas de la agencia. Una de ellas es la concepción *pura*, impermeable a las *bagatelas* de la causalidad, o al menos a la causalidad externa a la voluntad del agente. Kant representaría principalmente esta visión, con su libertad noumenal, pero también incluye imágenes no trascendentales de la agencia, que comparten la inmunidad a la suerte —como es el caso, en particular, de la concepción de Zimmerman. A esta concepción, Walker opone una imagen *impura* de la agencia, que reconoce el orden causal, y las diversas maneras en que es condicionada y condicionante, sin que podamos delinear un límite *unitario* o *fijo* a su ejercicio y para todos los propósitos; lo cual supone un reconocimiento de la participación de los agentes en el mundo y de su exposición a la contingencia y el riesgo.³⁰

Prescindiendo de una interpretación rígida de esta distinción — que de ningún modo puede concebirse como dicotómica— entre dos imágenes de la agencia, podemos coincidir en que la suerte moral sería un problema, principalmente, para aquellos que tienen una concepción purista de la agencia, o un excesivo celo por salvaguardar la independencia extrema del agente. Ciertamente, los agentes puros disfrutarán de una mayor libertad o independencia en relación a las diversas cargas que la responsabilidad puede imponer a un sujeto. Por el contrario, la agencia *impura* enfatiza la significación moral de nuestra conexión causal con el mundo y la importancia de nuestra respuesta a lo que resulta de esta conexión. Sin duda, las cargas resultantes serán mayores, pero es precisamente este reconocimiento de la propia vulnerabilidad a la suerte, como dice Walker, la que concede al agente una fiabilidad humanamente valio-

³⁰ Walker (1991), p. 244.

sa. Un mundo con este tipo de agentes será un lugar mucho mejor para vivir.³¹

No obstante, los defensores del principio de control (de la agencia pura) pueden también reconocer la importancia de la asunción de responsabilidades hacia aquello que resulta causalmente de la propia conducta —como vimos, en el capítulo anterior, que remarcaba Susan Wolf— y, a su vez, negar que eso suponga tener que reconocer una extensión del ámbito de la culpabilidad. No obstante, de nuevo esta respuesta depende de una regimentación en la idea de responsabilidad moral que, aunque socorrida, no parece sólida. En efecto, no es extraño ver que muchos filósofos que se ocupan de la responsabilidad moral comiencen matizando que el sentido de responsabilidad moral que está en el centro del debate filosófico es el “sentido en que implica merecimiento”, según el cual ser moralmente responsable implica *merecer* elogio o censura, con independencia de ningún beneficio consecuencialista que la asignación de responsabilidad moral pueda comportar.³² Sin embargo, los resultados teóricos de la Parte II, más los datos empíricos anteriores, que recogen cuán intrincado es el funcionamiento de nuestras prácticas de atribución de responsabilidad, muestran lo insostenible de esta regimentación. En particular, contra esta regimentación cabe aducir el hecho de que nuestras prácticas de elogio y censura derivan *parcialmente* de sus efectos en nosotros.³³ Esto no

³¹ Véase Walker (1991), p. 242.

³² Véase, entre otros, Strawson (1986), quien define y caracteriza adecuadamente este sentido.

³³ Véase Vargas (manuscrito) para el desarrollo de esta idea. Manuel Vargas ha recuperado recientemente la tan denostada concepción consecuencialista de la responsabilidad moral. En realidad, sólo afirma que muchas consideraciones consecuencialistas (o que persiguen la “influencia moral” en el agente) son relevantes para la atribución de responsabilidad moral, y no que esta última pueda reducirse a las anteriores. Concretamente, su propuesta, que tiene la virtud de superar las objeciones típicas a la concepción consecuencialista, sostiene que la justificación de nuestras normas de responsabilidad deben

significa que las prácticas o actitudes individuales tengan siempre que influir en un agente particular de una manera adecuada, ni mucho menos que tenga que ser el propósito de los agentes implicados en ellas. Por supuesto, muchas de estas prácticas y actitudes no lo hacen. En particular, la persona virtuosa se sentirá culpable por haber actuado incorrectamente incluso cuando tener esos sentimientos no produzca ningún beneficio social; esto es, encontrará esos sentimientos apropiados en sí mismos. Lo que trato de destacar es la eficacia global que el sistema de la responsabilidad moral tiene para influenciarnos en pro del cumplimiento de las normas morales y de la posesión de un carácter virtuoso.

Walker tiene aún otro argumento a favor de la superioridad moral de la concepción impura de la agencia, que llamaremos *argumento de la integridad moral*, y cuya reconstrucción es la siguiente:

- (i) Si aceptamos el hecho de que la integridad moral es un valor que merece ser cultivado, siendo mejor un mundo con personas íntegras que uno sin ellas; y
- (ii) resulta que la integridad es una cualidad del carácter “imposible de capturar completamente sin hacer referencia a las vicisitudes de la suerte moral”;³⁴ o, en otras palabras, que la existencia de la suerte moral es condición necesaria para la integridad;
- (iii) entonces un mundo con suerte moral es mucho más deseable y moralmente valioso que uno sin ella.

Cabe aclarar que, en la concepción de Walker, la integridad constituye una protección de nuestro yo moral, es “ese centro de compromisos morales en uno mismo del que surgen las respuestas moralmente apropiadas y

comprenderse como la justificación de un conjunto de normas que fundamentan y apoyan una intrincada red de prácticas, actitudes y juicios.

³⁴ Walker (1991), p. 241.

valiosas de un modo seguro y constante”.³⁵ Una persona íntegra es aquella en la que se puede confiar que llevará a cabo sus compromisos y que asumirá la responsabilidad por sus acciones, aun cuando haya excusas disponibles. En otras palabras, la integridad está irremediabilmente unida a la mala suerte, en este sentido: es constitutivo de ser una persona íntegra el no desentenderse de las consecuencias desagradables afectivas de la propia conducta, lo cual depende fundamentalmente de que éstas se den. De nuevo, en un mundo sin integridad moral, en el que los agentes “rutinariamente y sin justificación se distancian de los resultados dañinos, crueles e incluso desastrosos, para cuya producción sus acciones fueron críticas, si no suficientes”, sería insoportable de vivir.

Este argumento ha sido también criticado en virtud de la idea de que lo crucial no son las consecuencias, sino a la intención o negligencia mismas; así como tampoco lo son las acciones efectivas, sino las acciones contrafácticamente asignadas. De este modo, resulta que la integridad del agente no puede consistir primariamente en la asunción de su responsabilidad sólo por las consecuencias o acciones efectivas negativas, sino por la misma intención o negligencia, o por las acciones que se pretendía llevar a cabo. Así, la insistencia de Walker en que el agente íntegro se caracteriza por hacerse cargo de los resultados dañinos de sus acciones, y hacer por repararlos, distorsiona la cuestión; pues tiene la consecuencia de que el agente afortunado se considere libre de culpa. Lo fundamental, dirá el negador de la suerte moral, es en realidad que el agente se forme las intenciones correctas, que se esfuerce por actuar correctamente y no cometa

³⁵ Walker (1991), p. 242. Sin duda, esta concepción de la integridad, así como la noción de agencia impura, están íntimamente inspiradas en las disquisiciones de Williams al respecto.

negligencias; y esto es lo que las normas de atribución de responsabilidad deberían fomentar.³⁶

Sin embargo, esta crítica es injustificada, pues —como venimos viendo durante toda esta investigación— la suerte moral marcará una diferencia cuyo efecto es el aumento de la responsabilidad moral en el siguiente sentido. Las consecuencias negativas incrementan la responsabilidad, o suponen una responsabilidad añadida a la responsabilidad inicial por la intención o la negligencia que produjo estas consecuencias. Los conductores afortunado y desafortunado son igual de censurables por su negligencia, pero el desafortunado lo es más por las consecuencias trágicas de su negligencia. O la realización de una acción aumenta la responsabilidad —no la anula— de tener una intención o rasgo de carácter determinado. Pero no se sigue de la aceptación de la suerte moral, ni del argumento de la integridad moral, que debamos considerarnos libres de toda responsabilidad cuando incurriendo en una negligencia tenemos suerte y no causamos ningún daño; así como que la gente no deba esforzarse por deliberar y actuar correctamente, evitando cometer negligencias. Por otro lado, la exigencia de que los agentes deberían preocuparse (movidos por un afán retribuidor) tanto por la víctima real de sus actos como por la potencial, y en el mismo grado o intensidad, parece altamente implausible, por cuanto las necesidades de una y otra son muy diferentes.³⁷

En este apartado he tratado de defender, contra las objeciones principales, los argumentos normativos, principalmente elaborados por

³⁶ Esta objeción se encuentra principalmente en Domski (2004), y también en Domski (2005); véase Statman (2004) para una réplica contundente.

³⁷ No incidiré de nuevo aquí en el carácter extremadamente exigente y altamente revisionista de la idea de que la cantidad de censura que los agentes negligentes merecen, con independencia de los resultados, debe ser la misma. Volveré sobre esto en el capítulo siguiente.

Margaret Walker, a favor de la superioridad moral de un mundo con suerte moral (de agencia impura). La idea de asunción de responsabilidades —a través de los argumentos de la dependibilidad y la integridad moral— da una razón más a favor de la suerte moral; y, en concreto, contra la creencia, derivada de la defensa del Principio de Equidad, de que sería más deseable un mundo en el que la suerte moral no existiera. Este resultado es más sólido en conexión con el resultado del apartado anterior, según el cual la justificación de nuestras prácticas de elogio y censura derivan *parcialmente* de sus efectos positivos. Incluso la misma determinación de la responsabilidad causal, presupuesto de las atribuciones de censura, está constreñida por el daño realmente causado.

8.3. La insostenible fragilidad del juicio moral

Ciertamente, el rechazo del principio de control, y del Principio de Equidad, nos deja un cierto regusto amargo. No hay duda de que se trata de una pérdida importante, especialmente en relación a nuestra autoconcepción, que es una autoconcepción principalmente moderna o ilustrada. Aun cuando suponga un beneficio moral superior o diferente, el reconocimiento de la suerte moral parece lesionar, en alguna medida, los ideales ilustrados de justicia e igualdad. De hecho, cabe reconocer que una idea como la de dependibilidad, consecuencia de la existencia de la suerte moral, debe mantener un equilibrio prudente con otra idea, como es la de autonomía personal. Si bien es cierto que el excesivo énfasis en la autonomía del agente comporta el peligro de disociación entre el yo y el mundo —esto es, el peligro de ver como ajeno todo lo que no es el propio yo—; también abusar de la idea de dependibilidad puede llevarnos a conclusiones equivocadas.

Estas consideraciones deben frenar el razonamiento de que el rechazo del principio de control implica tener que aceptar que las personas pueden ser moralmente responsables por cualquier cosa. Esto es, si aceptamos que una persona puede ser moralmente responsable por cosas que no controla, el problema consiguiente es el de hallar un límite alternativo a aquello por lo que uno puede ser moralmente responsable, a no ser que estemos dispuestos a sostener que uno puede ser considerado responsable por cualquier cosa. Esto último, de hecho, no es tampoco tan descabellado. Algunos teóricos, tras reconocer los múltiples factores que influyen en la formación del carácter, la voluntad y las intenciones del agente, que le sitúan en unas circunstancias y no otras, o que hacen que se desencadenen unas u otras consecuencias, han llegado a la conclusión de que *todos nosotros compartimos las acciones de cada uno*.³⁸ Pero si “compartimos” quiere decir que, de alguna manera, somos igualmente responsables, creo que es una afirmación difícilmente aceptable. Recuérdese en este punto mi rechazo (en el capítulo 3) de la idea de Michael Zimmerman de que *somos igualmente responsables de cosas que ni nos imaginamos*. En sentido absoluto, esto supondría una mitigación inadmisibles de la responsabilidad moral. Si todos somos responsables por todo, en realidad nadie parece ser verdaderamente responsable de nada. No quiero negar con esto que haya un sentido en el que esta idea puede tener visos de plausibilidad. Por ejemplo, Beardsley ha defendido que los juicios de elogio y censura pueden adoptar la “perspectiva de la igualdad moral última”, que se centra en los factores causales “últimos”, que llevaron al agente a ser quien es y actuar como actuó. Esta perspectiva no discrimina entre agentes, sino que destaca su naturaleza común. Beardsley mantiene que esta es una perspec-

³⁸ Concretamente, esta es la posición de Browne (1992), p. 353. A una conclusión similar, desde presupuestos opuestos, llegaba Michael Zimmerman; véase capítulo 3.

tiva que “forma parte de la verdad moral completa” acerca del agente, que es adecuado subrayar para ciertos propósitos.³⁹ No obstante, esta perspectiva es sólo una de las diversas perspectivas posibles y, en todo caso, no puede constituir algo así como la perspectiva *suprema* de juicio moral — entre otras cosas, porque no hay tal perspectiva suprema. Trataré de explicar esto más detalladamente.

8.3.1. Jurisdicción, límites borrosos y perspectivas

El convencimiento de que debe haber algún límite a la atribución de responsabilidad moral es, sin duda, un convencimiento acertado. Es por ello que en este apartado propondré un límite alternativo al principio de control para las atribuciones de responsabilidad moral, que llamaré la *jurisdicción* de cada cual y cuya demarcación resultará inevitablemente borrosa. A continuación, intentaré describir ciertos rasgos de los juicios morales que, especialmente en contraste con los juicios legales, hacen patente su especial dificultad y fragilidad. Esto último, además de contribuir a reafirmar la imposibilidad de fijar unos límites nítidos al alcance de nuestra responsabilidad *moral*, nos conducirá a la vindicación de la prudencia en el juicio y la pertinencia de una actitud irónica en esta esfera de la actividad humana —contenido del siguiente apartado.

Como dije más arriba, una persona sólo puede ser moralmente responsable de cosas con las que está apropiadamente conectado. Esto implica, en primer lugar, que es una condición de la atribución de responsabilidad moral que el agente esté causalmente conectado con aquello por lo que es juzgado responsable. Pero ha de estarlo, además, de una manera *apropiada*. No vale cualquier conexión causal —como, por ejemplo, la de

³⁹ Beardsley (1960).

Kokoschka con las causas de la II Guerra Mundial. De hecho la misma determinación de que la conexión causal se dé o no depende en gran medida de consideraciones normativas. Esto se ve claramente en relación a las omisiones. Imaginemos, por ejemplo, a una mujer que circula con su coche por una carretera desierta. De repente, otro coche que se aproximaba en sentido contrario hacia ella sufre un pinchazo, se sale de la calzada y choca contra un árbol. El conductor, inmóvil, parece malherido. Si la mujer acelera y se desentiende del accidentado, está claramente cometiendo una omisión de auxilio. No es extraño afirmar que uno de los ingredientes que posibilitan la atribución de responsabilidad es la conexión causal de la mujer con el accidente. Sin embargo, esta misma conexión depende del hecho normativo de que prestar auxilio (en condiciones como las anteriores: ausencia de peligro, ser el único presente, etc.) es una obligación moral. En un mundo en el que prestar auxilio no constituyese una obligación moral, no podríamos censurar a la mujer; y podría incluso decirse que entre la mujer y el accidentado no se da ninguna conexión causal. Cabe afirmar, pues, que una omisión es un tipo de acción que sólo se produce en virtud de determinadas consideraciones normativas —y que, por tanto, las conexiones causales dependen de consideraciones normativas. Algo parecido podría decirse en relación a cualquier tipo de acción.

En todo caso, la tesis que quiero defender, como condición limitadora de la responsabilidad moral, es esta:

(J) Una persona sólo puede ser moralmente responsable por cosas con las que esta apropiadamente conectada; y esta *conexión apropiada* viene definida por consideraciones normativas (normas morales) particulares a las diferentes prácticas de juicio moral.

Llamo a esta condición necesaria J (de “jurisdicción”) porque todas aquellas cosas que la satisfacen conforman la esfera sobre la que un agente tiene *jurisdicción*; esto es, aquello por lo que debe *responder* o es susceptible de tener que *dar cuenta*.⁴⁰ Creo que J es una condición correcta para la atribución de responsabilidad moral, que puede desempeñar más adecuadamente el papel limitador que los negadores de la suerte moral asignan al principio de control. Sin duda, J es una condición menos rígida, pues los límites que impone pueden variar según el contexto o el uso del juicio moral particular. Pero, lejos de ser un defecto, este hecho concuerda mejor con la naturaleza de nuestras prácticas de juicio cotidianas y constituye un aspecto destacado de lo que podemos llamar la *fragilidad* del juicio moral —más sobre esto a continuación. Además, es evidente que esta idea requerirá ulterior precisión, si bien esto sólo será posible en relación a contextos y usos particulares.

Cabe destacar que J no es una mera paráfrasis del principio de control, ni mucho menos supone su recuperación bajo un nombre distinto; pues J es compatible con el reconocimiento de la suerte moral y, por lo tanto, irreconciliable con el principio de control. De hecho, el carácter intuitivo de la idea de que una persona sólo puede ser moralmente responsable por cosas con las que esté apropiadamente conectada, no es en sí incompatible con la existencia de la suerte moral. En particular, J nos permite mantener esta idea y rechazar, al mismo tiempo, el Corolario del Principio de Control; cosa que no nos permitiría la defensa de cualquier principio de control, por débil que fuera.⁴¹

⁴⁰ Véase, en conexión con esto, la última sección del capítulo 5.

⁴¹ Recuérdese el Corolario de PC: “Dos o más personas no deben ser moralmente evaluadas de manera diferente si las únicas diferencias entre ellas se deben a factores fuera de su control.”

Así pues, un primer aspecto de la fragilidad de los juicios morales se muestra en los límites borrosos de nuestras respectivas *jurisdicciones*. Lo que cae dentro de la jurisdicción de uno varía de persona a persona, de un contexto a otro y, en definitiva, siempre depende del uso del juicio moral particular. De hecho, el juicio moral es mucho menos *definitivo* (“absoluto” y preciso) de lo que la mayor parte de sus tratamientos filosóficos parecen indicar.

Cuando evaluamos casos particulares, muy diferentes elementos son tomados en consideración. Pensemos, de nuevo, en el caso de los conductores ebrios. Al juzgarlos hemos de tener en cuenta lo siguiente. En primer lugar, nos preguntamos si cumplen los requisitos (psíquicos, agenciales y morales) mínimos para ser considerados agentes susceptibles de atribución de responsabilidad moral. Supongamos que, en condiciones normales, los cumplen. Tenemos, sin embargo, el estado etílico, que parece eximirlos. Pero dado que tomaron la decisión de beber y conducir anteriormente, son moralmente culpables; tienen, pues, una responsabilidad derivada o que surge de un origen anterior. Superado este punto, vemos que en un sentido, los dos conductores son igualmente culpables, pues decidieron beber y conducir; arriesgaron (moralmente hablando) demasiado, y son igualmente culpables de poner en peligro a otras personas. Pero, desde otra perspectiva, resulta que una persona ha muerto, y sólo uno de los conductores ha atropellado y provocado la muerte de alguien. El otro conductor no sería igualmente culpable, desde esta perspectiva. Sin embargo, una ulterior perspectiva nos debería obligar a reconocer que ninguno de los dos tenía la intención directa de matar a nadie (aunque puede que alguien diga que tomar el riesgo de conducir borracho es ya acceder a una suerte de preintención).

Comparémoslo con el caso de George y Georg, que proponía Zimmerman. Tenemos que ambos cumplen (directamente) los prerequisites para la atribución de responsabilidad moral. Si avanzamos, veremos que, desde una perspectiva, ambos tenían la misma intención, matar a alguien, y una intención formada y llevada adelante —ambos llevaron a término las acciones necesarias para alcanzar este objetivo, por bien que Georg no lo consiguiera. Pero desde otra perspectiva, George mató a una persona y Georg no; y de ello Georg no puede ser responsable. Ulteriormente, serán culpables de algo, realmente vil, de lo que los conductores anteriores no lo eran: tenían la intención directa de asesinar a una persona.⁴²

Cabe precisar que este análisis es diferente de la estrategia típica contra la suerte moral que he desestimado en la Parte II de esta investigación. Vimos que Zimmerman consideraba en relación, por ejemplo, al caso de los dos conductores, que quien atropella a alguien y quien no lo hace eran igualmente responsables en el grado, aunque difieran en el alcance. Y, a la vez, distinguía otros tipos de evaluación moral. Pero lo que singulariza a esta clase de interpretaciones de propuestas como la que yo defiendo es que aquí no hay un elemento que sea el únicamente relevante para la atribución de responsabilidad moral, sino que los tipos de evaluación moral o las diversas perspectivas desde las que podemos evaluar moralmente un agente, constituyen una pluralidad no necesariamente jerárquica. En algunos casos, una ordenación jerárquica clara de las diferentes perspectivas puede resultar más razonable, pero eso no ha de ser necesariamente así. Hay ocasiones, como muestran los casos de suerte moral, en los que dos o más de estas *perspectivas* devienen irreductibles, sin que se

⁴² Por supuesto, el reconocimiento de la suerte moral, como ya he dicho en repetidas ocasiones, no va en contra de otorgar a todas estas distinciones su debida relevancia.

pueda apelar a una jerarquía previa que organice las perspectivas — sometiendo unas a otras— y nos permita alcanzar un veredicto o *juicio final*. Se trata de puntos de vista diversos y relativamente independientes, que a menudo pueden sopesarse y combinarse, con la posibilidad de llegar a juicios unificados en muchos casos; pero no siempre.

De hecho, un error típico aquí es el de concebir los juicios morales sobre el modelo de los juicios legales y aspirar a la misma precisión de éstos últimos, cuando se rechaza que los juicios morales puedan estar regidos por criterios convencionales o pragmáticos. En otras palabras, por un lado se espera que los juicios morales alcancen la precisión de los juicios legales, pero por otro se rechaza que los juicios morales puedan depender, en ninguna medida, de criterios convencionales, que son los que precisamente permiten la precisión de los juicios legales. Si esto es así, es obvio que no habrá más remedio que reconocer que los juicios morales, cuyo carácter no convencional parece irrenunciable, serán necesariamente imprecisos. A diferencia de los juicios legales, los juicios morales no poseen un carácter (convencionalmente) definitivo. Hasta que sepamos qué asunto depende de nuestro juicio, es difícil saber qué juicio tener. Y si hacemos abstracción completa de las cuestiones prácticas, no existirá modo racional alguno de alcanzar un juicio moral. En relación a esto, Joel Feinberg ha afirmado que:

si consideramos que esta cuestión demanda un *juicio* frío y exacto, un “veredicto” emitido sobre la base de sucesos pasados, anidado en el pensamiento pero no necesariamente exteriorizado, entonces la cuestión está mal formulada, pues nos exige establecer una especie de jurado moral en un juzgado en el que las normas y razones jurídicas no pueden tener ninguna relevancia, pero donde nin-

gún tipo de normas y razones de forma remotamente legal pueden ocupar su lugar.⁴³

En todo caso, la imposibilidad final de llegar a un juicio absoluto o último, no se nos ha de presentar como un problema, pues la evaluación de responsabilidad moral no tiene por qué aspirar, como muestran muchos casos de nuestra experiencia moral, a alcanzar un *veredicto* que unifique los diversos aspectos en consideración y concluya la evaluación. No hay duda de que una meta razonable será la de intentar conseguir la mayor integración posible de las diferentes perspectivas, pero nada asegura una integración completa.

Quizá una exigencia que cabría hacer al tipo de caracterización propuesto sería la de demandarle una definición clara de las diferentes perspectivas posibles. Ciertamente, se podría intentar. Por ejemplo, habría que hablar claramente de una perspectiva que atienda a la intención del agente, y de una perspectiva centrada en los resultados, o de una “perspectiva genealógica” que rastree la génesis del mecanismo de decisión y acción del agente; también de la perspectiva del grado de identificación del agente con la acción, o del coste que le puede suponer la realización, incluso desde la perspectiva de la *batalla interna*, la perseverancia o el autoengaño, etc., etc. Obviamente, la lista puede alargarse mucho más, y, de hecho, es parte de la idea que la lista ni puede completarse, ni las perspectivas precisarse con independencia del tipo de casos evaluado. En todo

⁴³ Feinberg (1970), p. 41. Feinberg habla del *Error Legalista*, que comete quien, al plantear una cuestión moral usando términos de estilo legal, presupone acriticamente la precisión de ese término en su sentido estrictamente legal, a la vez que excluye la apelación a tipos de criterios que pueden decidir su uso.

caso, el juicio moral dependerá de las distintas perspectivas adecuadas a cada caso, cuyo peso y posibilidad de integración variará de caso a caso.⁴⁴

Sin duda, juzgar correctamente requiere un esfuerzo importante, más aun si resulta que no podemos llegar siempre a un juicio inequívoco. La adquisición de sabiduría moral no consiste sólo en aprender cómo hacer juicios correctos desde cada perspectiva moral, sino también en aprender a correlacionar las perspectivas, a darles su debida importancia, o incluso a saber bajo qué circunstancias debe prevalecer una y no otra. A su vez, sopesar toda esta diversidad de perspectivas evaluadoras, nos permitirá alcanzar una mejor comprensión de la situación del agente y su actuación. El resultado plausible puede ser una mayor tolerancia y comprensión de los demás.

8.3.2. Censura, castigo e ironía

No obstante, el negador de la suerte moral puede seguir incidiendo en la idea de que el rechazo del principio de control trae necesariamente consigo un tipo de injusticia. Castigar a una persona más que a otra por factores que, en último término, están más allá de su control es injusto. Reconocer este hecho, como ya hice, es una verdad amarga acerca de

⁴⁴ Por otro lado, a menudo tenemos una mala comprensión del proceso mismo por el llegamos a un juicio moral, que refuerza la imagen de inevitabilidad del “veredicto” alcanzado. Para decidir, en casos difíciles al menos, muchas veces deliberamos, sopesamos las consideraciones a favor y las consideraciones en contra, y evaluamos su balance. Una vez hemos llegado a una decisión, normalmente la explicamos y la justificamos citando sólo las razones de más peso que hicieron que nos decantáramos en la dirección del signo de nuestro juicio. Si lo que inclinó el argumento fueron los pros, justificamos el juicio en referencia a éstos; si fueron los contras, lo justificamos con estos otros. Normalmente, no solemos referirnos más a las consideraciones contrarias a nuestra conclusión, que no prevalieron. De esta manera, las decisiones parecen mucho más seguras e inevitables de lo que lo fueron realmente. Pero, claramente, esta imagen final es inexacta.

nuestra condición de agentes. Para terminar esta sección, evaluaré una propuesta que trata de minimizar esta consecuencia desagradable.

Ante una acusación de este tipo, un defensor de la suerte moral como B. Browne sostiene que si la suerte moral parece un problema es porque conectamos equivocadamente evaluación moral y castigo. Así, deberíamos corregir esta creencia para que los casos de suerte moral dejen de parecer preocupantes. La solución sería, ante una acción moralmente incorrecta, no responder con ira y censura *contra el agente*, sino con una censura que no incluye hostilidad ni deseo de castigo.⁴⁵ La idea es que es correcto censurar a alguien que efectivamente ha actuado incorrectamente o ha causado un daño, más que a sus contrapartes; pero no reaccionar hostilmente o castigarles. De hecho, para Browne, el castigo impone un sufrimiento que no está claramente justificado. La única clase de sufrimiento que realmente satisface el crimen es el remordimiento que surge en el agente por haber actuado incorrectamente; un remordimiento que, a diferencia del sufrimiento impuesto desde fuera, está internamente conectado con los hechos pasados, y es por eso que resulta mucho más apropiado. Quien siente remordimientos sufre de una manera correcta; de una manera que no comporta ningún problema de trato injusto. Además, esta misma actitud del perpetrador ante el daño cometido puede ser muy importante para la víctima.

No cabe duda de que el castigo objetivo no satisface igualmente el mismo crimen en diferentes personas. En la novela autobiográfica *Recuerdos de la casa muerta*, Dostoyevski protesta por el mayor sufrimiento que, en comparación con otros presos, él experimentaba durante su cautiverio en Siberia. “Ahora toma un hombre con corazón, de mente cultivada y de consciencia delicada. Lo que siente lo mata más que el castigo mate-

⁴⁵ Browne (1992), p. 350ss.

rial. El juicio que él mismo pronuncia de su crimen es más apesadumbrado que el del tribunal más severo, con las leyes más draconianas.” Es cierto que a menudo los castigos se vuelven muy diferentes según las diferentes personas que los reciben. No obstante, la propuesta de Browne afronta dos objeciones importantes.

La primera es el hecho de que el remordimiento depende en exclusiva de la respuesta emocional del agente que ha cometido el daño, de su actitud con respecto al daño cometido. Como es obvio, habrá quien no sienta ningún tipo de remordimiento, y así, si sólo el remordimiento es un castigo adecuado, resultará imposible someterlo a castigo alguno. O quien, por otro lado, reciba un castigo excesivo al sentir un remordimiento que excede aquello que hizo. En todo caso, el castigo escapará a cualquier tipo de autoridad externa. Ante esto Browne puede responder que, aunque no se puede garantizar la efectividad y proporcionalidad de este tipo de castigo, de todas formas se trata de un requisito moral fundamental para la justicia auténtica. Y posiblemente lo sea; pero hay otro problema, más acuciante, para esta propuesta.

Aun si consideramos satisfactoria esta respuesta, existe el problema de qué hacer con la censura misma. Pues, parece que una parte importante de la acusación anterior de injusticia se halla en el mismo hecho de censurar al agente, con independencia de ulteriores castigos. La censura es en sí dolorosa, vaya o no acompañada de un “trato duro”.

Alternativamente, una manera de hacer inteligible la propuesta de Browne podría consistir en tratar de distinguir la censura misma de las actitudes reactivas adversas —o “actitudes retributivas”— que la acompañan; esto es, de actitudes como la indignación o el resentimiento. De

hecho, Galen Strawson ha defendido la posibilidad de una “aproximación Zen a la vida” que elimine todas las actitudes reactivas.⁴⁶

Sin embargo, esta propuesta choca frontalmente con el punto de vista que considera que las actitudes reactivas son constitutivas del propio juicio moral. Para esta concepción, censurar a una persona “conlleva más que sostener una creencia particular acerca de ella; implica la voluntad de adoptar ciertas actitudes hacia esa persona y comportarse con ella de cierto modo.”⁴⁷ Desde este punto de vista, aquél que es censurado no es sólo calificado negativamente, sino que también se le convierte en objeto de emociones negativas, tales como el resentimiento y la indignación, y puede ser sometido a diferentes tratos adversos como “la evitación, el reproche, la regañina, la denuncia, la queja y (en el límite) el castigo.”⁴⁸ Estas acciones y actitudes de censura son desagradables, onerosas o dañinas, y tienen serias repercusiones negativas para las relaciones del sujeto consigo mismo y con los demás. Además, parecen establecer un tipo de demanda o expectativa en relación a la persona censurado: si eres censurado de esta manera, tienes que justificar tu acción (mostrar que no fue mala, después de todo) o reconocer su maldad y hacer por enmendarla y evitarla en el futuro. Así, para la concepción según la cual las actitudes reactivas son constitutivas del juicio moral, la propuesta de eliminar las actitudes reactivas (por lo menos) negativas o más duras supone acabar con los mismos juicios morales que estas actitudes determinadas contribuyen a constituir; lo cual desacreditaría la misma propuesta. Pero es posible que esta concepción no sea la más correcta. Pues puede también defenderse

⁴⁶ Véase Strawson (1986), pp. 117-120. Gary Watson considera esta posibilidad de eliminar las actitudes reactivas retributivas en Watson (1987). Esta meta sería acorde con el ideal de amor humano universal, quizá históricamente tan importante para nuestra civilización como la noción de responsabilidad moral misma (p. 286).

⁴⁷ Fischer y Ravizza (1998), p. 1.

⁴⁸ Wallace (1994), p. 54.

que las actitudes reactivas no son constitutivas de los juicios morales, sino que la relación es la de compañeros típicos. En este caso, la propuesta de su eliminación no sería incoherente.

Pero aún si esta segunda posición fuese la correcta al respecto, hay un ulterior problema para la propuesta, a saber: su incompatibilidad con el podemos llamar el Principio de Realismo Psicológico.⁴⁹ La posibilidad de unas relaciones interpersonales sin actitudes reactivas negativas choca con todo lo que sabemos acerca de la psicología humana. Esto hace que la propuesta resulte altamente implausible. Pero, en todo caso, si la propuesta no fuese finalmente incompatible con el principio anterior, el problema persistiría: aunque las emociones negativas asociadas pudieran ser eliminadas, parece que el juicio de censura mismo posee una fuerza valorativa (positiva o negativa) que hace que pueda seguir siendo injusto.

Si la propuesta de Browne ha de cumplir con su objetivo de neutralizar la posibilidad del trato injusto, resulta también incompatible incluso con la mera censura sin actitudes reactivas asociadas; pues la censura en sí misma, a diferencia de la mera descripción, contiene una profundidad o fuerza características. Específicamente, la censura posee una fuerza que surge de la especial profundidad o importancia del objeto de tal juicio; a saber, la valía moral del agente. Por ejemplo, Susan Wolf distingue lo que llama censura “profunda” de la censura “superficial”; las personas “merecen un tipo de censura distintivo y más serio por ser men-

⁴⁹ Sigo aquí a Owen Flanagan, quien ha propuesto el siguiente Principio de Realismo Psicológico:

(PRP) Asegúrese cuando construya una teoría moral o proyecte un ideal moral que las prescripciones propuestas sean posibles, o por lo menos sean percibidas como posibles, para criaturas como nosotros.

Cabe remarcar que PRP no es suficiente para fijar la teoría moral correcta, dado que el conjunto de psicologías morales realizables es infinitamente grande. De hecho, muchas de las que son realizables no serán buenas. PRP establece un criterio para evaluar las teorías que es fruto de una aspiración naturalista. Véase Flanagan (1990), cap. 2.

tirosas o mezquinas de lo que lo merecen los cerdos por ser desaliñados o los libros por estar desgastados.”⁵⁰ Además, ser moralmente censurado conlleva un tipo de crítica más serio que si te dicen que tu charla no ha estado muy bien, o que lo que cocinaste estaba soso, o que tu conversación fue aburrida. (Esta concepción suele ir unida a la imagen más clásica de la atribución de responsabilidad moral, que la asemeja a una especie de entradas en la cuenta de crédito o expediente moral de una persona, en el que se le suma o resta valía de acuerdo a su comportamiento.⁵¹)

Así, cabe afirmar que la misma censura moral ya incluye una fuerza característica que está contenida en el juicio moral mismo. No hay duda de que, analíticamente, es posible distinguir entre simplemente notar que una persona realizó un determinado acto incorrecto y *acusar* a esa persona por cometer ese acto. Pero la censura no tiene por qué identificarse sólo con lo segundo. El mero hecho de describir la conducta de un agente de un modo determinado (considerando que esa es una descripción *correcta*) ya comporta la censura.

Cabe remarcar que propuestas como la de Browne constituyen una deflación significativa de nuestras nociones de censura y responsabilidad moral que comportaría un cambio fundamental en nuestras prácticas morales cotidianas; un cambio tan fundamental que deberíamos preguntarnos

⁵⁰ Wolf (1990), p. 64.

⁵¹ Zimmerman afirma: “Se puede decir que elogiar a alguien supone juzgar que hay un “crédito” en su “libro de cuentas vital”, o un “signo positivo” en su “parte vital”, o un “lustre” en su “expediente como persona”; que su “expediente” ha sido “pulido”; que su “estatus moral” ha “mejorado”. Censurar a alguien supone juzgar que hay un “descrédito” o “débito” en su “libro de cuentas”, o un “signo negativo” en su “parte” (1988, p. 38). Véase también Feinberg (1970), pp. 124-5 y Glover (1970), p. 27. Es concepción de la atribución de responsabilidad moral no es incompatible con el reconocimiento del papel desempeñado por los sentimientos asociados, ni su utilidad. La disputa, parcialmente independiente, se da entre considerarlos o constitutivos o meramente complementarios al juicio de responsabilidad moral. Para este contraste, véase Watson (1987); Fischer y Ravizza (1998), pp. 1-4; y Hieronymi (2004), pp. 119-125.

si seguirían siendo las mismas nociones o, al menos, *nuestras nociones*. La salida hacia adelante, si uno está dispuesto a seguir en la línea de Browne, sería la eliminación de la misma censura moral. Precisamente, Michael Slote ha defendido esta vía: eliminar de las nociones de censurabilidad, culpabilidad, etc.; sustituyéndolas por las nociones de crítica moral y de carácter virtuoso o vicioso del agente, pero sin implicar con ellas el elogio o la censura morales.⁵² Para ello se necesitaría una reforma parcial del vocabulario y aparato conceptual éticos. En definitiva, lo que se promueve es el abandono de todo tipo de actitud moralmente adversa hacia los demás.

Sin embargo, esto conllevaría el abandono de una manera básica en que reaccionamos al mal, y de una manera en que cotidianamente tratamos de *hacer justicia*, aunque sea de manera irremediamente imperfecta. Supondría renunciar a la posibilidad misma de impartir *moralmente* justicia. Por otro lado, resulta extraño decir que no se debería censurar a alguien aunque uno crea que ha realizado una acción moralmente negativa o, en general, tenga una opinión pobre de él.⁵³ Además, afirmar que las críticas y reproches no incorporan una censura *moral* presupone de nuevo una regimentación muy estricta de las prácticas de juicio moral (algo así como tipos cerrados en sí mismos e independientes unos de otros), que no se corresponde con nuestras prácticas y actitudes cotidianas. Finalmente, las actitudes reactivas (constitutivas o no del juicio moral) desempeñan — como destacó Peter Strawson— un papel tan fundamental en nuestras relaciones personales y en nuestra propia autocomprensión que dejar de elogiar y censurar no puede ser la solución.⁵⁴ Las acciones humanas de-

⁵² Véase Slote (1994).

⁵³ Sobre este punto, véase Adams (1984), pp. 21-4.

⁵⁴ Strawson es muy claro al respecto: “Pero nuestra proclividad *general* a estas actitudes y reacciones está, de manera inextricable, estrechamente vinculada a esa participación en

mandan una amplia variedad de respuestas que incluyen muy diversos matices emocionales.

Cabe notar que la propuesta anterior tiene mucho en común con la actitud típica de rechazo del juicio moral; esto es, la acusación de que sólo los presuntuosos o fatuos osan juzgar a los demás. Los clichés más socorridos suelen ser los de que “nadie tiene derecho a juzgar a los demás”, o que “nadie está libre de pecado”. Por supuesto, la irreflexiva severidad para con los demás (y también para con uno mismo) es del todo desaconsejable. Sin duda, debemos estar alerta ante el peligro de caer en una actitud censora de las faltas morales ajenas; pero esto no arruina la legitimidad del juicio moral mismo. Como afirmó Hannah Arendt:

La reflexión de que quizá uno se hubiera comportado mal, en el caso de encontrarse en las circunstancias de quienes así se comportaron, quizá dé lugar al nacimiento de cierto espíritu de perdón, pero aquellos que en la actualidad traen [esto] a colación [...] parecen encontrarse un tanto confundidos.⁵⁵

La meta principal del juicio moral es la justicia, no la misericordia —y la justicia incluye reconocer los derechos dañados de las víctimas. Además, disolver nuestra indignación hacia el autor de un crimen, sobre la base de la influencia de la suerte, supondría su deshumanización, además de evidenciar una despreciable actitud paternalista por nuestra parte. La aceptación de la suerte moral, tanto en la formación de nuestro carácter, como

interrelaciones sociales y personales que comienza al tiempo que nuestras vidas, que se desarrolla y complica de manera muy diversa a lo largo de la misma y que es, podríamos decir, una condición de nuestra humanidad. Ese compromiso ineludible constituye un hecho natural, algo tan profundamente enraizado en nuestras naturalezas como nuestra existencia como seres sociales.” (1985, II, 1; 2003, p. 81.)

⁵⁵ Arendt (2006), p. 430.

en las circunstancias que afrontamos y en las consecuencias de nuestras acciones, nos conduce al reconocimiento de la contingencia de nuestro yo moral y de nuestro merecimiento y logros; lo que, finalmente, no puede más que “infundir —como ha dicho Watson— *ironía* a las actitudes reactivas.”⁵⁶ El reconocimiento de la suerte moral debe, sin duda, impregnar nuestros juicios morales de una cierta actitud irónica que facilita el perdón; pero no debe impugnarlos. Creo que este espíritu queda muy bien recogido en las siguientes palabras de Arendt, al enjuiciar a Eichmann:

Has contado tu historia con palabras indicativas de que fuiste víctima de la mala suerte, y nosotros, concedores de las circunstancias en que te hallaste, estamos dispuestos a reconocer, hasta cierto punto, que si éstas te hubieran sido más favorables muy difícilmente hubieras llegado a sentarte ante nosotros o ante cualquier tribunal penal. Si aceptamos, a efectos dialécticos, que tan solo a la mala suerte se debió que llegaras a ser voluntario instrumento de una organización de asesinato masivo, todavía queda el hecho de haber, tú, cumplimentado y, en consecuencia, apoyado activamente, una política de asesinato masivo.⁵⁷

8.4. Diferencias culturales en la atribución de responsabilidad

Para finalizar este capítulo y, en parte, a modo de adenda, me gustaría decir algo sobre una cuestión importante en relación a la concepción general de la responsabilidad moral que he defendido. Dado que he mantenido que para entender la noción de responsabilidad moral tenemos necesariamente que tomar en consideración la conexión del concepto con nuestras prácticas ordinarias al respecto, pudiera suceder que éstas últimas variaran sustantivamente entre culturas y grupos sociales, de manera que

⁵⁶ Watson (1987), p. 275. También Vogel (1993) ha incidido en esta idea.

⁵⁷ Arendt (2006), p. 405. Mis cursivas.

el análisis ofrecido quedara en algún sentido comprometido. Así, la cuestión con la que quiero finalizar este capítulo es la del alcance tanto histórico como geográfico del concepto de responsabilidad moral. La cuestión que mueve esta indagación es si el concepto de responsabilidad moral es culturalmente variable o si, por el contrario, es interculturalmente estable; y, en todo caso, cabrá saber en qué medida lo es o no. A continuación, intentaré extraer las consecuencias que se derivarían de la respuesta a estas cuestiones para nuestro debate. Puedo adelantar que mi conclusión será tentativa y condicional. Sugeriré que la existencia de diferencias culturales sustantivas en el concepto de responsabilidad moral, si las hay, sería significativa para las diferentes actitudes de aceptación o rechazo del fenómeno de la suerte moral.

Ciertamente, el uso del predicado “responsable” y la práctica de tener a alguien por tal, así como las actitudes implicadas, es algo que viene de antiguo y cuya extensión a lo largo del planeta es universal. Sin embargo, parece que sus usos han cambiado con el tiempo. Siglos atrás las atribuciones de responsabilidad fueron mucho más generalizadas, en el sentido de que se responsabilizaba a alguien por ciertas acciones cuyo control hoy consideramos que está claramente fuera de su poder. No obstante, éste bien puede ser un cambio en la *aplicación* del concepto de responsabilidad y no en el concepto mismo. Algo así sucedería en las condenas por brujería; o, más contemporáneamente, en el caso de Phineas Gage, respecto al cual sólo recientemente se ha descubierto que su incapacidad para observar las convenciones sociales y comportarse éticamente (e incluso tomar decisiones ventajosas para sí mismo) no se debía a un quebrantamiento voluntario —o involuntario pero indirectamente controla-

ble—, sino a los daños sufridos en sistemas cerebrales específicos.⁵⁸ Lo que ha cambiado, pues, son una serie de creencias en cuanto a las condiciones de aplicación, debido a un aumento del conocimiento fáctico disponible. En particular, ha cambiado el conocimiento de los daños cerebrales, de su alcance y sus formas, y de que ciertos estados mentales y acciones se deben a ellos; pero no la creencia de que una privación sería constituye un eximente de la responsabilidad moral.

Sin embargo, este tipo de variación no es el único posible, ni el que aquí más nos interesa. Lo que se pretende indagar es la variabilidad o no del propio concepto de responsabilidad moral, esto es, de las *creencias* comúnmente compartidas y directamente relacionadas con las atribuciones de responsabilidad moral.

La cuestión es: ¿de qué datos disponemos acerca de la universalidad del concepto de responsabilidad moral? En este punto, lo adecuado sería examinar la información disponible al respecto, que los diferentes estudios descriptivos, tanto históricos, etnográficos o sociológicos, como de historia de las ideas e incluso literarios, nos brindan. Se trataría, en cualquier caso, de indagar en las diferentes prácticas culturales y sus presupuestos. Como resultará obvio, no dispongo del espacio, ni del conocimiento, suficientes, para llevar a cabo un estudio en extenso de las diversas fuentes—que requeriría una tesis doctoral independiente. No obstante, lo que sí que puedo hacer es recoger ciertos puntos de vista de especial relevancia en torno a la cuestión e intentar extraer de ellos ciertas conclusiones. Una cautela adicional a tener en cuenta es la de que, obviamente, estos estudios nunca son meramente descriptivos, sino también en gran medida interpretativos.

⁵⁸ Véase Damasio (1994).

Existe una célebre distinción, en relación a la articulación social de las diferentes asunciones y reglas respecto a la responsabilidad moral, entre culturas de la *vergüenza*, por un lado, y culturas de la *culpa*, por el otro. La descripción más común es que, mientras que una cultura de la vergüenza suele estar dominada por estándares de comportamiento correcto de tipo social y por la aspiración individual al reconocimiento público, en una cultura de la culpa el individuo tiende a alcanzar sus propios estándares morales, los cuales se imponen a la conveniencia social: aquí el árbitro último es la propia consciencia. Sin embargo, esta distinción debe ser matizada y enriquecida.

A grandes rasgos, los estudios psicológicos generalmente nos dicen que el sentimiento de culpa refleja la conciencia del agente de haber violado alguna regla, lo cual le lleva a centrar su atención en el daño causado y en la propia víctima; éste ansía la reparación y el propio castigo, mientras provoca el odio y el resentimiento en los demás. Por otro lado, la vergüenza se focaliza en la identidad del propio agente, en relación a su creencia de cómo es visto por los demás; el sujeto deviene consciente de sus malas cualidades, de haber fallado en la satisfacción de su propio ideal del yo, con lo que su auto-imagen se resiente. La reacción del sujeto consiste en esforzarse por mejorar, o simplemente quitarse de enmedio; la reacción de los demás es el escarnio y el desprecio. En síntesis, pues, si la culpa parece apuntar a acciones y transgresiones, la vergüenza se focaliza en la propia identidad del agente. No obstante, la misma pretensión de establecer una distinción nítida entre ambas emociones es cuestionable. En concreto, la vergüenza puede también surgir a causa de comisión de un acto que viola de una regla, así como una conducta vergonzosa puede provocar el ansia de reparación, la indignación en los demás y la expecta-

tiva de ser castigado. Igualmente, el sentimiento de culpa va naturalmente asociado a la automejora y daña la identidad del sujeto.⁵⁹

En todo caso, la determinación de un sentimiento culturalmente central en relación a la responsabilidad moral ha servido para contraponer diferentes tipos de culturas. En este sentido, hay que destacar la interpretación usual de la cultura griega clásica como una cultura de la vergüenza; o el uso de la idea de evolución de una cultura de la vergüenza a una cultura de la culpa para describir el paso de la sociedad medieval, dominada por la religión, al mundo moderno occidental individualista; y aún, sincrónicamente, para caracterizar las diferencias entre Oriente —para el cual se ha acuñado también el calificativo de “cultura del honor”— y Occidente. Cabe destacar el clásico estudio que la antropóloga norteamericana Ruth Benedict realizó durante la II Guerra Mundial y en el que hace uso de esta noción como marco de comprensión de las claves de la cultura japonesa.⁶⁰ En resumen, frente a las culturas de la vergüenza, el paradigma de cultura de la culpa lo constituiría la cultura occidental moderna, con su herencia judeocristiana.

No obstante, esta distinción debe considerarse sólo aproximativa y no *categorica*. Ya E. R. Dodds hizo hincapié en la creciente sofisticación en el desarrollo de la épica y el drama griegos, que evolucionaron desde una concepción arbitraria del mundo y el orden moral y de un sujeto al capricho de los dioses, a una comprensión posterior de los límites de la

⁵⁹ Véase Stocker (2007) para una crítica importante del modo usual de caracterizar y distinguir el sentimiento de culpa y la vergüenza, en la que se basan estas consideraciones.

⁶⁰ Véase Benedict (1945). Cabe aclarar que Benedict presenta la distinción como una cuestión gradual o de énfasis; ha sido la posterior popularización la que la ha convertido en una distinción categorica. Cabría destacar asimismo, a nivel de pensamiento, los resultados más recientes de los estudios de Robert Nisbett en psicología cultural que indican diferencias sistemáticas entre asiáticos y occidentales debidas al mayor énfasis “individualista” occidental frente al mayor “holismo” oriental. Véase Nisbett (2003).

responsabilidad moral.⁶¹ De manera más radical, Bernard Williams ha denunciado recientemente la comprensión equivocada de las diferencias entre los griegos y *nosotros* mismos “en términos de un desplazamiento en las concepciones éticas básicas de la actuación, la responsabilidad, la vergüenza y la libertad”;⁶² un error que podemos achacar a una visión progresista injustificada, que tacha de primitivas las nociones de agente y responsabilidad de la cultura griega clásica y, por ende, del Medievo y de Oriente, en contraposición a las de la cultura occidental moderna.

Ciertamente, la distinción entre culturas de la vergüenza y culturas de la culpa está inmersa en una idea más amplia de progreso, de formas más simples a formas más complejas, en el desarrollo tanto de individuos como de culturas. Es de destacar que, en sus primeras interpretaciones de la vergüenza, psicoanalistas y antropólogos coincidieron en considerarla una emoción inmadura en contraste con la culpa. Esta idea fue reforzada por la observación de que la vergüenza precede a la culpa en el desarrollo individual. Y, así, el resultado ha sido la tendencia a ver en la culpa una emoción mucho más civilizada y avanzada, dentro de un esquema progresista sospechosamente favorable al Occidente contemporáneo.

Sin embargo, además de lo dudoso que resultaba —como hemos visto— la creencia de que existen unos criterios exactos de discriminación entre las emociones de culpa y vergüenza, no está tampoco justificado identificar por completo el dominio de la culpa o la vergüenza con una

⁶¹ Véase Dodds (1951).

⁶² Williams (1993b), p. 7. Williams, como Dodds, no rechaza la distinción misma, sino las simplificaciones a que ha sido sometida, así como la visión progresista que asocia el énfasis en la vergüenza con un estadio primitivo del desarrollo moral. Asimismo, mantiene que los griegos no carecieron de la noción de culpa, sino que “no hicieron de esas reacciones [las que nosotros asociamos con la culpa] la cosa especial en que se convirtieron” (p. 91).

determinada cultura, pues dentro de una misma área cultural o incluso en una misma sociedad se dan variaciones significativas entre grupos de población diversos. Por ejemplo, los psicólogos culturales Nisbett y Cohen han defendido la existencia de una “cultura del honor” entre los hombres blancos del sur de Estados Unidos, caracterizada por unos patrones de conducta y unas actitudes peculiares, que explicaría la gran tasa de homicidios allí existente. La “reputación por la fuerza —afirman estos autores— es la esencia de la cultura del honor, el individuo que insulta a alguien debe ser forzado a retractarse; si el instigador rechaza hacerlo, tiene que ser castigado mediante la violencia e incluso con la muerte”.⁶³ Lo fundamental del estudio es que este grupo es mucho más proclive a verse envuelto en homicidios que resultan de discusiones, que no en homicidios ocurridos en el curso de un atraco u otro delito. Así, parece que la distinción se encuentra en una concepción diferente de aquello por lo que alguien debe responder y no meramente en una mayor tendencia general a la violencia.

Pero aun restringiéndonos a instancias centralmente características de la cultura occidental contemporánea, proclamada paradigma de la cultura de la culpa, es difícil no ver el papel significativo que también la vergüenza juega en ella.⁶⁴ La vergüenza es fundamental, en muchos aspectos,

⁶³ Nisbett y Cohen (1996), p. 5. Nisbett y Cohen apoyan su hipótesis en datos demográficos y encuestas, además de en una variedad importante de experimentos. Por ejemplo, en un estudio de laboratorio registraron cómo, de los sujetos participantes (hombres blancos del norte y del sur de EEUU), sólo los sujetos sureños experimentaron un incremento drástico del nivel de cortisona (hormona asociada con altos niveles de estrés, ansiedad y excitación) y de testosterona (hormona asociada con la conducta agresiva y dominante) tras un incidente en el que un ayudante encubierto del estudio fue molestado por un sujeto no sospechoso que pasaba por el pasillo y aquél, fingiendo fastidio por la interrupción, le golpeó y le insultó (pp. 45-48).

⁶⁴ Williams ha criticado la asunción de que la culpa por sí sola caracterice mejor nuestras reacciones ante el fallo moral; afirma que ésta “no puede por sí misma ayudar a alguien a comprender su relación con esos sucesos [malas actuaciones, etc.], o a reconstruir el yo que ha hecho esas cosas y el mundo en que ese yo tiene que vivir. Sólo la vergüenza

para nuestras prácticas sociales, tanto a nivel educativo, como en las relaciones laborales, personales, etc.⁶⁵ En realidad, la culpa y la vergüenza se solapan significativamente en la cultura de Occidente, y ambas son constitutivas de nuestras prácticas morales. Y, en general, culpa y vergüenza han convivido y conviven en todas las culturas y grupos sociales; en todo caso, se trataría de una cuestión de mayor o menor énfasis.

Recopilando todas estas consideraciones, se puede afirmar que los estudios localizan normalmente la cultura de la vergüenza en las sociedades o grupos sociales más colectivistas, del pasado o del presente, y la cultura de la culpa en las sociedades occidentales modernas, paradigmáticamente individualistas. De hecho, podría haberme limitado, sin más, a la distinción clásica entre colectivismo e individualismo, sin embargo poner el énfasis en la idea de emociones dominantes enriquece la caracterización del contraste. En todo caso, la conclusión (de carácter eminentemente provisional) a la que llego es que las creencias compartidas que regulan la responsabilidad y la censura morales parece que divergen entre los diferentes grupos humanos; esto es, que existen diferentes concepciones de la responsabilidad, más o menos influidas por la vergüenza o por la culpa, o de carácter más colectivista o más individualista. Pero ésta divergencia no puede identificarse sin más con un corte entre la cultura occidental contemporánea y el resto de culturas, pues también en ésta se da una variación interna al respecto; tanto la vergüenza como la culpa juegan su papel, y la confrontación entre colectivismo (moderado) e individualismo está bien viva.

puede hacer eso, porque ella incorpora las concepciones de lo que uno es y de cómo uno se relaciona con los demás” (Williams, 1993b, p. 94).

⁶⁵ Una persona puede sentir vergüenza por cosas sobre sí misma que no hizo intencionalmente, o incluso que no dependieron en absoluto de ella; de este modo puede darse cuenta de que debería intentar controlar o cambiar esto.

Cabe destacar que este resultado puede ser un problema para las teorías contemporáneas de la responsabilidad moral, que tienen como característica común la universalidad —para todos los agentes, en todas las sociedades— de las condiciones o criterios que proponen para la responsabilidad moral.⁶⁶ Pero, ¿cómo afecta esta variabilidad social a la cuestión de los límites de la atribución de responsabilidad moral? Un elemento central de toda concepción de la responsabilidad moral es la importancia de la voluntariedad del agente; si bien, no obstante, ninguna concepción se limita exclusivamente a lo voluntario, ni a lo intencional. Un caso obvio: un acto involuntario fruto de la ignorancia culpable del agente no le exime típicamente de la culpabilidad.⁶⁷ Pero aún su restricción a aquello que el agente controla no resulta satisfactoria, a no ser que se interprete de una manera no muy precisa. Pero, en este caso, la noción de *jurisdicción* que propuse más arriba, recoge mucho mejor, a mi entender, esta idea, en tanto que claramente la hace depender de un contexto y uso determinado del juicio de responsabilidad moral. Esto no sólo es claramente compatible con el reconocimiento de su variabilidad cultural, sino que viene reforzado por ella.

La variabilidad cultural es principalmente un problema para aquellos que comparten una visión de la responsabilidad moral que deriva de la *concepción purista de la agencia*, caracterizada por el rechazo de ideas como la de dependibilidad —que vimos más arriba. De nuevo, éstos pueden responder a lo anterior arguyendo que, por ejemplo, la concepción propia de las culturas del honor confunde diferentes ideas acerca de la

⁶⁶ Para una indagación de esta cuestión, véase Sommers (en prensa).

⁶⁷ Estas consideraciones conectan, en particular, con lo dicho en las dos últimas secciones del capítulo 5.

responsabilidad moral, que deben de ser disociadas para un correcto tratamiento de la cuestión. Se confunde *lo que el agente controla* con *aquello a lo que se espera que responda*. Para la concepción purista de la agencia, lo verdaderamente relevante en cuanto a la responsabilidad moral es lo primero, y no lo segundo; y considera que éste es nuestro verdadero concepto de responsabilidad moral. Como traté de argüir más arriba esta concepción, aunque tiene su fuerza, es errónea. La afirmación de que el *verdadero* concepto de responsabilidad moral es el de *control* estricto no es exacta ni siquiera referida exclusivamente a la cultura occidental moderna, y aun limitada a los estratos “más educados” de ésta. Entre nuestro mundo occidental moderno, se espera que uno asuma su responsabilidad para con su familia, sus hijos, padres, ancianos, enfermos, etc.; que responda por cosas que, en general, desbordan la esfera de su control estricto.

Dónde estamos y adónde vamos

En este capítulo, he abordado la significación y consecuencias del rechazo del principio de control y he defendido que un mundo con suerte moral no es en general peor que uno sin ella —la idea contraria es fruto de una concepción equivocada de la agencia y de la moralidad.

En el capítulo siguiente me ocuparé de ciertas cuestiones metateóricas relacionadas con nuestra cuestión. Ello me dará la oportunidad de acabar de caracterizar mi posición en relación a la suerte moral.

9. Teoría, práctica y choque de intuiciones

Consideraciones metateóricas. Por un revisionismo teórico moderado

- 9.1. Intuiciones comunes y su papel en la construcción teórica
 - 9.1.1. Naturaleza de las intuiciones
 - 9.1.2. Análisis conceptual y filosofía experimental
 - 9.1.3. Algunos desiderata metodológicos
- 9.2. Intuiciones acerca de la responsabilidad moral
 - 9.2.1. La apelación a las intuiciones populares
 - 9.2.2. Experimentos: ¿somos incompatibilistas ‘naturales’?
 - 9.2.3. La cuestión de la compatibilidad y la suerte moral
- 9.3. Conflicto de intuiciones
 - 9.3.1. Tipos de conflicto
 - 9.3.2. Investigación de las fuentes del conflicto concreto / abstracto
- 9.4. Suerte moral: conflicto y revisionismo
 - 9.4.1. Suerte moral y tipos de revisionismo
 - 9.4.2. Revisionismo moderado; ¿el resultado de un “equilibrio reflexivo”?

Tengo menos confianza en cualquier teoría general...
de la que tengo en muchos juicios particulares...

Robert M. Adams, *Involuntary Sins*.

Hemos estado viendo a lo largo de esta investigación que el fenómeno de la suerte moral se planteaba en términos de choque de intuiciones. Por un lado, nos encontrábamos con la intuición de que las personas son sólo moralmente evaluables por aquello que está bajo su control —lo que responde y apoya el Principio de Control. Por el otro, resulta que de hecho las evaluamos también por cosas que escapan a su control. Ante este resultado, algún tipo de revisión, de uno u otro elemento, aparecía como inevitable.

Este último capítulo tendrá un carácter más metodológico, o de reflexión de segundo orden, en el que trataré de comparar, generalizar y sistematizar los principales resultados de los capítulos precedentes. No debe extrañar, pues, que el tratamiento directo de la cuestión de la suerte moral quede un tanto marginada en algunas secciones. En particular, empezaré refiriéndome a la teorización general acerca de la naturaleza de nuestras intuiciones y su papel en la construcción teórica. A continuación, presentaré el debate más particular en torno a las intuiciones populares sobre la responsabilidad moral. El primer objetivo será conectar el conflicto de intuiciones explicitado por la reflexión acerca de la suerte moral con este debate. Para ello, describiré un grupo de estudios experimentales que tratan de descubrir cuáles son, en concreto, las intuiciones populares acerca de la llamada Cuestión Compatibilista y trazaré un paralelismo con las intuiciones que entran en juego en la caracterización de la suerte moral. Tras diagnosticar cuáles son las intuiciones populares acerca de la suerte moral, cabrá contrastarlas con las consideraciones sustantivas de los capítulos anteriores. Esto nos llevará, finalmente, a prescribir algún tipo de revisión. Ofreceré una clasificación de las clases y grados de revisionismo y caracterizaré mi posición como un revisionismo moderado de un tipo particular. Para finalizar, trataré de dictaminar si mi solución supone un tipo de “equilibrio reflexivo”.

9.1. Intuiciones comunes y su papel en la construcción teórica

9.1.1. Naturaleza de las intuiciones

Recientemente ha adquirido una gran importancia la cuestión metafísica del papel que las intuiciones o juicios intuitivos juegan o deben jugar en la construcción de teorías filosóficas. Es cierto que no dispo-

nemos de una comprensión satisfactoria de cómo funcionan las intuiciones y, en general, de su estatus epistemológico —esto es, de la deseada correlación entre el hecho de tener la intuición de que P y que sea verdadero que P—; sin embargo, se asume mayoritariamente que, para toda teoría, apelar a *nuestras* intuiciones es una virtud, o un demérito no adecuarse a ellas.¹ Por ejemplo, una gran mayoría de la enorme bibliografía que surgió, en el campo de la epistemología, como respuesta al clásico artículo de Gettier usa intuiciones acerca de casos específicos para poner a prueba los análisis del concepto de conocimiento propuestos.² También en el debate en torno al libre albedrío y la posibilidad de la responsabilidad moral se apela a menudo a intuiciones comunes, que constituyen restricciones para las teorías propuestas. Por ejemplo, Susan Wolf afirma que su teoría sólo tendrá éxito si se acomoda adecuadamente a nuestras intuiciones al respecto.³

De esta manera, las intuiciones parecen jugar en filosofía un papel análogo al que juega la evidencia empírica en la ciencia; constituirían lo que podemos llamar *evidencia filosófica*. Sin embargo, el paralelismo es problemático. Como resulta obvio para cualquiera, la gran fiabilidad de los modos de observación e instrumentos científicos hace que la evidencia empírica goce de un prestigio del que carecen las intuiciones —incluso las intuiciones *filosóficas* (o intuiciones de los “educados filosóficamente”). En concreto, las intuiciones no pueden ser calibradas de la manera en que lo son los métodos científicos. Además, la misma naturaleza de las intuiciones es controvertida. En primer lugar, cabría que nos preguntáramos a qué nos referimos exactamente con el término *intuición*. Por ejem-

¹ Estas cuestiones están ganando gran importancia conforme la filosofía (analítica) se ha vuelto más autoconsciente de sus métodos.

² Por supuesto, me refiero a Gettier (1963).

³ Wolf (1990), p. 112.

plo, Gopnik y Schwitzgebel han caracterizado la noción de juicio intuitivo como “un juicio [...] no emitido sobre la base de ningún tipo de razonamiento explícito que una persona pueda observar conscientemente.”⁴ Pero esta caracterización es demasiado liberal, pues incluiría juicios perceptivos, introspectivos y recuerdos. Más adecuado sería decir, con George Bealer, que las intuiciones son “apariencias intelectuales que se presentan a sí mismas como necesarias y que son distintas de fenómenos como las ‘intuiciones físicas’, los experimentos mentales, las creencias, las corazonadas, los juicios de sentido común y los recuerdos.”⁵ Aun así, constituyen una clase demasiado amplia y variopinta de cogniciones. Las intuiciones parecen ser criaturas intelectualmente extrañas. Cuando tenemos una intuición (en el sentido relevante), hay algo, principalmente un contenido proposicional, que se nos aparece como cierto, que nos parece que es el caso, sin que seamos capaces de rastrear su origen en inferencias o percepciones sensoriales de que eso es así. Por supuesto, una intuición puede dar origen a una reflexión acerca de su objeto intencional, que puede acabar en una justificación de la intuición sobre la base explícita de buenas razones, a través de inferencias, percepciones, etc.; pero este resultado no forma parte de la intuición misma. Lo característico de las intuiciones es que “[c]uando se dan, frecuentemente sobresalen psicológicamente, pero sin que sus orígenes nos estén disponibles”.⁶

9.1.2. Análisis conceptual y filosofía experimental

La que podemos llamar *concepción heredada* (típica de la concepción tradicional de la filosofía como análisis conceptual) asume que las

⁴ Gopnik y Schwitzgebel (1998), p. 77.

⁵ Bealer (1998), p. 213.

⁶ Weinberg (2007), p. 318.

intuiciones constituyen puntos de partida para la investigación filosófica, al tiempo que suponen una constricción que las teorías deberían respetar. Sin embargo, esta asunción ha sido criticada sobre la base de que la aceptación sin más de las intuiciones como puntos de partida incuestionables es peligrosa. De hecho, las intuiciones pueden no ser más que convicciones internalizadas a partir de generalizaciones irreflexivas de la experiencia o de burdas analogías. No es implausible pensar que un examen detallado de sus posibles fuentes subyacentes —las teorías explícitas, las creencias cotidianas, las teorías tácitas, o los mismos conceptos de nuestro lenguaje— muestre que las intuiciones (populares y filosóficas) son artefactos que reflejan antes algún rasgo o aspecto de la persona o grupo particular, más que una genuina información acerca del tema en cuestión.⁷

En esta línea, algunos filósofos han defendido que no podemos dar por sentado el poder de las intuiciones, y mucho menos nuestra fiabilidad a la hora de detectar cuáles son estas intuiciones y su uniformidad. Si las teorías filosóficas han de ser fieles a las intuiciones comúnmente compartidas, como se afirma, la cuestión que primero habrá que establecer es la de cuáles son estas intuiciones comúnmente compartidas. Para esta empresa, los límites del reconocimiento individual, por parte de cada pensador aisladamente, son más que obvios. Por ello, los integrantes de la corriente autodenominada “filosofía experimental” han defendido que el método correcto para averiguar cuáles son estas intuiciones es la experimentación.

El método experimental utilizado consiste, en general, en realizar estudios empíricos entre sujetos sin ningún contacto previo con el estudio filosófico de los temas en cuestión (que eviten el sesgo que éste puede producir), que permita comprobar las intuiciones populares. Por ejemplo,

⁷ Gopnik y Schwitzgebel (1998), p. 258; Cummings (1998), p. 116.

en relación a los casos Gettier a que me referí más arriba, Weinberg, Nichols y Stich presentaron descripciones como la siguiente a diferentes grupos de sujetos:

Bob tiene un amigo, Jill, que ha tenido un Buick durante muchos años. Bob piensa, por ello, que Jill conduce un coche americano. No obstante, no es consciente de que ésta ha vendido recientemente su Buick, y lo ha cambiado por un Pontiac, que es un tipo diferente de coche americano. ¿Sabe realmente Bob que Jill tiene un coche americano, o sólo lo cree?⁸

Tras leer la historia, los sujetos tenían que elegir entre “Realmente Sabe” o “Sólo Cree”. En una población de estudiantes de grado de la universidad de Rutgers, sólo el 26% de los que se habían identificado como culturalmente occidentales pensaron que Bob tenía conocimiento, mientras que el 57% de los participantes originarios de Asia oriental y el 61% de los originarios de la India subcontinental se decantaron por “Realmente Sabe”. Estas diferencias en las intuiciones epistémicas mostradas por los sujetos participantes en el experimento son estadísticamente significativas y contrastan con la opinión ortodoxa en epistemología de que en los casos Gettier la intuición compartida es que Bob no sabe, sino que sólo cree. Weinberg, Nichols y Stich ven en este y otros resultados semejantes muestras de variaciones fundamentalmente culturales entre los grupos analizados.⁹ Por otro lado, también se encontraron diferencias en las intuiciones epistémicas de sujetos de diferentes estatus educativos. En particular, la mención de una posibilidad no realizada de error produjo una

⁸ Weinberg, Nichols y Stich (2001), p. 443.

⁹ Específicamente, ellos conectan estos resultados con las diferencias entre el énfasis “individualista” occidental de analizar los objetos aisladamente al contexto y el énfasis “holista” oriental en las relaciones entre los objetos y su contexto. Sin duda, esto conecta con lo dicho en la última sección del capítulo anterior, en relación a las diferentes concepciones de la responsabilidad moral.

respuesta más crítica por parte de los sujetos con estatus socioeconómico alto que por parte de los sujetos de estatus bajo.

En definitiva, el hecho de que las intuiciones varíen de un grupo a otro es un resultado sorprendente que socava la confianza de los filósofos en su capacidad de detectar aisladamente y desde su sillón cuáles son las intuiciones que comúnmente compartimos. Según Weinberg, Nichols y Stich, la epistemología actual, basada en las intuiciones de quien teoriza, sería en realidad *etnoepistemología*; esto es, un estudio de las actitudes epistémicas contingentes a la subcultura del teórico. Como resultado, afirman, la epistemología (y la filosofía, en general) que se apoya en intuiciones no puede darnos normas acerca de cómo debe pensar todo el mundo.¹⁰

Por supuesto, los resultados de experimentos como éste plantean muchas cuestiones. Por ejemplo, Ernesto Sosa —defensor del análisis conceptual tradicional— ha afirmado que no está nada claro que la variación de las respuestas al caso Gettier de Bob y Jill puedan adscribirse propiamente a variaciones culturales en relación a lo específicamente epistémico, más bien que a variaciones culturales en el modo en que los lectores rellenan el trasfondo de detalles no epistémicos de las historias presentadas.¹¹ No obstante, este tipo de estudios plantea, en general, un reto importante a la *concepción heredada*. Los autores de experimentos como el anterior afirman que el descubrimiento de que ciertas intuiciones acerca de cuestiones filosóficas, respecto a las cuales se ha dado por sentado que son uniformes, de hecho varían en relación a factores culturales y educacionales, pone en cuestión “la adecuación de las intuiciones para

¹⁰ Swan, Alexander y Weinberg (en prensa), para resultados semejantes en relación a los casos *Truetemp* de Keith Lehrer contra el fiabilismo.

¹¹ Sosa (2008).

desempeñar ningún tipo de rol evidencial.”¹² Sin embargo, el cuestionamiento sistematizado de nuestras intuiciones de sentido común no puede más que conducir al escepticismo radical. De hecho, la principal objeción a la apelación a las intuiciones, a saber, que no son fiables, se aplica igualmente a la percepción y a la introspección, a las que presumiblemente nadie quiere renunciar.¹³ En particular, para Timothy Williamson el escepticismo sobre las intuiciones, como el escepticismo sobre la percepción, nos insta a que concibamos la evidencia como un hecho acerca de nuestros mismos estados psicológicos internos; en concreto, hechos acerca de nuestras inclinaciones a emitir juicios sobre determinadas cuestiones más que hechos sobre las cuestiones mismas. Pero esta presión debe resistirse, pues descansa en una mala epistemología: una que está guiada por el ideal inalcanzable de una evidencia identificable de manera no problemática.¹⁴ Por otro lado, también se ha defendido que nuestras intuiciones no pueden estar completamente equivocadas, en tanto que son al menos parte constitutiva de los conceptos a los que se refieren.¹⁵

9.1.3. Algunos desiderata metodológicos

Creo que con esto bastará para que el lector se haga una idea del contenido del debate metafilosófico acerca del papel que las intuiciones juegan o deben jugar en la construcción de teorías filosóficas, que es reflejo y a su vez repercute en la misma concepción de la filosofía que uno tiene, y del estatus epistemológico de las mismas intuiciones. Por supuesto, con esto no pretendo haber sido exhaustivo, ni creo que se pueda llegar

¹² Alexander y Weinberg (2007), p. 63.

¹³ Véase Sosa (1998) y Bealer (1998).

¹⁴ Williamson (2004).

¹⁵ Por ejemplo, Jackson (1998).

a ninguna conclusión estricta. Pero el mero hecho de seguir la dialéctica de este debate nos permite sospechar algunas cosas.

Vimos que era presupuesto de la *concepción heredada* que las intuiciones (i) constituyen puntos de partida necesarios para la investigación filosófica y (ii) suponen una constricción que las teorías han de respetar. Los resultados experimentales deben hacernos sospechar de los límites o peligros de la apelación indiscriminada a nuestras intuiciones individuales. Sin embargo, tenemos que partir de algún lugar y las intuiciones populares son un buen punto para hacerlo en filosofía. Creo que salta a la vista que, si no queremos vernos abocados al escepticismo general, no hay más remedio que aceptar (i). ¿De dónde más podrían partir nuestras investigaciones filosóficas? Además, el descubrimiento, por medio de la experimentación, de que las intuiciones populares respecto a una cuestión filosófica determinada son diferentes a las asumidas por los filósofos, o que varían entre diferentes grupos de población, vendrá a contribuir a una mayor fiabilidad de este método más que a su bancarrota.¹⁶ En la medida en que el experimentalismo puede contribuir a desenmascarar intuiciones no fiables, es algo que todo filósofo ha de ver con buenos ojos. Así, experimentos como el anterior pueden ser aliados de la misma apelación a intuiciones.

Sin embargo, (ii) es mucho más cuestionable. En concreto, si algo que juzgamos claramente intuitivo resulta que contradice ciertos resultados teóricos, empíricos, o simplemente choca con otras intuiciones más bien sustentadas, entonces parece que no habrá más remedio que proceder a su revisión. Si nuestras intuiciones no son tan fiables, uniformes y unitarias como en la concepción ortodoxa del análisis conceptual parecía su-

¹⁶ En este sentido, contra los “experimentalistas radicales”, Liao (2007) ha defendido que los estudios presentados por éstos sugieren que de hecho algunas intuiciones son fiables.

poner, otras consideraciones, como la simplicidad, la resistencia, la coherencia lógica o la consonancia con los resultados de las ciencias naturales, pueden prevalecer, con lo que cierta antiintuitividad puede resultar necesaria en beneficio de la ganancia total en el sistema teórico construido. Esto es algo que resulta irremediable en relación al conflicto de intuiciones planteado por la suerte moral.

Creo que estas dos maneras de comprender (i) y modificar (ii) son relativamente poco controvertidas, por lo que las consideraré dos desiderata razonables. (Por supuesto, repito, que no se siguen estrictamente de la discusión anterior.) Los únicos que, en rigor, pueden oponerse a esta propuesta serán aquellos que otorguen un carácter a priori e infalible a nuestras intuiciones —especie que no debe abundar mucho hoy en día—, así como, en el extremo opuesto, los naturalistas radicales que no acepten como fuente de evidencia más que la observación empírica. Sin embargo, la apelación a intuiciones como punto de partida de la investigación filosófica no debe resultar sospechosa a los filósofos de orientación naturalista moderada, pues —repito— no se trata de que aquéllas constituyan un límite incorregible e infranqueable, sino que quizá muchas de ellas deberán ser finalmente rechazadas.

Pues bien, si resulta que las intuiciones compartidas en relación a una cuestión determinada constituyen un punto de partida necesario para la investigación filosófica de esa cuestión, entonces será importante aclarar cuáles son estas intuiciones. De este modo, el objeto primario de nuestra atención no serán, en este punto, las diferentes teorizaciones filosóficas acerca de la responsabilidad, sino las intuiciones populares acerca de la atribución de responsabilidad moral. Es por ello que en la sección siguiente analizaré un conjunto de experimentos recientes que han arrojado una gran cantidad de datos empíricos acerca de cuáles son estas intuicio-

nes. Los datos empíricos resultarán relevantes para el tema de la suerte moral.

Con esto no quiero implicar —repito— que la determinación correcta de cuáles son las intuiciones populares acerca de un problema filosófico pueda constituir la solución misma a este problema. De lo que se trata, en primer lugar, es sólo de aclarar cuáles son estas intuiciones. Esta es la tarea de lo que podemos llamar el *proyecto descriptivo*, cuya meta sería sacar a la luz las intuiciones populares sobre un dominio determinado con la esperanza de poder esbozar la teoría popular subyacente sobre ese dominio.¹⁷ Este proyecto contrasta con la empresa ulterior de determinar si las concepciones populares son correctas y si debemos preservar o revisar nuestros conceptos y prácticas a la luz de los resultados.

9.2. Intuiciones y experimentos acerca de la responsabilidad moral

9.2.1. La apelación a las intuiciones populares

Hay que comenzar diciendo que muchos de los filósofos que se han ocupado del problema del libre albedrío y la posibilidad de la responsabilidad moral no han dudado en apelar a intuiciones compartidas en relación a estas nociones, que sus diferentes teorías al respecto supuestamente recogen. Esta apelación es mucho más destacada entre los incompatibilistas, quienes suelen empezar afirmando que es una intuición comúnmente compartida que alguien sólo es verdaderamente responsable si posee libre albedrío, siendo éste incompatible con el determinismo; y uti-

¹⁷ Véase Nichols (2006); también Vargas (2005a). Nichols distingue tres proyectos: el descriptivo, el sustantivo y el prescriptivo; Vargas, dos: el diagnóstico, que trata de reflejar los hechos acerca de nuestro concepto de responsabilidad moral y sus condiciones de aplicación, y el prescriptivo, que aspira a generar una teoría correcta, que nos diga qué debemos pensar y hacer al respecto.

lizan esta intuición como apoyo de su posición. Más aún, muchos libertaristas afirman que es una idea ampliamente compartida la de que somos seres morales y que para serlo hemos de ser libres, y que, por ello, no podemos más que serlo. Por ejemplo, Robert Kane ha resaltado el “incompatibilismo natural” con que, según él, toda persona se acerca al problema del libre albedrío.¹⁸ También Laura Ekstrom ha insistido en nuestro “incompatibilismo preteórico”.¹⁹ Y Galen Strawson ha afirmado que “la intuición incompatibilista [es] un hecho natural acerca de seres cognitivos como nosotros”;²⁰ está “en nuestra naturaleza sentir que el determinismo plantea un serio problema para nuestras nociones de responsabilidad y libertad.”²¹ Asimismo, Thomas Pink considera que “la mayoría de nosotros comienza con una importante asunción sobre la libertad. Tendemos naturalmente a asumir que nuestra libertad de acción debe ser incompatible con que nuestras acciones estén determinadas”.²²

Aunque en menor medida, también los compatibilistas han apelado a intuiciones compartidas acerca de la responsabilidad moral que, en su opinión, niegan el presupuesto del indeterminismo. Por ejemplo, Daniel Dennett defiende que cuando las personas cotidianamente atribuyen responsabilidad moral “simplemente no importa en absoluto [...] si el agente en cuestión podría haber actuado de otro modo en esas circunstancias”.²³ Asimismo, los casos Frankfurt están diseñados para tratar de desencadenar la intuición de que la responsabilidad moral es compatible con el determinismo.²⁴ Finalmente, Susan Wolf motiva, en parte, su versión

¹⁸ Véase Kane (1999), p. 218.

¹⁹ Ekstrom (2002), p. 310.

²⁰ Strawson (1986), p. 88.

²¹ Strawson (1986), p. 30.

²² Pink (2004), p. 12.

²³ Dennett (1984), p. 558.

²⁴ Ver Frankfurt (1969).

del compatibilismo afirmando que “parece estar de acuerdo con, y dar cuenta de, el conjunto completo de nuestras intuiciones acerca de la responsabilidad.”²⁵

Pues bien, ante este *hecho*, una cuestión importante es la de si realmente las personas somos *naturalmente* incompatibilistas o resulta que nuestras intuiciones acerca de la libertad y la responsabilidad no presuponen el indeterminismo. Esta es una cuestión eminentemente empírica, por lo que nada mejor, para descubrir cuáles son estas intuiciones, que el estudio empírico. A continuación, describiré una serie de experimentos que se han propuesto averiguar cuáles son las intuiciones comúnmente compartidas al respecto. El método general consiste en presentar descripciones breves de agentes que actúan en escenarios deterministas y preguntar a los participantes en el experimento si éstos son o no moralmente responsables de sus acciones. Para sorpresa de muchos, los primeros resultados apuntaron en contra de que seamos incompatibilistas *preteóricos*.

9.2.2. Experimentos: ¿somos incompatibilistas ‘naturales’?

Cronológicamente, los primeros estudios al respecto fueron los de Viney y colegas, cuya conclusión fue que las intuiciones de la gente son más bien compatibilistas.²⁶ En él, los investigadores usaron un cuestionario inicial para distinguir entre sujetos que creían que el universo era determinista y aquellos que no. Todos los sujetos recibieron unas preguntas, en respuesta a las cuales podían justificar los actos de castigo. El hallazgo más importante fue que los deterministas ofrecieron justificaciones retributivistas en el mismo porcentaje que los indeterministas, lo que sugiere que los deterministas son predominantemente compatibilistas. No obstante,

²⁵ Wolf (1990), p. 89.

²⁶ Viney *et al.* (1982) y (1988).

estos experimentos han sido criticados por la medida defectuosa usada para identificar a los deterministas. La afirmación que servía para este propósito (“Creo en el determinismo”) era sólo un ítem de los siete que contribuían a la puntuación total en “la escala libre albedrío-determinista”, por lo que personas que rechazaron la afirmación determinista central podía aún terminar siendo etiquetadas como deterministas.²⁷

Sin embargo, Woolfolk, Doris y Darley han llegado recientemente a una conclusión similar usando una metodología radicalmente diferente.²⁸ En concreto, pusieron en marcha una serie de experimentos en los que los sujetos recibieron breves descripciones de agentes que tenían que actuar bajo grandes niveles de constrictión. En una de ellas, un personaje llamado Bill era capturado por unos terroristas que le obligaban a tomar la “droga de la sumisión” con el fin de inducirlo a asesinar a un amigo:

Sus efectos [de la “droga de la sumisión”] son similares al impacto de la hipnosis hábilmente administrada; el resultado es la sumisión total. Para poner a prueba los efectos de la droga, el jefe de los secuestradores ordenó a Bill que se abofeteara a sí mismo. Para su asombro, Bill observó como su propia mano derecha golpeó su mejilla izquierda, aunque no experimentó haber querido que su mano hiciera este movimiento. El jefe le entregó la pistola a Bill con una bala en la recámara y le ordenó que disparara a Frank en la cabeza...

Los investigadores manipularon el grado en el que el agente se *identificaba* con la conducta que se le había ordenado realizar. A unos sujetos se les dijo que Bill no quería matar a Frank; a otros, que Bill estaba contento de haber tenido la oportunidad de matar a Frank. El resultado mostró que los sujetos se decantaron más por considerar a Bill moralmente responsable

²⁷ Nichols (2006), pp. 11-2.

²⁸ Woolfolk, Doris y Darley (2006).

cuando se identificaba con la conducta que cuando no. Así, los participantes asignaron más responsabilidad cuando hubo mayores niveles de identificación, aun a pesar de que la conducta del agente fuera completamente forzada. Así pues, el estudio parece aportar evidencia a favor de la idea de que consideramos a los demás como moralmente responsables por su conducta, aun cuando no pudieran haber actuado de otro modo; lo que supone un refrendo experimental de los famosos casos Frankfurt. Sin embargo, los sujetos mostraron una baja asignación general de responsabilidad, lo que hace que estos resultados no supongan un reto directo a la idea de que somos incompatibilistas *naturales*. Además, estos estudios pueden ser criticados por lo mismo que se ha criticado a Harry Frankfurt; esto es, cabría aclarar si los participantes en el experimento han presupuesto que la identificación o no del agente con su acción es algo que en sí mismo está libre o no de constricción.²⁹ Los resultados serían realmente importantes si los sujetos hubieran considerado mayoritariamente que Bill es moralmente responsable por matar a Frank aun cuando se hubiera explicitado el hecho de que la misma identificación de Bill con sus actos es algo que está determinado.

Más sólidos son los resultados de la serie de estudios llevados a cabo por Nahmias, Morris, Nadelhoffer y Turner. En ellos, los sujetos participantes recibieron breves descripciones de casos de agentes que actuaron inmoralmente en un mundo determinista, y se les preguntó por la responsabilidad moral de las acciones de éstos. Por ejemplo, los sujetos leyeron esta historia:

²⁹ Para una interesante discusión y crítica de los casos Frankfurt en virtud de la idea sobre la que se basa mi objeción, véase Moya (2007), cap. 2. Sobre la noción de identificación de Frankfurt, véase Frankfurt (1969) y el capítulo 5, § 1 de esta investigación.

Imagina que en el siglo que viene descubrimos todas las leyes de la naturaleza, y construimos un superordenador que pueda deducir a partir de esas leyes de la naturaleza y del estado actual del mundo todo lo que sucederá en cualquier punto temporal del futuro. El superordenador puede verlo todo acerca de cómo es el mundo y predecirlo todo acerca de cómo será con un 100% de precisión. Supón que ese superordenador existe y observa el estado del universo en un momento determinado del 25 de marzo de 2150, veinte años antes de que Jeremy Hall naciera. A partir de esa información y de las leyes de la naturaleza, el ordenador predice [deduce] que Jeremy asaltará el Banco Fidelidad a las 6:00 PM del 26 de enero de 2195. Como siempre, la predicción del superordenador es correcta; Jeremy asalta el banco a las 6:00 PM del 26 de enero de 2195.³⁰

Los experimentadores preguntaron si Jeremy es censurable por robar el banco y el 83 % de los sujetos respondieron que sí. Además, Nahmias y colegas obtuvieron resultados similares usando otros dos escenarios diferentes.³¹ Estos resultados aportan, pues, evidencia significativa en contra de la tesis de que somos *incompatibilistas naturales*.

No obstante, Shaun Nichols y Joshua Knobe, guiados por la idea de que las encuestas informales a estudiantes suelen llegar a la conclusión contraria, se preguntaron si los resultados anteriores de Nahmias y colegas no se deberían a ciertos rasgos particulares en la confección de las historias que presentaron a los sujetos. En particular, la hipótesis inicial de Nichols y Knobe era la de que los resultados dependían de la descrip-

³⁰ Nahmias *et al.* (2005), p. 566. Los autores afirman que utilizan una noción laplaciana de determinismo. Laplace define ‘determinismo’ en términos de un ser inteligente tal que si tuviera un conocimiento completo de las leyes de la naturaleza y el estado actual del universo, podría conocer todos los sucesos pasados y predecir todos los sucesos futuros (véase la nota 15). La predictibilidad presupone, aquí, el determinismo.

³¹ En estos otros estudios trataron de focalizar más claramente la atención de los sujetos en el carácter determinista de las leyes de la naturaleza, más que en su predictibilidad. Esto respondería a la acusación de confundir ‘determinado’ con ‘predictible’.

ción en términos concretos de las acciones en cuestión, que provocaban reacciones fuertemente afectivas, y que los resultados no serían los mismos si la descripción de las acciones fuera más abstracta.³²

Para poner a prueba esta hipótesis, Nichols y Knobe presentaron a los sujetos la descripción de un Universo A, en el que los sucesos siempre tienen lugar según leyes deterministas, y un Universo B, en el que tenía cabida el indeterminismo, descritos de este modo:

La diferencia fundamental, pues, es que en el Universo A todas las decisiones de sus habitantes están completamente causadas por lo que sucedió antes de la decisión —dado el pasado, cada decisión *tiene que suceder* del modo en que lo hace. Por el contrario, en el Universo B, las decisiones no están completamente causadas por el pasado, y cada decisión humana *no tiene que suceder* del modo en que lo hace.³³

A continuación, se separó a los sujetos en dos grupos. A los del grupo de la “condición abstracta” se les preguntó si es posible que en el Universo A “una persona sea por completo moralmente responsable de sus acciones.” A los sujetos en la “condición concreta” se les facilitó la siguiente descripción:

En el Universo A, un hombre llamado Bill se ve atraído por su secretaria, y decide que el único modo de estar con ella es matar a su mujer y tres hijos. Él sabe que es imposible escapar de su casa si se produce un incendio. Antes de iniciar un viaje de negocios, monta un mecanismo en el sótano que hace que la casa arda completamente y acaba con su familia. ¿Es Bill por completo moralmente responsable de matar a su esposa e hijos?³⁴

³² Nichols y Knobe (2007).

³³ Nichols y Knobe (2007), p. 669.

³⁴ Nichols y Knobe (2007), p. 670.

Los resultados confirmaron significativamente la hipótesis inicial: el 72% de los sujetos respondieron “sí” en la condición concreta, dando con ello una respuesta compatibilista; mientras que sólo 5% respondieron “sí” en la condición abstracta.

Estos resultados me parecen muy interesantes. En primer lugar, en relación a la cuestión de si somos o no incompatibilistas *naturales*, el resultado es que no hay una respuesta unívoca a esta pregunta, sino que lo que se observa es que en ciertas condiciones las intuiciones de la gente tienden a ser compatibilistas y en otras tienden a ser incompatibilistas. Este resultado *dual* es un hallazgo importante. Pero yendo más allá de esta misma cuestión, un resultado más general es el de que tendemos a responder de manera distinta según una cuestión se plantee en términos abstractos o concretos. Este resultado es consistente con la comprensión de la suerte moral como un choque entre una intuición abstracta y una intuición concreta —lo que, a su vez, supone un apoyo indirecto de esta comprensión. Desarrollaré esta idea en el apartado siguiente.

9.2.3. La cuestión de la compatibilidad y la suerte moral

La dialéctica del apartado anterior, que —recordemos— tiene que ver con el carácter intuitivo o contraintuitivo de la Cuestión de la Compatibilidad, esto es, de la compatibilidad o incompatibilidad de la responsabilidad moral con el determinismo, podría reconstruirse en términos de la compatibilidad o incompatibilidad entre la responsabilidad moral y la falta de control, que es el tema de la suerte moral. Hemos visto que la posición por defecto en relación a la primera cuestión parece ser la de que somos incompatibilistas —incompatibilistas *preteóricos*. Ante la falta de una base sólida (experimental) para esta afirmación (empírica), algunos filósofos la han puesto a prueba experimentalmente, con el resultado de

que parece ser falsa: los experimentos muestran que la atribución de responsabilidad moral es posible aún en escenarios deterministas. Sin embargo, parece que esta conclusión viene propiciada por la manera de presentar los casos ante los que los sujetos participantes han de responder: las descripciones concretas desencadenan respuestas compatibilistas, mientras que las descripciones abstractas producen respuestas incompatibilistas.

Traduzcamos ahora esta dialéctica a nuestro tema. La posición por defecto, o *prima facie*, parece ser la de que la responsabilidad moral presupone el control, o es proporcional al grado de control, del agente. Pero, no es difícil pensar en experimentos que pusieran a prueba las intuiciones populares sobre esta cuestión que diesen resultados adversos a esta posición por defecto. Por ejemplo, consideremos un caso paradigmático de suerte resultante, el de los dos conductores borrachos, cuya descripción (neutral) puede ser esta:

Dos amigos están bebiendo unas cervezas en un bar. La euforia por la victoria de su equipo favorito provoca que tomen alguna cerveza de más. Tras el partido, ambos vuelven a casa con sus respectivos coches y, en un punto determinado del camino, ambos invaden el arcén. Pero sólo uno de ellos, Pepe, atropella a un ciclista; mientras que cuando el coche de Juan invade ese mismo trozo del arcén ningún ciclista circula por allí.

Ante esta descripción los sujetos tendrían que decir si creen que “Ambos son igualmente responsables” o que “Pepe es más responsable que Juan”. A la gran mayoría de personas (sin estudios filosóficos al respecto) a las que he planteado este caso se han decantado por la segunda opción, que va en contra del principio de control. Aunque, por otro lado, tampoco es difícil concebir un ulterior experimento, confeccionado según el patrón

del de Nichols y Knobe, que mostrase que las intuiciones a favor de la suerte moral se deben a la manera concreta de describir las situaciones, y que en condiciones abstractas la gente tiene intuiciones favorables al principio de control (contrarias, por lo tanto, a la suerte moral). En concreto, parece que proponiendo preguntas como la siguiente no sería difícil obtener respuestas contrarias a la suerte moral:

¿Crees que es justo juzgar diferentemente a dos personas (a Pepe y a Juan) cuando la diferencia entre las acciones o resultados de las acciones de ambos se deben a factores que escapan a su control?

Por supuesto, estas encuestas informales no tiene el mismo valor que la realización de los experimentos (bajo el debido procedimiento), y el resultado podría bien ser muy diferentes. Además, deberían realizarse experimentos individuales para cada tipo de suerte moral, y no es irrazonable pensar que pudiesen darse resultados diferentes para tipos diferentes de suerte moral. En todo caso, y a falta de estos experimentos, lo único que pretendo mostrar aquí es la plausibilidad del paralelismo con los resultados de los estudios empíricos acerca de la Cuestión de la Compatibilidad; en ambos casos las respuestas intuitivas varían de la misma manera según el modo de presentación de los casos.

Quiero remarcar que el paralelismo que trato de enfatizar supondría un apoyo experimental (aunque sea indirecto) a la idea de que en el caso de la suerte moral lo que fundamentalmente provoca el conflicto de intuiciones es la distinción concreto/abstracto. Por otro lado, la viabilidad de este paralelismo haría que la solución propuesta al conflicto de intuiciones que supone la suerte moral sea trasladable (pueda ser también una solución) a la Cuestión de la Compatibilidad. En la sección siguiente ana-

lizaré el fenómeno del conflicto de intuiciones, y de las salidas posibles a él.

9.3. Conflicto de intuiciones

El siguiente paso, que es todavía parte de lo que hemos llamado el *proyecto descriptivo*, consiste en el diagnóstico exacto de ante qué tipo de conflicto nos encontramos. Para ello, podemos empezar haciendo algunas distinciones generales en relación a los tipos de conflicto. A continuación, me ocuparé de las posibles respuestas, lo que nos llevará a la cuestión del revisionismo.

9.3.1. Tipos de conflicto

En general, debemos identificar dos tipos diversos de conflicto. En primer lugar, cabe distinguir entre conflictos sobre el mismo caso y conflictos sobre casos que parecen el mismo, pero en realidad son diferentes. Matthew Liao ilustra esta distinción con el contraste entre los dos casos siguientes.³⁵ Por un lado, una persona puede tener la intuición de que el aborto no es permisible porque atenta contra el futuro desarrollo y vida del feto, pero a la vez tener la intuición contraria de que el aborto es permisible sobre la base de que las mujeres deben poder decidir en relación a su cuerpo; estaríamos aquí ante un conflicto del primer tipo. Por otro lado, una persona puede también tener la intuición de que el aborto no es permisible pero que la investigación con células madre embrionarias sí que lo es. Aunque ambos temas pueden parecer el mismo, estrictamente no lo son, por lo que realmente no estaremos ante un caso de conflicto de

³⁵ Liao (2007), p. 13. Las distinciones de este y el siguiente párrafo siguen las propuestas por Liao, § 5.

intuiciones. A este respecto, parece claro que el conflicto de intuiciones entre el principio de control y la atribución de responsabilidad moral en ausencia de control (en el sentido relevante) es un conflicto acerca de lo mismo y no de algo similar pero diferente.

En segundo lugar, podemos establecer otra distinción, independiente de la anterior, entre conflictos externos y conflictos internos. Conflictos externos son los que se dan entre diferentes sujetos, como es el caso, por ejemplos, de los estudios citados más arriba de Weinberg y colegas (y Swan y colegas). En ellos, lo que entra en conflicto son las intuiciones de sujetos diferentes, que en esos casos podían agruparse en grupos étnicos. Sin embargo, los conflictos también pueden ser internos a un mismo sujeto, y es este tipo de conflicto el que está implicado en el reconocimiento de la suerte moral.³⁶

Pues bien, estamos ante un conflicto sobre el mismo tema e interno a cada uno de nosotros. Sin embargo, aún podemos preguntarnos si estamos ante un conflicto genuino, o se trata más bien de un conflicto aparente o fruto de prejuicios o sesgos cognitivos. Lo primero que cabe descartar es que estemos ante un mero desacuerdo verbal. La verdad es que si en este punto de la investigación el lector todavía creyese que estamos ante una cuestión meramente verbal me sentiría verdaderamente frustrado. Más interesante será averiguar si el conflicto de intuiciones se debe a prejuicios o sesgos.³⁷ Para ello habrá que averiguar a qué mecanismos psicológicos se deben las diferentes intuiciones en conflicto. En realidad, abordar esta cuestión es ya avanzar hacia el *proyecto sustantivo*,

³⁶ Por supuesto, podría haber diferencias entre individuos y grupos de individuos acerca del reconocimiento o no de la suerte moral. En especial, podría haber diferencias significativas entre grupos culturales y socioeconómicos diversos. Creo que esto es algo que merecería ser investigado empíricamente.

³⁷ Véase Liao (2007) y Sinnott-Armstrong (2006) para interesantes discusiones de diferentes prejuicios que pueden actuar en el conflicto de intuiciones.

esto es, hacia la determinación de qué intuiciones populares son correctas y cuáles incorrectas. Si resulta que unas intuiciones son fruto de mecanismos psicológicos racionalmente sesgados o poco fiables, se convertirán en obvias candidatas a la revisión.

De hecho, los mismos Nichols y Knobe han sugerido una explicación de los mecanismos psicológicos que generaban la variación de las respuestas de los sujetos ante la dualidad concreto/abstracto, que han querido testar experimentalmente. Cabe adelantar que su hipótesis principal es que las respuestas compatibilistas son fruto de las fuertes emociones despertadas en los sujetos por las descripciones concretas de los casos; mientras que las descripciones abstractas no provocan emociones fuertes, permitiendo así razonamientos teóricos que conducen a respuestas incompatibilistas.

9.3.2. Investigación de las fuentes del conflicto concreto/abstracto

Nichols y Knobe han propuesto diversos modelos explicativos, que han tratado de poner a prueba mediante la realización de un experimento. Entre estos modelos, ellos mismos destacan los dos siguientes.³⁸

Según el *modelo del error afectivo*, las reacciones afectivas fuertes suelen sesgar y distorsionar los juicios de la gente. Cotidianamente emitimos juicios de responsabilidad sobre la base de una teoría tácita, pero cuando afrontamos una violación flagrante de las normas morales —lo cual es especialmente destacado en las descripciones concretas de los casos— experimentamos una reacción afectiva fuerte que nos impide aplicar la teoría tácita correctamente.³⁹ Por el contrario, el *modelo de la com-*

³⁸ Nichols y Knobe (2007), § 5.

³⁹ No obstante, como los mismos autores resaltan, aceptar este modelo no implicar rechazar que el afecto pueda ser una parte de la competencia fundamental subyacente a los

petencia afectiva —principal rival del anterior—, en vez de suponer que el afecto sirve sólo para sesgar o distorsionar nuestros juicios teóricos, destaca que las reacciones afectivas constituyen el núcleo del proceso por el cual las personas asignamos responsabilidad. Así, y en resumen, si para el primer modelo las reacciones afectivas interferirían en la operación normal de juicio, para el segundo el proceso está primariamente gobernado por el afecto.⁴⁰

Para poner a prueba estas hipótesis, Nichols y Knobe idearon una segunda fase para su experimento anterior. De nuevo, todos los sujetos recibieron la descripción inicial de dos universos, A y B, y fueron asignados arbitrariamente al grupo *afecto débil* o al grupo *afecto fuerte*. Los del grupo de *afecto fuerte* tuvieron que responder a esta pregunta:

Como ha hecho muchas veces en el pasado, Bill acecha y viola a una extraña. ¿Es posible que Bill sea moralmente responsable por violar a la extraña?

A los sujetos del grupo de *afecto débil* se les preguntó:

Como ha hecho muchas veces en el pasado, Mark lo arregla todo para defraudar al erario público. ¿Es posible que Mark sea por completo moralmente responsable por defraudar al erario público?

juicios de responsabilidad moral. Es una opción que el afecto sea tanto esto como un factor que a veces puede producir errores performativos.

⁴⁰ Ambos modelos tienen su apoyo científico. A favor del modelo del error afectivo se puede recordar la gran cantidad de resultados que en psicología social indican que la influencia de sesgos afectivos o motivacionales suelen hacer disminuir la probabilidad de que el agente reconozca determinados tipos de información relevantes, forme apropiadamente sus creencias o use recursos críticos. Por su parte, el segundo modelo concuerda con estudios recientes de gente con deficiencias en el procesamiento de las emociones debidas a enfermedades psicológicas y amputaciones cerebrales.

Adicionalmente, para la mitad de los integrantes de cada grupo la pregunta estipulaba que el agente se hallaba en el Universo A; para la otra mitad, en el Universo B.

El resultado fue que, incluso entre escenarios concretos, hubo una clara diferencia entre casos de afecto fuerte y casos de afecto débil. En particular, fue mucho más probable que los sujetos dieran respuestas incompatibilistas en los casos de afecto débil que en los de afecto alto, siendo siempre el universo determinista. Mucha gente dijo que no es posible que el defraudador de impuestos sea totalmente responsable, mientras que una gran mayoría de personas dijo que era posible que el violador fuera moralmente responsable.⁴¹ El resultado es que el afecto parece jugar un papel importante en el proceso que genera las intuiciones compatibilistas. Trasladándolo a nuestro tema, podemos predecir igualmente el impacto del afecto en las intuiciones acerca de la atribución de responsabilidad moral en casos paradigmáticos de suerte moral: las respuestas serán más contrarias a la suerte moral cuando la reacción afectiva generada por la comprensión de la situación sea más débil, que cuando sea más fuerte.

Sin embargo, esta segunda fase del experimento me parece defectuosa. En efecto, en un caso la reacción emocional es más fuerte que en el otro, pero lo que últimamente produce la distinción es más bien la mayor gravedad moral de una conducta sobre la otra. Lo que más pesa en el caso es la oposición mayor/menor gravedad de la trasgresión moral y el daño cometido. En otras palabras, es el hecho de que el agente lleve a cabo una violación, en contraposición a un fraude, lo que —ciertamente a través del

⁴¹ Por el contrario, los sujetos que respondieron a casos de agentes en universos indeterministas dijeron en su mayoría que es posible para el agente ser moralmente responsable, con independencia de si defraudó o violó. Los resultados globales fueron estos. Agentes en un universo indeterminista: afecto fuerte: 95%; afecto débil: 89%. Agentes en un universo determinista: afecto fuerte: 64%; afecto débil: 23%. Nichols y Knobe (2007), p. 266.

desencadenamiento de una reacción emocional más fuerte— repercute en una condena más contundente y en un juicio compatibilista. Y es esta diferencia en la gravedad moral lo que la descripción concreta pone en primer plano y que puesto en términos abstractos pasa a un segundo plano. De hecho, el experimento, tal y como fue concebido y conducido, no puede discriminar si lo relevante es la gravedad o la emoción. Una manera más adecuada de poner a prueba si lo que en último término produce la diferencia es realmente la fuerza de las emociones experimentadas, sería describir el mismo caso —con la misma gravedad moral— de dos maneras distintas; una, más general y la otra, más detallada.

Por otro lado, tampoco me resulta satisfactoria su explicación de los mecanismos causantes de estos resultados. Según Nichols y Knobe, el *modelo del error afectivo* proporciona una mejor explicación de los resultados experimentales, del efecto concreto/abstracto. Es obvio que las emociones muy fuertes pueden provocar distorsiones en nuestro juicio, pero no es necesario que lo hagan. En realidad, lo que vemos en el caso del defraudador de impuestos es que, cuando la influencia del afecto se minimiza, la gente da respuestas dramáticamente diferentes dependiendo de si el agente está en un universo determinista o indeterminista. Estas respuestas revelarían, según este modelo, la genuina competencia en la atribución de responsabilidad, pero cuando el afecto se maximiza, como en el caso del violador, las respuestas se igualan, debido a que las emociones sesgan los juicios. Sin embargo, si mi objeción anterior era correcta, parece que lo que en último término marca la diferencia es la gravedad moral del daño inflingido, que provoca emociones reactivas más o menos fuertes y con ellas hace más o menos apremiantes el juicio moral adverso.

De hecho, ya vimos en el capítulo anterior que un efecto de este tipo ha sido empíricamente corroborado. Recuérdese el estudio de S.

Walster, contundentemente replicado con posterioridad, que mostraba que es más probable que la gente considere responsables a los agentes que causan un daño severo que a los que causan un daño ligero, incluso si la acción del agente se describe en ambos casos igualmente negligente.⁴² Entre otras cosas, los datos muestran, sin la menor duda, el papel fundamental que la severidad del daño desempeña en la atribución de responsabilidad moral —en este caso, afectando a la exigencia con la que determinamos el umbral de la negligencia. De nuevo, se podría afirmar que este efecto responde a un sesgo o prejuicio. No obstante, como ya dije en el capítulo anterior, es mucho más convincente la idea de que responde a la importancia capital de las necesidades de la víctima. Es razonable que la gente sea menos exigente en sus condiciones para atribuir responsabilidad en los casos en los que el daño es severo que en los que el daño es leve, pues en los casos más severos la necesidad de respuesta y restitución de la víctima es mucho mayor.

Para concluir, la pretendida demostración experimental de que el *modelo del error afectivo* explica más adecuadamente la dualidad de respuestas concreto/abstracto que otros modelos alternativos, en especial el *modelo de la competencia afectiva*, resulta fallida. A falta de otro tipo de pruebas experimentales, cabe concluir que la divergencia entre intuiciones concretas e intuiciones abstractas, a que remite la diferencia entre las respuestas a enunciados concretos y a enunciados abstractos, no puede ser decidida a favor de las intuiciones abstractas, en virtud del hecho de que las respuestas a enunciados concretos se deban a un mecanismo sesgado. Así pues, el conflicto entre estas dos diferentes clases de intuiciones se resiste a ser resuelto de manera meramente explicativa o descriptiva. Esto, sin duda, refuerza la posición de Thomas Nagel, para quien nos hallamos

⁴² Walster (1966). Véase la sección 8.2.1 del capítulo anterior.

ante un fenómeno paradójico, cuyo carácter conflictivo es irresoluble. No obstante, queda una salida: optar por algún tipo de revisión.

9.4. Suerte moral: conflicto y revisionismo

Ante el resultado anterior, esto es, el conflicto interno genuino de intuiciones que supone la suerte moral, cabe investigar qué intuiciones son las más poderosas o, a la inversa, cuáles las más fácilmente revisables. En este punto, la meta es aportar argumentos filosóficos que permitan resolver el conflicto, bien elaborando una teoría positiva en relación a por qué deberían incluirse ciertas intuiciones, bien aportando una teoría negativa, del error, en relación a por qué otras intuiciones rivales deberían ser excluidas. Esto no es inusual en filosofía. Por ejemplo, Nelson Goodman subrayó (respecto a las paradojas de la confirmación) que un choque de intuiciones o convicciones preteóricas justifica al teórico a deshacerse de algunos de estos elementos y construir una teoría general (para un área determinada, en su caso la de la confirmación), compatible con los desiderata teóricos y explicativos, que salve tantas de nuestras convicciones originales como sea posible.⁴³ En particular, en los últimos años ha adquirido una gran popularidad un método que participa de este espíritu: el llamado “equilibrio reflexivo”.

En esta última sección, acabaré de caracterizar mi solución al conflicto de intuiciones que nos presenta la suerte moral. Para ello, describiré las diferentes clases de revisionismo disponibles y explicitaré la opción que se ha favorecido en esta investigación. Daré algunas razones de la, a

⁴³ Goodman (1965), p. 68. Slote (1994) defiende explícitamente esta metodología para la construcción teórica de la ética, a la luz de las inconsistencias y paradojas de nuestro pensamiento cotidiano sobre la moralidad.

mi entender, mayor moderación y plausibilidad de mi posición frente a las demás. Finalmente, explicaré en qué sentido mi solución a la cuestión puede o no interpretarse como un ajuste mutuo de intuiciones en la línea de un proceso de *equilibrio reflexivo*.

9.4.1. Suerte moral y tipos de revisionismo

Hemos visto a lo largo de esta investigación que la suerte moral se presentaba como un choque de intuiciones entre (i), intuiciones particulares o concretas acerca de las atribuciones de responsabilidad moral, y (ii) una intuición general o abstracta (el principio de control).⁴⁴ Esto hace suponer que toda solución posible implicará algún tipo de *revisionismo* de (i) o (ii). De hecho, yo he venido ofreciendo diferentes argumentos a favor de mantener (i) y rechazar o revisar (ii). Sin embargo, la revisión puede ser de diversos tipos y grados; por lo que resultará útil, en este punto, clarificar los posibles tipos de revisionismo.

Por un lado, hay que distinguir entre la clase de revisión referida a las prácticas mismas de juicio o atribución de responsabilidad moral (en adelante, *revisionismo de las prácticas*) y la referida a nuestro concepto o teorías de cómo funciona la atribución de responsabilidad (*revisionismo teórico*). Por otro lado, podemos distinguir al menos tres grados de revisionismo. En primer lugar, el *revisionismo radical* impulsa una revisión radical de uno de los elementos anteriores del sistema de la responsabilidad moral; constituye una *teoría del error* acerca de aquello que pretende revisar. En el extremo opuesto, el *revisionismo débil* es un revisionismo acerca de *lo que las personas piensan que piensan*; es la idea de que la gente no se da cuenta de la naturaleza de sus verdaderos compromisos

⁴⁴ Una ulterior opción era la defendida por Nagel, a saber, que nos encontramos ante un fenómeno paradójico, cuyo carácter conflictivo es irresoluble.

prácticos y conceptuales sobre la responsabilidad moral. Por último, a medio camino entre los otros dos tipos, el *revisionismo moderado* es un revisionismo sobre *lo que la gente piensa*.⁴⁵

Así, entre las opciones teóricas que hemos explorado y rechazado en los capítulos anteriores había quien simplemente pretendía alguna clase de revisión débil, como en el caso de los defensores del Argumento Epistémico, cuya idea era reinterpretar las prácticas cotidianas de juicio mediante una serie de distinciones que salvaban tanto las prácticas como el principio de control. Pero también vimos una manera más radical de defender el principio de control. En concreto, Michael Zimmerman defendía una teoría del error acerca de nuestras prácticas de asignación de responsabilidad moral. Según la clasificación propuesta, se trataría de un *revisionismo radical de las prácticas*. Por otro lado, había también cabida para un tipo de revisionismo práctico desde posiciones favorables a la suerte moral. En particular, B. Browne proponía disociar las actitudes reactivas del castigo; o Michael Slote llegaba a proponer una revisión radical de la moralidad que eliminara la práctica de la censura moral. No obstante, la respuesta más natural de quienes reconocemos el fenómeno es pedir alguna clase de *revisionismo teórico*, consistente en la tesis de que al menos algunas de nuestras creencias e intuiciones sobre la responsabilidad moral están equivocadas y deben ser revisadas si queremos que reflejen nuestras prácticas reales. Dentro de esta categoría cabría incluir a Bernard Williams, Michael Moore, Margaret Walker, Robert Adams y a mí mismo.

⁴⁵ Quien más ha desarrollado el tema del revisionismo en relación a la responsabilidad moral es Manuel Vargas; véase, en especial, Vargas (2005a). Vargas también ha defendido el *revisionismo* como una posición particular que pretende dar respuesta al problema del libre albedrío y la responsabilidad moral; véase, por ejemplo, Vargas (2004) y (2007).

Sin embargo, la caracterización clásica (o nageliana) de la suerte moral en términos de un conflicto entre teoría y prácticas, siendo correcta y útil, puede ser precisada. De hecho, ya al final del capítulo primero —§ 1.4— presenté preliminarmente una distinción entre no dos, sino tres niveles de elementos que conforman el *sistema de la responsabilidad moral* y que son susceptibles de entrar en conflicto. Recuperémosla.

En primer lugar, tenemos el nivel de las *prácticas* de atribución de responsabilidad moral, que incluyen los juicios de responsabilidad moral, las disposiciones o actitudes psicológicas asociadas (“actitudes reactivas”), así como ulteriores usos de estas atribuciones (premios, castigos, etc.). En segundo lugar, está el nivel de la *teoría tácita*, que regula las prácticas anteriores y suele incluir intuiciones ordinarias o de sentido común, pero que no siempre es transparente a nuestras creencias explícitas.⁴⁶ Finalmente, el nivel de las creencias abstractas y los principios que nos parecen intuitivamente autoevidentes, que consideramos que regulan las prácticas anteriores.

Esquemáticamente, la propuesta quedaría así:

- i) *Prácticas* de atribución de responsabilidad: los juicios prácticos y las disposiciones o actitudes psicológicas características (“actitudes reactivas”).
- ii) *Teoría tácita* reguladora de estas prácticas.
- iii) *Creencias y principios abstractos* sobre la atribución de responsabilidad moral.

⁴⁶ En la expresión “teoría tácita”, procedente de la ciencia cognitiva, la palabra *teoría* no tiene el sentido fuerte que tiene en la expresión, que viene a continuación, de “concepciones teóricas”. A grandes rasgos, una teoría en este sentido es cualquier representación interna de un cuerpo de información.

Cabe subrayar que entre estos tres niveles tiene lugar una compleja dialéctica. En general, pensamos que nuestras atribuciones de responsabilidad moral satisfacen diversas condiciones o descripciones generales, además de una serie de intuiciones que parecen funcionar como guía. Por ejemplo, parece que las personas se ven generalmente como responsables de aquello que hacen que es resultado de sus rasgos de carácter o tendencias dentro de un espectro normal y con causas normales. Pero sucede que a menudo encontramos desacuerdos destacados, como respecto a si somos o no responsables de aquello que hacemos como resultado, en parte, de serias privaciones. Es fundamental darse cuenta del hecho de que estas intuiciones entran a menudo en conflicto, y que en casos particulares diversos divergimos en cuanto a considerar si ciertas condiciones generales son o no sostenibles. Además, entre el nivel de la teoría tácita y el de las concepciones teóricas se da un proceso de retroalimentación claro. Quizá, cabría distinguir aún dos niveles en relación a las creencias: las de tipo *general*, como ‘es necesario que uno controle, en alguna medida, su acción para que sea responsable por ella’, que suelen ser más solidarias de las concepciones teóricas; y las creencias *particulares*, es decir, creencias o intuiciones que tenemos ante casos particulares, que están más estrechamente conectadas con las prácticas cotidianas de juicio moral, y que —como hemos podido observar a lo largo de esta investigación— a menudo chocan con las generales.

Con esta triple distinción a la vista, cabe añadir un *revisionismo* de la teoría tácita. Así, mi propuesta —que centra la cuestión en la relación entre (ii), esto es, la teoría tácita, y (iii), las concepciones teóricas— puede ahora ser adecuadamente caracterizada como un revisionismo moderado referido directamente a las concepciones teóricas y, en tanto que éstas influyen en la teoría tácita, indirectamente a nuestras intuiciones

generales. He tratado de mostrar que ni las prácticas de juicio moral, ni las intuiciones y creencias populares referidas a casos concretos (en tanto que libres o no del todo permeables a la influencia directa de concepciones teóricas sobre las atribuciones de responsabilidad moral) necesitan ser revisadas. Lo que tiene que ser *reinterpretado* son nuestras creencias más abstractas e influidas por la teoría.

Creo que esta propuesta es ventajosa en tanto que es mucho más asimilable que las clases de revisionismo que proponen las alternativas previas.

9.4.2. Revisionismo moderado; ¿el resultado de un ‘equilibrio reflexivo’?

Son dos, las razones principales a favor de la mayor plausibilidad práctica de mi propuesta sobre las demás.

La primera de ellas es el carácter moderado de esta posición. Toda propuesta de revisión radical tiene no sólo que vencer en el debate, sino además convencer. Debe mostrar, de sobra, que la revisión propuesta es realmente necesaria y ha de estar avalada por un plan de viabilidad. Una propuesta de revisión moderada no tiene que cumplir requisitos tan exigentes. Recordemos que el revisionismo moderado propone modificar *lo que la gente piensa*. Esto es, dado que parece que la gente piensa que el control es una condición necesaria para la atribución de responsabilidad moral, mi posición propone modificar esta creencia general —en concreto a favor de la idea de jurisdicción que defendí en el capítulo 8. En todo caso, la sustitución del principio de control por la idea de jurisdicción no parece muy traumática. En realidad, esta propuesta, siendo moderada, está más cerca del revisionismo débil que del radical; esto es, de ser un revisionismo acerca de *lo que las personas piensan que piensan*. De hecho,

ha sido parte de mi tesis que es la idea de jurisdicción la que realmente guía nuestras atribuciones de responsabilidad moral, aunque la gente puede que no se dé cuenta de esto y crea que sus prácticas de juicio se comprometen, en cambio, con el principio de control. De este modo, mi propuesta de revisión teórica es también más débil que otras propuestas favorables a la suerte moral.

La otra razón que, a mi juicio, juega a favor de esta posición es que la revisión propuesta se dirige a los que la gente piensa acerca de sus prácticas y no a sus prácticas mismas. Si bien, como dije en el capítulo 3, el carácter recalcitrante de nuestras intuiciones y prácticas supone un fuerte obstáculo para cualquier tipo de teoría revisionista, la revisión de las prácticas sociales siempre es más difícil que la revisión de creencias generales con una gran carga teórica. Además, la pervivencia de éstas prácticas es un hecho que juega a su favor. Las propuestas de innovaciones en las prácticas, sobre todo si son radicales —como, por ejemplo, la de Zimmerman y la de Browne y Slote—, siempre comportan el peligro de que puedan resultar en un desastre con graves consecuencias adversas. Esto siempre y cuando sean posibles. Recuérdese que el capítulo 3 criticó la propuesta revisionista de Zimmerman por ser conceptualmente incoherente; y, en el capítulo 8, hice lo propio con la de Browne y Slote, en virtud de su implausibilidad (¿irrealizabilidad?) psicológica.⁴⁷

⁴⁷ Estos dos tipos de crítica parecen ir en la línea de los dos límites al revisionismo propuestos por Manuel Vargas: la *adecuación normativa*, que prohíbe innovaciones que no estén “justificadas y bien integradas en la red de normas y prácticas que se apoyan mutuamente”; y la *plausibilidad naturalista*, que prohíbe innovaciones “implausibles bajo una concepción amplia del naturalismo sustantivo”. Véase Vargas (2004), pp. 230-3.

Para finalizar, me gustaría responder a la pregunta de si, a fin de cuentas, no he llegado a esta posición mediante un proceso de *equilibrio reflexivo*.⁴⁸

El método del *equilibrio reflexivo*, popularizado por John Rawls, procede avanzando y retrocediendo entre nuestros juicios intuitivos acerca de casos particulares, principios o reglas que creemos que los gobiernan y consideraciones teóricas que creemos que sostenemos al aceptar esos juicios considerados, principios o reglas, revisando cualquiera de esos elementos cuando sea necesario, de modo que se produzca un reajuste mutuo del sistema que, finalmente, produzca una coherencia aceptable entre todos sus elementos.⁴⁹ De hecho, se trata de un método mucho más extendido de lo que puede parecer, pues usualmente corregimos nuestras intuiciones acerca de casos particulares haciéndolas más coherentes con nuestros principios generales considerados y corregimos nuestros principios generales haciéndolos más coherentes con nuestros juicios acerca de casos particulares.

De un modo más técnico, se dice que el método parte de unas convicciones antecedentes de tipo diverso (creencias, juicios e intuiciones), y mediante (i) la construcción de teorías que sistematizan y explican esas convicciones y (ii) reflexionando y resolviendo los diversos conflictos que emergen, uno llega finalmente a una concepción sistemática idealmente coherente sobre la cuestión. Supuestamente, todas las convicciones relevantes han de ser consideradas.

El método es conservador en tanto que empezamos con nuestras concepciones presentes e intentamos hacer los menores cambios que

⁴⁸ Agradezco a Josep Corbí la sugerencia de indagar si el método del equilibrio reflexivo pudiera dar solución al conflicto de intuiciones planteado por la suerte moral.

⁴⁹ Véase Rawls (1971). Asimismo, Daniels (2003) ofrece una exposición breve y clara de los rasgos básicos del método, que he tenido en cuenta para mi caracterización.

promuevan la coherencia de nuestra concepción global. Algunas de nuestras intuiciones, que Rawls llama “puntos fijos”, pueden ser más firmes que otras; pero ni hay puntos fijos *a priori*, sino que todas las intuiciones o convicciones iniciales permanecen revisables.⁵⁰

En este sentido, el método del equilibrio reflexivo encaja bastante fielmente con el espíritu de mi propuesta, que se caracteriza por ser lo más conservadora posible, dentro de lo inevitable de tener que optar por algún tipo de revisionismo. Sin embargo, en tanto que el reconocimiento de la suerte moral supone necesariamente el rechazo del principio de control, el resultado no puede considerarse, en sentido estricto, como el resultado de un reajuste mutuo. En nuestro tema, el choque frontal entre principio y prácticas, nos obliga a optar por uno u otro extremo. Es decir, si tenemos A o B y optamos simplemente por A (o por B), no puede decirse que se haya negociado, o reajustado mutuamente, nada. En concreto, en relación al principio de control, no hay negociación posible; o se acepta o se rechaza. La idea de *jurisdicción*, con la que intenté suplir el hueco dejado por el rechazo del principio de control, no puede considerarse, en sentido estricto, una revisión o modificación de éste último, en tanto que es compatible con el reconocimiento de que la suerte pueda marcar diferencias en el juicio moral. (Recuérdese que ‘suerte’ se definió precisamente como ‘falta de control’ y que del Principio de Control se seguía el Corolario que impedía juzgar diferentemente a las parejas de contrapartes de los ejemplos de suerte moral.⁵¹)

⁵⁰ Por supuesto, en la línea de lo visto en la sección 9.1.2, algunos críticos han acusado este método de dar demasiado peso a nuestras intuiciones iniciales.

⁵¹ Véanse, sobre esto, las secciones 2.2.1, 4.3. y 8.3.1 de esta investigación. Por supuesto, el Principio de Control podría ser mantenido desvinculándolo del Corolario e incluso de la definición de la suerte como falta de control, pero entonces quedaría significativamente desfigurado, hasta el punto de ya no ser la misma noción. En concreto, dejaría de

No obstante, una vez sustituido el principio de control por la idea de jurisdicción (que no es por definición incompatible con la suerte y de la que no se sigue el corolario anterior), el proceso de equilibrio reflexivo, de reajuste mutuo, entre ésta y nuestras prácticas cotidianas de atribución de responsabilidad, queda abierto.

En ningún caso pretendo afirmar que estas prácticas, en las que estamos todos inmersos, no sean susceptibles de revisión y cambio. De hecho, es razonable pensar que muchos de nuestros juicios morales cotidianos serán erróneos, e incluso injustos; razón por la cual la actitud prudente e irónica, de la que hablé en el capítulo anterior, es muy recomendable. Pero no lo serán por las razones paradigmáticamente esgrimidas por el negador de la suerte moral.

Dónde hemos llegado

En este último capítulo, de carácter principalmente metodológico, he tratado de sistematizar algunos de los resultados de los capítulos precedentes. Tras presentar el debate acerca de a las intuiciones populares sobre la responsabilidad moral, he destacado la conexión entre el conflicto de intuiciones explicitado por los estudios experimentales acerca de la Cuestión Compatibilista y el conflicto de intuiciones con que el planteamiento típico de la suerte moral nos confronta; lo que refuerza la fuerza de este último. He rechazado que la causa de este fenómeno sean ciertos mecanismos psicológicos sesgados que originan uno de los tipos de intuiciones en conflicto. Posteriormente, he ofrecido una clasificación de los

ser el Principio de Control necesario para motivar la cuestión de la suerte moral. Esta nueva noción, que sustituiría al control, es la que he llamado *Jurisdicción*.

tipos y grados de revisionismo y he caracterizado y defendido la superioridad de mi opción revisionista moderada, en virtud de dos criterios razonables. Para finalizar, he rechazado que la solución al choque de intuiciones planteado por la suerte moral pueda constituir, en lo fundamental, el resultado de un proceso de equilibrio reflexivo.

Conclusions

Throughout this dissertation I have been arguing for the reality of moral luck. In particular, my central aim has been to rebut the deniers' arguments against moral luck. I would like to finish by briefly summarizing the main conclusions reached.

Having devoted Part I to characterizing the phenomenon of moral luck and the map for the discussion, in part II I addressed directly the most important groups of responses. Particularly, in Chapter 3 I tried to reject what was called the Global Case Against Moral Luck—i.e., the case against moral luck in general. The remaining chapters of this part (4–7) were devoted to each of the four previously distinguished kinds of moral luck: constitutive, formative, circumstantial and resultant. In each chapter the same methodology was followed: I began by reconstructing the best arguments against the particular kind of moral luck at stake and I then proceeded to put forward a rebuttal of them. My overall argumentative strategy combined a general rejection of the Global Case with a particular positive defense of every kind of moral luck.

My articulation of the Global Case against Moral Luck distinguished two main strategies. On the one hand, the Moderate Strategy (MS) vindicates the Control Principle by means of the Epistemic Argument (EA) and is committed to the Conservative Claim—that we can acknowledge the error EA unmasks and also maintain our practices because our ordinary judgments depend on the available evidence. Additionally, MS defenders commit themselves to counterfactual moderate claims

about moral judgment like this: we judge people for the way that it is *plausible for us to think* they would have acted, provided that they had had the chance. Then, the locus of judgment is the agent's actual character, dispositions or intentions. Regarding this strategy, I showed that it is insufficient as a global case, as long as it is fundamentally restricted to resultant and circumstantial moral luck. By its very nature, MS cannot be extended to antecedent kinds of moral luck, failing then to become a real Global Case against moral luck—its extension would involve renouncing to the moderate commitments.

This led us to consider the Radical Strategy (RS). It is possible to pursue “the implications of the denial of the relevance of luck to moral responsibility” to their “logical conclusion”. RS starts from the idea that the Control Principle is the only guide to determine moral desert. Its defenders have no problem in claiming that the agent's actual intentions, will, character or dispositions cannot be the basis for attributing responsibility, because they depend, at least partially, on factors beyond the agent's control. So, RS defenders appeal to the idea of basing (true) desert on those dispositions that the agent would have had given her possible counterfactual histories, i.e. dispositions that she would have had if her factual history had been different, among the whole of her life's possible histories. Of course, RS gets rid of every commitment (to actual intentions, the “plausible” restriction, etc., characteristic of MS), beyond the control principle itself. In sum, RS is a more ambitious and comprehensive enterprise than MS, but that enterprise is achieved by losing its intuitive character and possibilities of application.

Regarding RS, I posited two objections that affect the applicability of this proposal. Firstly, by splitting up an agent's actual moral record from her true desert, a big gap emerges between the actual assessments

we do and people's true deserts. This is, no doubt, an undesirable consequence. However, the denier might acknowledge that it is really difficult to make a judgment about true desert or essential moral worth, but insist that this does not imply radical scepticism about true desert. Secondly, this proposal would imply an implausible radical revisionism, which would actually be impossible to carry out. How can we revise (change) our actual judgments on the base of a multiplicity of essentially counterfactual judgments? On the other hand, this idea involves a universalization of positive and negative attributions of moral responsibility that threatens to dissolve the very idea of moral responsibility; in the end, it would eliminate distinctions in personal desert.

However, worries about applicability were not my main objection against the denier. My main contention was that the notion of (essential) true desert turns out to be impossible, not only to know, but also to fix, even in ideal conditions. By making the move of going backwards and relying on the dispositions the agent would have had given her *counterfactual possible histories*, the proponent of an essential true desert takes progressive steps backwards that ultimately reduce the agent's identity to nothing, to a bare self with no properties. Pursued to its logical conclusions, the denier's view, which rests on the idea that what ultimately matters is *only what exclusively depends on the agent*, becomes meaningless, since it happens that finally nothing *exclusively* depends on the agent. No character or identity remains to be assessed as the agent's deep self—the putative locus of her true desert. In other words, there is finally no agent on whom anything might depend. Therefore, the very idea of an *essential true desert*—as actual-enactment independent (indeed radically independent of the actual world)—becomes, at the end of the day, unintelligible.

Neither MS nor RS accomplished their intended role of constituting a Global Case against moral luck. This rebuttal of the Global Case does not involve that all kinds of moral luck are, *ipso facto*, vindicated; but only that to find a single general strategy is not possible. Nevertheless, it was still possible to reject the existence of moral luck (in all its kinds) by means of a mixed strategy, consisting of a combination of RS with a particular argument against constitutive moral luck.

Two key arguments against constitutive moral luck were laid out. The first one denied the coherence of the very notion of constitutive luck. The very constitution of an agent, this argument goes, cannot be luck for anybody, since it is required somebody or something previous on *whom* or *which* luck rests; so, the notion of factors under or beyond one's control cannot be applied to received personality or character. In this sort of objection the very notion of luck plays a fundamental role. However, I reached the conclusion that it is undeniable that the notion of constitutive luck, which does not presuppose that agents might have had a different essential constitution, is far from being incoherent. Moreover, one's character, so long as it is definitory of one's moral identity, plays an outstanding role in our moral appraisal of others, even though this does not exclude the relevance of the relationship of an agent with his own character and degree of control over it.

The second argument rested upon the idea of self-formation of one's character. It is true that we have different sets of traits that we simply receive or inherit, that is, we simply find within us; nevertheless—the argument goes—we do not judge and are not judged for this, but for our *formed character*, in which our intervention is crucial. Consequently, in Chapter 5 I explored the issue of the agent's power in forming and reforming her own moral character within her life's circumstances, which

leads each person to build herself (to some extent). This self-formative power could serve—it was the denier’s hope—to achieve the longed-for neutralization of constitutive inequalities. So, I examined the rational and the historical projects of overcoming moral luck, and I concluded that neither achieves the aim of neutralizing constitutive and, in general, antecedent moral luck. A more general conclusion was this: the aim of fully eliminating the uncontrollable factors involved in personal constitution and circumstances of formation and development from the sphere of moral responsibility is just a vain wish, doomed to failure. It is not risky to say that this is simply a fact about how we acquire our condition of morally responsible agents in the real world. Luck is not only an inescapable factor in our constitution and formation, it is a necessary factor.

At that point, it remained a move opened to the denier: to concede that moral luck cannot be globally rejected—accepting that constitutive and formative luck are inescapable—but to hold nonetheless that circumstantial and resultant luck should be rejected; that is, to defend a *hybrid strategy*. In fact, most of the participants in the debate have focused on the former, while putting the latter kinds aside or even completely ignoring them.

Consequently, Chapter 6 addressed the denial of circumstantial moral luck, i.e. the claim according to which the same intentions deserve the same moral judgment, independently of their leading or not to actual actions. One might understand the intentions alleged to be the only locus for moral responsibility as either the product of the agent’s decisions or as a part of her character. In any case, it seems that a stable mental quality is looked for. I focused on the claim that what we should judge is the agent’s (quality of) character, and explored some difficulties for this view. I began by pointing out how misleading it is the transparent picture

of agency as consisting of some agency-related mental traits, from which, to be sure, actions spring, but which are independent of these actions themselves—the view that character is fixed completely independently of overt action. On the contrary, I claimed that there exists a dialectical relationship among character, actions and situations, which is far from being transparent.

On the other hand, the claim that what we should judge is the agent's character presupposes a particular view of character: a specially stable, or robust, view of it. However, it is not evident that this view of character is empirically appropriate. In order to clarify that point I examined and evaluated the discussion about the nature of character traits. Although I rejected eliminativism about character traits, I accepted the importance of the situationist findings for our conception of character. That allowed me to defend, without giving up the view that character traits can rightly be the locus for attributions of moral responsibility, that the idea of restricting moral responsibility attributions exclusively to character, on the assumption that it is a stable property of the agent, is unjustified given the actual instability of character. The ruin of a too self-sufficient view of the agent and her character is an undeniable consequence of these findings. Moreover, they have some particular repercussions for personal autonomy, moral responsibility and, so, for moral luck that need to be stressed. The experiments make it clear that we often—more often than we usually think—ignore the true motives of our actions, which supposes an important deficiency in self-knowledge or that we do not manage to turn our motives or intentions into actions, which involves a deficiency in self-control.

Coming back to cases of circumstantial luck, what determines ascription of moral responsibility, the denier claims, can be neither the

agents' actions, nor their decisions, given that they acted and decided differently. It seems that intentions are the best candidate. But both agents in the example cannot share exactly the same intentions. They can share some general intentions (or determinable intentions, or plans) but not the specification of these intentions (determinate intentions, plan's fulfillment)—different circumstances shape different intentions. My claim was that this difference in their determinate intentions justifies a difference in moral judgment. The denier could reply that the power of the example is due to its ambiguity: if details were refined, we would see that either both counterparts are so similar that there will be no relevant moral difference between them, or that they are so dissimilar that they will not be considered counterparts in the sense relevant to the case any more. But this move dissolves the issue. It bypasses the intended significance of cases of circumstantial luck: how different circumstances give room for different opportunities of action, as well as raise different demands on the agent. I also argued that traits are simply undetermined until certain behavior fixes them. One's character depends—partially, but unavoidably—on one's actual intentions, decisions and actions.

Finally, in Chapter 7 I replied to specific arguments against resultant moral luck, partially drawing on Williams' argument. For the denier, the agent's will or decision and its consequent action is what exclusively counts for attribution of moral responsibility. However, in defense of moral luck, I vindicated the role played by the different reaction that each result triggers in us, and the fact that such distinct results have divergent repercussions in the victim and the agent's future character and identity.

I argued for the claim that the retrospective perspective is an ingredient of moral assessment. My contention was that *what actual harmful consequences do is raising the standard of moral assessment*. This is

particularly clear in cases of negligent behavior and decisions under uncertainty. For instance, regarding the subclass of negligent behavior the denier claims that assignation of moral responsibility depends exclusively on the moral risk assumed, and the moral risk taken should be calculated from the agent's perspective at the time of action, or *ex ante*. Against this view, I argued that, at least in some cases, *it seems that the very negligence exists just in case the outcome is harmful*. The rationale for this claim is that we ordinarily take risks—*moral* risks included—but it would be a too rigorous view of morality and life in general, and indeed impracticable, to equalize moral blame with independence of the actual outcome. The standard of demand cannot ordinarily be so high. We cannot judge both drivers with the same harshness; neither can we treat them with the same leniency. In general, what the claim of the retrospective nature of justification stresses is that one is linked with one's actions in a way that exceeds one's relation to one's own actions as an *ex ante* rational deliberator.

On the other hand, I argued for the relevance of actual causation of harm in ascribing moral responsibility and vindicated the role of our attitudes and emotional reactions in finding out what we value and care about. In my interpretation of Williams' insistence in the importance of agent-regret, what the existence of this sentiment reveals is the significance of actual harm for moral appraisal—and this significance is easy to understand if we consider morality as particularly concerned with guaranteeing respect for persons. I finished considering the moral sentiments of guilt, indignation and resentment—which contrast with agent-regret, as long as they are necessarily connected with voluntariness or moral fault. The existence of the sort of reactive attitudes discussed highlights our concern for harm done, as well as the agent's contribution to it, at once.

The rebuttal of the arguments against circumstantial and resultant moral luck completed my defense of moral luck in all of its kinds. Part III was devoted to the main repercussions of acknowledging the existence of moral luck.

In particular, Chapter 8 addressed the significance and consequences of rejecting the Control Principle. I began by rebutting Nagel's view of moral luck as essentially paradoxical on the basis of the strong conception of control assumed in his argument. Next, I proceeded to supplement my results in Part II with a series of experimental findings pointing to the same upshot. I reported some empirical data that appears to show, on the one hand, that people's attributions of intentional action are often influenced by moral considerations (the *Knobe effect*) and, on the other hand, that the two steps of the mechanism for attributing moral responsibility are deeply entangled. In particular, I pointed out that some of those findings show that folk attributions of moral responsibility are influenced by considerations concerning the *actual* behavior's moral polarity and the *actual* consequences of actions.

On the other hand, as long as the Control Principle is vindicated as a moral aim by means of the Fairness Argument, it could happen that it clashed with other moral aims, so that it turned out that its sacrifice leads to an overall profit. Consequently, and following M. Walker, I argued for the moral superiority of Dependability and the virtue of *taking responsibilities*, over the Control Principle. After conceptualizing the incompatibility between both principles in terms of the notions of pure and impure agency, I proceeded to rebut two objections. On the one hand, I rejected the claim that this view unjustifiedly conflates two distinct senses of responsibility on the basis of both my results in Part II and the preceding empirical data. The second objection was answered by arguing that it

does not follow either from acceptance of moral luck, nor from the argument from moral integrity, that we should consider people free of culpability when negligence does not derive in harm; and neither that people should not make the effort to deliberate and avoid reckless behavior. Moreover, the demand that agents should care about their actual victim to the very same degree or intensity as about their potential victim is striking (given that each one's needs are different) and highly revisionist.

After that, I proposed a new limiting principle for attribution of moral responsibility, the idea of one's *Jurisdiction* (J), in substitution of the Control Principle—with which I intend to pick up the intuitive character of the idea that a person can only be morally responsible for things she is appropriately connected with, but without having to reject moral luck. J can play the role of an abstract limiting principle, acknowledging at the same time the inefficacy of a substantive principle really able to fix in abstract the conditions for attributions of moral responsibility—which agrees better with how our ordinary practices actually work.

Nevertheless, I recognized the remaining feeling of injustice after accepting moral luck, although I rejected a moral-luck-friendly remedy consisting in the separation of moral judgment and punishment or adverse feelings. This proposal would imply giving up a basic way in which we react to wrongdoers, a way in which we ordinarily try to *do justice*. Moreover, reactive attitudes play a fundamental role both in our interpersonal relationships and in our own self-understanding, so that the idea of dropping ordinary praise and blame cannot be the right answer. Acceptance of moral luck leads us to the acknowledgement of the contingency of our moral self and desert. This instills irony into our reactive attitudes and judgments—an irony that makes forgiveness easier, but does not rebut those attitudes and judgments.

Finally, Chapter 9 was devoted to some related meta-theoretical issues. I argued that Nichols and Knobe's results about folk intuitions concerning the Compatibility Question—a conflict between concrete and abstract intuitions—can straightforwardly be transferred to the moral luck issue: here, the default position (in abstract) is that attributions of moral responsibility presuppose control, or are proportional to the degree of control; but it is not difficult to think of concrete presentation of cases that will yield the opposite results. If this projection is correct, it will provide (indirect) experimental support to the idea that moral luck is fundamentally a conflict between abstract and concrete intuitions. However, I rejected Nichols and Knobe's putative demonstration of the affect-biased origin of concrete intuitions. In my view, what lastly is in the origin of the different reactions is the distinct moral seriousness of behavior. Anyway, the experiment's design prevents from discriminating between seriousness and affect in bringing about the effect, which undermines Nichols and Knobe's conclusion. With no further experimental data available, we are justified to think this is a genuine conflict.

Consequently, I proceeded to explore the possible ways out of this conflict. I proposed a typology of revisionism and tried to fill it in with the different proposals in the debate. I characterized my own proposal as a sort of *moderate revisionism* that directly concerns to our theoretical conceptions and—provided that these influence the tacit theory—indirectly affects to our general intuitions. I intended to show that neither practices of moral judgment, nor folk intuitions and beliefs referred to concrete cases—as long as they are not completely permeable to the direct influence of theoretical conceptions—are in need of revision. What needs to be *re-interpreted* are our more abstract and theory-laden beliefs.

I judged that this proposal is more advantageous than the rest as long as it entails an easier-to-assimilate sort of revisionism.

I just finish by restating the obvious: luck is pervasive, and morality is not a metaphysically privileged or *sui generis* sphere immune to it.

Bibliografía

- Adams, Robert M. 1985. "Involuntary Sins", *Philosophical Review* 94: 3-31.
- 2006. *A Theory of Virtue: Excellence in Being for the Good*. Oxford / Nueva York: Oxford University Press.
- Adler, Jonathan. 1987. "Luckless Desert is Different Desert", *Mind* 96: 247-249.
- Anderson, Elisabeth. 1999. "What is the Point of Equality", *Ethics* 109: 339-349.
- Andre, Judith. 1983. "Nagel, Williams, and Moral Luck", *Analysis* 43: 202-207. Reimpresión en Statman 1993.
- Alexander, Jonathan, y Jonathan Weinberg. 2007. "Analytic Epistemology and Experimental Philosophy", *Philosophy Compass* 2 (1): 56-80.
- Arendt, Hanna. 2006. *Eichmann en Jerusalén*. Traducción de Carlos Ribalta. Barcelona: Debolsillo. (Edición original: *Eichmann in Jerusalem: A Report on the Banality of Evil*. Nueva York: The Viking Press, 1963.)
- Arneson, Richard. (en prensa). "Rawls, Responsibility, and Distributive Justice", en M. Salles y J.A. Weymark, eds. *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls*. Cambridge: Cambridge University Press.
- Aristóteles. 1960. *Ética a Nicómaco*, edición bilingüe, traducción de M. Araujo y J. Marías. Madrid: Instituto de Estudios Políticos.

——— *Física*, varias ediciones.

Arpaly, Nomy. 2003. *Unprincipled Virtue. An Inquiry into Moral Agency*.
Oxford: Oxford University Press.

Arpaly, Nomy y Timothy Schroeder. 1999. "Praise, Blame and the Whole
Self", *Philosophical Studies* 93: 161-188.

Asch, Solomon. 1951. "Effects Group Pressure upon the Modification of
Distortion of Judgment", en Harold Guetzkow, ed. *Groups, Lead-
ership, and Men*. Pittsburgh: Carnegie Press, pp. 177-90.

Athanassoulis, Nafsika. 2005. *Morality, Moral Luck and Responsibility:
Fortune's Web*. Basingstoke: Palgrave.

Audi, Robert. 1991. "Responsible Action and Virtuous Character", *Ethics*
101: 304-21.

Bealer, George. 1998. "Intuition and the Autonomy of Philosophy", en
DePaul y Ramsey 1998, pp. 201-39.

Beardsley, Elisabeth L. 1960. "Determinism and Moral Perspectives",
Philosophy and Phenomenological Research 21: 1-20.

Benedict, Ruth. 1945. *The Chrysanthemum and the Sword: patterns of
Japanese culture*. Boston: Houghton Mifflin.

Bennett, Jonathan. 1995. *The Act Itself*. Oxford: Clarendon Press.

——— 1974. "The Conscience of Huckleberry Finn", *Philosophy* 49: 123-
134.

Bernstein, Mark. 2005. "Can We Ever Be Really, Truly, Ultimately,
Free?", *Midwest Studies in Philosophy* 29: 1-12.

Blackburn, Simon. 1998. *Ruling Passions: A Theory of Practical Reason*.
Oxford: Oxford University Press.

Bratman, Michael. 1987. *Intention, Plans, and Practical Reason*. Cam-
bridge, Mass.: Harvard University Press.

- Browne, Brynmor. 1992. "A Solution to the Problem of Moral Luck", *Philosophical Quarterly* 42: 343-356.
- Buss, Sarah, y Lee Overton, eds. 2002. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, Mass.: MIT Press.
- Canetti, Elias. 2005. *Fiesta bajo las bombas*. Barcelona: Galaxia Gutenberg / Círculo de Lectores. (Referencia a la traducción inglesa: *Party in the Blitz: The English Years*. Londres: The Harvill Press, 2005.)
- Card, Claudia. 1990. "Gender and Moral Luck", en Flanagan y Rorty 1990, pp.199-218.
- 1996. *The Unnatural Lottery: Character and Moral Luck*. Filadelfia: Temple University Press.
- Clark, Randolph. 1993. "Toward a Credible Agent-Causal Account of Free Will", *Noûs* 27: 191-203.
- 2003. *Libertarian Accounts of Free Will*. Oxford: Oxford University Press.
- Coffman, E. J. 2007. "Thinking about Luck", *Synthèse* 158: 385-398.
- Cohen, Gerald A. 1989. "On the Currency of Egalitarian Justice", *Ethics* 99: 906-44.
- Corbí, Josep. 2004. *Un lugar para la moral*. Madrid: Antonio Machado Libros.
- Cummings, Robert. 1998. "Reflections on Reflective Equilibrium", en DePaul y Ramsey 1998, pp. 113–27.
- Damasio, Antonio. 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. Nueva York: Grosset / Putnam.
- Daniels, Norman. 2003. "Reflective Equilibrium", *The Stanford Encyclopedia of Philosophy (Spring 2003 Edition)*, Edward N. Zalta, ed. URL = <<http://plato.stanford.edu/entries/reflective-equilibrium/>>

- D'Arms, Justin, y Daniel Jacobson. 2000a. "The Moralistic Fallacy: On the 'Appropriateness' of emotion". *Philosophy and Phenomenological Research* 61: 65-90.
- 2000b. "Sentiment and Value", *Ethics* 110: 722–748.
- 2003. "The Significance of Recalcitrant Emotions (Or Anti-QuasiJudgmental-ism)", en A. Hatzimoysis, ed. *Philosophy and the Emotions*. Cambridge: Cambridge University Press.
- Davis, M. 1986. "Why Attempts Deserve Less Punishment Than Completed Crimes", *Law and Philosophy* 5: 1-32.
- Darley, John, y Daniel Batson. 1973. "'From Jerusalem to Jericho': A Study of Situational and Dispositional Variables in Helping Behavior", *Journal of Personality and Social Psychology* 27: 100-08.
- Deigh, John. 1995. "Empathy and Universalizability", *Ethics* 105: 743-763.
- Dennett, Daniel. 1984. *Elbow Room: Varieties of Free Will Worth Wanting*. Cambridge, Mass: MIT Press.
- DePaul, Michael, y William Ramsey, eds. 1998. *Rethinking Intuitions. The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield.
- Dodds, E. R. 1951. *The Greeks and the Irrational*. Berkeley: Univ. of California Press.
- Domsy, Darren. 2005. "Tossing the Rotten Thing Out: Eliminating Bad Reasons not to Solve the Problem of Moral Luck", *Philosophy* 80: 531-541.
- 2004. "There is No Door: Finally Solving the Problem of Moral Luck", *Journal of Philosophy* 101: 445-464.
- Doris, John. 2002. *Lack of Character*. Cambridge, Mass.: Cambridge University Press.

- Doris, John, Joshua Knobe y Robert L. Woolfolk. 2007. "Variantism about Moral Responsibility", *Philosophical Perspectives*, 21 (*Philosophy of Mind*): 183-214.
- Driver, Julia. 2006. "Luck and Fortune in Moral Evaluation", manuscrito para la *First Online Conference in Philosophy* (con comentario de Hans Maes).
- Dworkin, Ronald. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge: Harvard University Press.
- Ekstrom, Laura. 2002. "Libertarianism and Frankfurt-style Cases", en Kane 2002a.
- Enoch, David, y Andrew Marmor. 2007. "The Case Against Moral Luck", *Philosophy and Law* 26: 405-436.
- Feinberg, Joel. 1995. "Equal Punishment for Failed Attempts: Some Bad But Instructive Arguments Against It", *The Arizona Law Review* 37: 117-133.
- 1970. *Doing and Deserving: Essays in the Theory of Responsibility*. Princeton, NJ: Princeton University Press.
- 1962. "Problematic Responsibility in Law and Morals", *Philosophical Review* 71: 340-351. Reimpreso en Feinberg 1970.
- Fischer, John M. 2006. "The Cards That Are Dealt You", *Journal of Ethics* 10: 107-29.
- 1999. "Recent Work on Moral Responsibility", *Ethics* 110: 93-139.
- Fischer, John M. y Robert H. Ennis. 1986. "Causation and Liability", *Philosophy and Public Affairs* 15 (1): 33-40.
- Fischer, John M., y Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

- Flanagan, Owen. 1991. *Varieties of Moral Personality*. Harvard University Press.
- Flanagan, Owen, y Amélie Rorty, eds. 1990. *Identity, Character, and Morality*. Cambridge, Mass.: MIT Press.
- Frankfurt, Harry. 2002a. "Reply to John Martin Fischer", en Buss y Overton 2002.
- 2002b. "Reply to David Velleman", en Buss y Overton 2002.
- 1999. *Necessity, Volition, and Love*. Cambridge: Cambridge University Press.
- 1999b. "On the Necessity of Ideals", en Frankfurt 1999.
- 1994. "Autonomy, Necessity, and Love", en H. F. Fulda y R. P. Horstmann, eds. 1994. *Vernunftbegriffe in der Moderne: Stuttgarter Hegel-Kongress 1993*. Stuttgart: Klett-Cotta. Reimpreso en Frankfurt 1999.
- 1988. *The Importance of What We Care About*. Cambridge University Press.
- 1988b. "Rationality and the Unthinkable", en Frankfurt 1988, pp. 177-190.
- 1988c. "The Importance of What We Care About", en Frankfurt 1988, pp. 80-94.
- 1987. "Identification and Wholeheartedness", en Schoeman 1987. Reimpreso en Frankfurt 1999.
- 1976. "Identification and Externality". En A. Rorty 1976, pp. 239-251.
- 1971. "Freedom of the Will and the Concept of a Person", *Journal of Philosophy* 68: 5-20. Reimpreso en Frankfurt 1988.
- 1969. "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy* 66: 828-839. Reimpreso en Frankfurt 1988.

- Gendler, Tamar S., y John Hawthorne, eds. 2002. *Conceivability and Possibility*. Nueva York: Oxford University Press.
- Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?", *Analysis* 23: 121–123.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge, Mass.: Harvard University Press.
- Greco, John. 1995. "A Second Paradox Concerning Responsibility and Luck", *Metaphilosophy* 26: 81-96.
- Goldie, Peter. 2000. *The Emotions*. Oxford: Clarendon Press.
- Gopnik, Alison y E. Schwitzgebel. 1998. "Whose concepts are they anyway? The role of philosophical intuition in empirical psychology", en DePaul y Ramsey 1998.
- Goodman, Nelson. 1965. *Fact, Fiction, and Forecast*. Indianápolis: Bobbs-Merril.
- Glover, Jonathan. 1970. *Responsibility*. Oxford: Routledge & K. Paul.
- Haji, Ishtiyaque. 2002. *Deontic Morality and Control*. Cambridge University Press.
- 1998. *Moral Appraisability*. Nueva York: Oxford University Press.
- Haney, Craig, Curtis Banks y Philip Zimbardo. 1973. "Interpersonal Dynamics in a Simulated Prison." *International Journal of Criminology and Penology* 1:69–97.
- Harman, Gilbert. 2003. "No Character or Personality", *Business Ethics Quarterly* 13: 91-92.
- 2000. "The Nonexistence of Character Traits", *Proceedings of the Aristotelian Society* 100: 224.
- 1999. "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error", *Proceedings of the Aristotelian Society* 99: 315-331.

- Hartshorne, Hugh y Mark May. 1928. *Studies in the Nature of Character*, vol. I: *Studies in Deceit*. Nueva York: Macmillan.
- Hart, H. L. A. 1968. *Punishment and Responsibility*. Nueva York: Oxford University Press.
- Herman, Barbara. 1993. *The Practice of Moral Judgment*. Cambridge: Harvard University Press.
- Hieronymi, Pamela. 2004. "The Force and Fairness of Blame", *Philosophical Perspectives* 18 (*Ethics*): 115-148.
- Honderich, Ted. 1988. *A Theory of Determinism: The Mind, Neuroscience, and Life-Hopes*. Oxford: Clarendon Press.
- Hume, David. 1748/1975. *Enquiry concerning Human Understanding*, in *Enquiries concerning Human Understanding and concerning the Principles of Morals*. Edición de L. A. Selby-Bigge, 3ª edición revisada por P. H. Nidditch. Oxford: Clarendon Press.
- 1739-40/2000. *A Treatise of Human Nature*, editado por David Fate Norton y Mary J. Norton. Oxford / Nueva York: Oxford University Press.
- Hurley, Susan. 2003. *Justice, Luck, and Knowledge*. Cambridge: Harvard University Press.
- 2001. "Luck and Equality", *Proceedings of the Aristotelian Society* 75 (supplementary volume): 51-72.
- 1993. "Justice without Constitutive Luck", *Ethics: Royal Institute of Philosophy Supplement* 35: 179-212.
- Isen, Alice, y Paula Levin. 1972. "Effect of Feeling Good on Helping", *Journal of Personality and Social Psychology* 21:384-88.
- Jackson, Frank. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.

- Jacobs, Jonathan. 2001. *Choosing Character: Responsibility for Virtue and Vice*. Ithaca, NY: Cornell University Press.
- Jensen, Henning. 1984. "Morality and Luck", *Philosophy* 59: 323-330. Se cita la reimpresión en Statman 1993.
- Kamtekar, Rachana. 2004. "Situationism and Virtue Ethics on the Content of Our Character", *Ethics* 114: 458-91.
- Kane, Robert, ed. 2002a. *The Oxford Handbook of Free Will*. Nueva York: Oxford University Press.
- 2002b. "Some Neglected Pathways in the Free Will Labyrinth", en Kane 2002a.
- 1999. "Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism", *Journal of Philosophy*, 96: 217-240.
- 1996. *The significance of Free Will*. New York: Oxford University Press.
- 1994. "Free Will: The Elusive Ideal", *Philosophical Studies* 75: 25-60.
- Kant, Immanuel. 1785/1986. *Fundamentación de la Metafísica de las Costumbres*. Traducción de Manuel García Morente. Madrid: Espasa-Calpe, 6ª ed. Original: *Grundlegung der Metaphysik der Sitten*; en *Kants esammelte Schriften*, editados por la *Königliche Akademie der Wissenschaften* y sucesores, Ak. IV (Berlín: George Reimer / W. de Gruyter, 1902-).
- 1797/1902. *Metaphysik der Sitten*, Ak. VI.
- Kessler, Kimberly D. 1994. "The Role of Luck in Criminal Law", *University of Pennsylvania Law Review* 142 (6): 2183-2237.
- Knobe, Joshua. 2006. *Folk Psychology, Folk Morality*. Tesis doctoral, Princeton University.

- 2003. “Intentional Action and Side Effects in Ordinary Language.” *Analysis* 63: 190–93.
- Knobe, Joshua, y John Doris. (en prensa). “Strawsonian Variations: Folk Morality and the Search for a Unified Theory”, en John Doris, ed. *The Handbook of Moral Psychology*. Oxford: Oxford University Press.
- Knobe, Joshua, y Ben Fraser. 2008. “Causal Judgment and Moral Judgment: Two Experiments”, en Walter Sinnott-Armstrong, ed. *Moral Psychology – Volume 2: The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, Mass.: MIT Press.
- Kripke, Saul. 1980. *Naming and Necessity*. Harvard University Press.
- Kunda, Z. 1999. *Social Cognition: Making Sense of People*. Cambridge, Mass.: MIT Press.
- Kupperman, Joel. 1991. *Character*. Nueva York / Oxford: Oxford University Press.
- Latané, Bibb, y John Darley. 1968. “Bystander Intervention in Emergencies: Diffusion of Responsibility.” *Journal of Personality and Social Psychology* 8: 377-83.
- 1970. *The Unresponsive Bystander: Why Doesn't He Help?* Englewood Cliffs: Prentice Hall.
- Latané, Bibb, y Judith Rodin. 1969. “A Lady in Distress: Inhibiting Effects of Friends and Strangers on Bystander Intervention”, *Journal of Experimental Social Psychology* 5: 189-202.
- Latus, Andrew. 2003. “Constitutive Luck”, *Metaphilosophy* 34: 460-75.
- 2001. “Moral Luck”, *The Internet Encyclopedia of Philosophy*, J. Feiser, ed.
URL = <<http://www.iep.utm.edu/m/moralluc.htm>> (Revisado en 2005.)

- . 2000. “Moral and Epistemic Luck”, *Journal of Philosophical Research* 25: 149-172.
- Lehane, Dennis. 2001. *Mystic River*. (Se cita la traducción castellana de Maria Via. Barcelona: RBA Libros, 2003.)
- Levy, Ken. 2005. “The Solution to the Problem of Outcome Luck: Why Harm Is Just as Punishable as the Wrongful Action that It Causes”, *Law and Philosophy* 24: 263- 303.
- Levy, Neil. (en prensa). “Restrictivism Is a Covert Compatibilism”, en N. Trakakis, ed., *Essays on Free Will and Moral Responsibility*. Cambridge Scholar Press.
- . 2005. “The Good, The Bad and The Blameworthy”, *Journal of Ethics and Social Philosophy* 1: 1-16.
- Lewis, David. 1989. “The Punishment That Leaves Something to Chance”, *Philosophy and Public Affairs* 96: 227-242.
- Liao, Matthew S. 2007. “A Defense of Intuitions”, *Philosophical Studies*, en prensa.
- Louden, Robert B. 2007. “The Critique of the Morality System”, en Thomas 2007.
- Lucas, John R. 1993. *Responsibility*. Oxford: Clarendon Press.
- Mackie, John. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
- Malle, Bertram. 2004. *How the Mind Explains Behavior*. Cambridge, Mass.: MIT Press.
- Malle, Bertram, y R. Bennett. 2004. “People’s Praise and Blame for Intentions and Actions: Implications of the Folk Concept of Intentionality”, manuscrito, University of Oregon.

- Mathews, K.E., y L.K. Cannon. 1975. "Environmental Noise Level as a Determinant of Helping Behavior", *Journal of Personality and Social Psychology* 32:571-7.
- McDowell, John. 2007. "The Burdens of Autonomy", conferencia leída en el *Workshop on John McDowell*, Universitat de València.
- Mele, Alfred. 2006. *Free Will and Luck*. Oxford: Oxford University Press.
- Mendus, Susan. 1988. "The Serpent and the Dove", *Philosophy* 63: 331-343.
- Milgram, Stanley. 1963. "Behavioral Study of Obedience", *Journal of Abnormal and Social Psychology* 67: 371-78.
- . 1974. *Obedience to Authority*. Nueva York: Harper and Row Publishers.
- Miller, Christian. 2003. "Social Psychology and Virtue Ethics", *Journal of Ethics* 7: 365-92.
- Montmarquet, James. 1993. *Epistemic Virtue and Doxastic Responsibility*. Lahnman: Rowman & Littlefield.
- Moody-Adams, Michele. 1990. "On the Old Saw That Character Is Destiny", en Flanagan y Rorty 1990, pp. 111-31.
- Moore, A.W. 1990. "A Kantian View of Moral Luck", *Philosophy* 65:297-321.
- Moore, Michael. 1997. *Placing Blame: A Theory of the Criminal Law*. Oxford: Clarendon Press.
- . 1994. "The Independent Moral Significance of Wrongdoing", *Journal of Contemporary Legal Issues* 237-281; reimpresso en Moore 1997, cap. 6.
- Moran, Richard. 2001. *Authority and Estrangement*. Princeton University Press.

- Morton, Adam. 1980. *Frames of Mind*. Oxford: Clarendon Press.
- Moya, Carlos. 2006. *Moral Responsibility. The Ways of Scepticism*. Oxford: Routledge.
- Mulhal, Stephen. 2007. "Luck, Mystery and Supremacy: D. Z. Phillips Reads Nagel and Williams on Morality", *Philosophical Investigations* 30 (3): 266-84.
- Mulligan, Kevin. 2008. "On Being Struck by Value: Exclamations, Motivations and Vocations", en Barbara Merkel, ed. *Wohin mit den Gefühlen? Emotionen im Kontext*. Berlin: Mentis-Verlag.
- 1998. "From Appropriate Emotions to Values", *The Monist* 161-188.
- Nagel, Thomas. 1986. *The View from Nowhere*, Nueva York: Oxford University Press.
- 1979. "Moral Luck", en Nagel 1979, pp. 24-38.
- 1979b. *Mortal Questions*. Nueva York: Cambridge University Press.
- 1976. "Moral Luck", *Proceedings of the Aristotelian Society (supplementary volume)* 76:136–50.
- 1970. *The Possibility of Altruism*. Oxford University Press.
- Nahmias, Eddy. 2007. "Agent's Autonomy and Social Psychology", en Marraffa, Massimo, Mario De Caro y Francesco Ferretti, eds. 2007. *Cartographies of the Mind: Philosophy and Psychology in Intersection*. Dordrecht: Springer, pp.169-185.
- Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer y Jason Turner. 2005. "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility", *Philosophical Psychology* 18: 561–84.
- 2006. "Is Incompatibilism Intuitive?", *Philosophy and Phenomenological Research* 73: 28-53.

- Nelkin, Dana. 2004. "Moral Luck", *The Stanford Encyclopedia of Philosophy (Spring 2004 Edition)*, Edward N. Zalta, ed. (Revisado en 2008.)
- URL = <<http://plato.stanford.edu/archives/spr2004/entries/moral-luck/>>
- 2005. "Freedom, Responsibility and the Challenge of Situationism", *Midwest Studies in Philosophy* 29: 181-206.
- 2007. "Do We Have a Coherent Set of Intuitions about Moral Responsibility", *Midwest Studies in Philosophy* 31: 243-259.
- Nelly, Wright. 1974. "Freedom and Desire", *Philosophical Review* 83: 32-54.
- Nichols, Shaun. 2006. "Folk Intuitions on Free Will", *Journal of Cognition and Culture* 6 (1): 57-86.
- Nichols, Shaun, y Joshua Knobe. 2007. "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions", *Nous* 41 (4): 663-685.
- Nisbett, Richard. 2003. *The Geography of Thought*. New York: Simon & Schuster.
- Nisbett, Richard, y D. Cohen. 1996. *Culture of Honour: The Psychology of Violence in the South*. Boulder, CO: Westview Press.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, Mass.: Harvard University Press.
- 1974. *Anarchy, State, and Utopia*. Nueva York: Basic Books, Inc., Publishers.
- Nussbaum, Martha. 1986. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.

- O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. Nueva York: Oxford University Press.
- Ortega y Gasset, José. 1947. "Pidiendo un Goethe desde dentro", en *Obras Completas*, vol. IV. Madrid: Revista de Occidente, pp. 395-419.
- Oshana, Marina. 2006. "Moral Taint", *Metaphilosophy* 37 (3-4): 353-375.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Pereboom, Derk. 2001. *Living without Free Will*. Cambridge: Cambridge Univ. Press.
- Pettit, Phillip. 2001. *A Theory of Freedom*. Oxford: Oxford University Press.
- Pink, Thomas. 2004. *Free Will: A Very Short Introduction*. Nueva York: Oxford University Press.
- Pizarro, David, Eric Uhlmann y Peter Salovey. 2003. "Asymmetry in Judgments of Moral Blame and Praise: The Role of Perceived Metadesires", *Psychological Science* 14: 267-72.
- Pritchard, Duncan. 2005. *Epistemic Luck*. Oxford: Oxford University Press.
- 2006. "Moral and Epistemic Luck", *Metaphilosophy* 37 (1): 1-25.
- Pritchard, Duncan, y M. Smith. 2005. "The psychology and philosophy of luck", *New Ideas in Psychology* 22: 1-28.
- Ramírez, Erick. (en preparación). *An Argument against Basic Emotions: Salvaging Sentimentalism*. Tesis doctoral, University of California - San Diego, 2008.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Rescher, Nicholas. 1995. *Luck: The Brilliant Randomness of Everyday Life*, New York: Farrar, Straus and Giroux. (Se cita: *La suerte*).

- Aventuras y desventuras de la vida cotidiana*, trad. de Carlos Gardini; Santiago de Chile: Andrés Bello, 1997.)
- 1990. “Luck”, *American Philosophical Association Proceedings* 64 (3): 5-20. Versión revisada, con el título de “Moral Luck”, en Statman 1993. (Paginación de la reedición.)
- Richards, Norvin. 1986. “Luck and Desert”, *Mind* 65: 198-209. Se cita la reimpresión en Statman 1993.
- Robbennolt, J. 2000. “Outcome Severity and Judgments of ‘Responsibility’: A Meta-analytic Review”, *Journal of Applied Social Psychology* 30 (2): 575–609.
- Roemer, John. 1996. *Theories of Distributive Justice*. Cambridge: Harvard University Press.
- Rorty, Amélie. 1980. “Agent Regret”, en Rorty, ed. 1980. *Explaining Emotions*. Los Angeles: University of California Press.
- ed. 1976. *The Identities of Persons*. University of California Press.
- Rorty, Amélie, y David Wong. 1990. “Aspects of Identity and Agency”, en Flanagan y Rorty 1990.
- Rosebury, Brian. 1995. “Moral Responsibility and Moral Luck”, *Philosophical Review* 104: 499-524.
- Rosell, Sergi. 2009. “Tres concepciones de la libertad”, manuscrito, Universitat de València.
- 2008. “Doxastic Voluntarism and the Aim of Belief”, manuscrito presentado en el *IV Meeting on Pragmatism: Truth and the Ethics of Belief*, Universidad de Murcia.
- 2007. “On An Attempt to Undermine Reason-Responsive Compatibilism Appealing to Moral Luck. A response to Gerald K. Harrison”, *Sorites* 19: 7-13.

- Rosen, Gideon. 2004. "Scepticism about Moral Responsibility", *Philosophical Perspectives* 18 (*Ethics*): 295-313.
- Ross, Lee, y Richard Nisbett. 1991. *The Person and the Situation*. Filadelfia: Temple University Press.
- Ross, Lee 1988. "Situationist Perspectives on the Obedience Experiments: Review of A.G. Miller", *Contemporary Psychology* 33: 101-104.
- Royzman, E., y R. Kumar. 2004. "Is Consequential Luck Inconsequential? Empirical Psychology and Moral Luck", *Ratio* 17: 329-344.
- Russell, Paul. 2008. "Free Will, Art and Morality", *Journal of Ethics* 12: 307-325.
- 1999. "Smith on Moral Sentiment and Moral Luck", *History of Philosophy Quarterly* 16: 37-58.
- 1995. *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility*. Oxford / Nueva York: Oxford University Press.
- Sabini, John, Michael Siepmann y Julia Stein. 2001. "The Really Fundamental Attribution Error in Social Psychological Research", *Psychological Inquiry* 12: 1-15.
- Sabini, John, y Maur Silver. 2005. "Lack of Character: Situationism Critiqued", *Ethics* 115 (3): 535-562.
- Sartre, Jean Paul. 1946/1996. *L'existentialisme est un humanisme*. París: Gallimard.
- Scanlon, Thomas. 1998. *What we Owe each Other*. Cambridge: Harvard Univ. Press.
- Scheffler, S. 2003. "What is Egalitarianism", *Philosophy and Public Affairs* 31: 5-39.
- Schoeman, Ferdinand, ed. 1987. *Responsibility, Character, and the Emotions*. Ithaca: Cornell University Press.

- Shaver, S. 1970. "Redress and Conscientiousness in the Attribution of Responsibility for Accidents", *Journal of Experimental Social Psychology* 6: 100–110.
- Sher, George. 2006. "Out of Control", *Ethics* 116: 285-301.
- 2005. "Kantian Fairness", *Philosophical Issues* 15: 179–92.
- 2001. "Blame for Traits", *Nous* 35 (1): 146-161.
- Sinnott-Armstrong, W. 2006. "Moral Intuitionism Meets Empirical Psychology", en T. Horgan y M. Timmons, eds. *Metaethics after Moore*. Nueva York: Oxford University Press, pp. 339–365.
- Slote, Michael. 1994. "The Problem of Moral Luck", *Philosophical Topics* 22 (1-2): 397-409.
- Smart, J.J.C., y Bernard Williams. 1983. *Utilitarianism For & Against*. Cambridge: Cambridge University Press.
- Smilansky, Saul. 2000. *Free Will and Illusion*. Oxford: Oxford University Press.
- 1994. "Fortunate Misfortune", *Ratio* (new series) 7: 153–63.
- Smith, Adam. 1790/1976. *The Theory of Moral Sentiments*, editado por D. D. Raphael y A. L. Macfie. Oxford: Clarendon Press.
- Smith, Angela. 2003. "Responsibility for Attitudes: Activity and Passivity in Mental Life", *Ethics* 115: 236-271.
- Smith, Holly. 1983. "Culpable Ignorance", *Philosophical Review* 92: 543–71.
- Solomon, Robert, ed. 2003. *Thinking About Feeling: Contemporary Philosophers on Emotions*. Nueva York: Oxford University Press.
- Sommers, Tamler. 2007. "Metaskepticism about Moral Responsibility", presentación en la *Conference on Responsibility, Agency, and Persons*, Universidad de San Francisco.

- Sosa, Ernest. 2008. "A Defense of the Use of Intuitions in Philosophy", en Michael Bishop y Dominic Murphy, eds. *Stich and His Critics*. Oxford: Blackwell.
- 1998. "Minimal Intuition", DePaul y Ramsey 1998, pp. 257–69.
- Statman, Daniel. 2005. "Doors, Keys, and Moral Luck: A Reply to Domsky", *Journal of Philosophy* 422-36.
- ed. 1993. *Moral Luck*, Albany: State University of New York Press.
- 1991. "Moral and Epistemic Luck", *Ratio* 4:146-156.
- Stocker, Michael. 2007. "Shame, Guilt, and Pathological Guilt", en Thomas 2007.
- Strawson, Peter F. 1985. *Skepticism and Naturalism: Some Varieties*. Nueva York: Columbia University Press. (Se cita por la versión castellana: *Escepticismo y naturalismo*; trad. de Susana Badiola. Madrid: Antonio Machado Libros, 2003.)
- 1974. "Freedom and Resentment", en *Freedom and Resentment and Others Essays*. Londres, Methuen.
- Strawson, Galen. 1994. "The Impossibility of Moral Responsibility", *Philosophical Studies* 75: 5-24.
- 1986. *Freedom and Belief*. Nueva York: Oxford University Press.
- Sverdlik, S. 1988. "Crime and Moral Luck", *American Philosophical Quarterly* 25: 79-86. Reimpresión en Statman 1993.
- Swan, S., J. Alexander y J. Weinberg. 2008. "The Instability of Philosophical Intuitions: Running Hot & Cold on Truetemp", *Philosophy and Phenomenological Research*, en prensa.
- Tannenbaum, Julia. 2007. "Emotional expressions of moral value", *Philosophical Studies* 132: 43-57.
- Taylor, Charles. 1976. "Responsibility for Self", en Rorty 1976, pp. 281-99.

- Taylor, Craig. (manuscrito). "Moral Luck", réplica a Pritchard 2006.
- Thomas, Alan, ed. 2007. *Bernard Williams*. Cambridge: Cambridge University Press.
- Thompson, Janna. 2000. "The Apology Paradox", *Philosophical Quarterly* 50 (201): 470-5.
- Thomson, Judith Jarvis. 1993. "Morality and Bad Luck". Reimpresión en Statman 1993.
- Trianosky, Gregory. 1990. "Natural Affection and Responsibility for Character: A Critique of Kantian Views of the Virtues," en Flanagan y Rorty 1990, pp. 93-109.
- Vargas, Manuel. (manuscrito). "Moral Influence, Moral Responsibility". Universidad de San Francisco, 2008.
- 2007. "Revisionism", en Fischer, John, Robert Kane, Derk Pereboom y Manuel Vargas. *Four Views on Free Will*. Blackwell.
- 2006. "Philosophy and the Folk: On Some Implications of Experimental Work for Philosophical Debates on Free Will", *Journal of Cognition and Culture* 6 (1-2): 239-254.
- 2005a. "The Revisionist's Guide to Responsibility", *Philosophical Studies* 125 (3): 399-429.
- 2005b. "Problems for Tracing", *Midwest Studies in Philosophy* 29: 269-290.
- 2004. "Responsibility and the Aims of Theory: Strawson and Revisionism" *Pacific Philosophical Quarterly* 85: 2.
- Velleman, David. 2002. "Identification and Identity", en Buss y Overton 2002, pp. 91-123.
- Viney, W., P. Parker-Martin y S.D.H. Dotten. 1988. "Beliefs in Free Will and Determinism and Lack of Relation to Punishment Rationale and Magnitude", *Journal of General Psychology* 115: 15-23.

- Viney, W., D. Waldman y J. Barchilon. 1982. "Attitudes toward Punishment in Relation to Belief in Free Will and Determinism", *Human Relations* 35: 939–49.
- Vogel, Lawrence. "Understanding and Blaming: Problems in Attribution of Moral Responsibility", *Philosophy and Phenomenological Research* 53 (1): 129-142.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press.
- Walker, Margaret Urban. 1991. "Moral Luck and the Virtues of Impure Agency", *Metaphilosophy* 22: 14-27. Se cita la reimpresión en Statman 1993.
- Walster, E. 1966. "Assignment of Responsibility for an Accident", *Journal of Personality and Social Psychology* 3: 73–79.
- Watson, Gary. 2001. "Reason and Responsibility", *Ethics* 111: 374-94.
- 1996. "Two Faces of Responsibility," *Philosophical Topics* 24: 227–48. Reimpre-so en Watson. 2004. *Agency and Answerability*. Oxford: Clarendon Press, pp. 260-288. Paginación de esta última edición.
- 1987. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme", en Schoeman 1987, pp. 256-86.
- 1975. "Free Agency", *Journal of Philosophy* 72: 205-20.
- Weinberg, Jonathan M. 2007. "How to Challenge Intuitions Empirically Without Risking Skepticism", *Midwest Studies in Philosophy* 31: 318-343.
- Weinberg, Jonathan M., Shaun Nichols y Stephen Stich. 2001. "Normativity and Epistemic Intuitions", *Philosophical Topics* 29 (1-2): 429-460.

- Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton University Press.
- 1995. *Making Sense of Humanity*. Cambridge University Press.
- 1995b. “Resisting one’s own existence”, en Williams 1995.
- 1995c. “Moral Incapacity”, en Williams 1995.
- 1993. “Postscript”, en Statman 1993, pp. 251-58. Reimpreso en Williams 1995.
- 1993b. *Shame and Necessity*. Berkeley: University of California Press.
- 1985. *Ethics and the Limits of Philosophy*. Londres: Fontana.
- 1981. “Moral Luck”, en Williams 1981b, pp. 20-39.
- 1981b. *Moral Luck*. Cambridge: Cambridge University Press.
- 1976. “Moral Luck”, *Proceedings of the Aristotelian Society* 50 (supplementary volume): 115–35.
- 1973. *Problems of the Self: Philosophical Papers 1956-1972*. Cambridge: Cambridge University Press.
- 1973b. “Imagination and the Self”, en Williams 1973, pp. 26-45.
- 1973c. “Deciding to Believe”, en Williams 1973, pp. 136-151.
- 1973d. “Ethical Consistency”, en Williams 1973, pp. 166-186.
- 1973e. “Morality and the Emotions”, en Williams 1973, pp. 207-229.
- Williamson, Timothy. 2004. “Philosophical ‘Intuitions’ and Skepticism about Judgment”, *Dialectica* 58: 109–153.
- Winch, Peter. 1972. *Ethics and Action*. Londres: Routledge & Kegan Paul.
- Wittgenstein, Ludwig. 1988. *Investigaciones Filosóficas*. Edición bilingüe; traducción de Alfonso García Suárez y Ulises Moulines. Madrid: Crítica.

- Wolf, Susan. 1980. "Asymmetrical Freedom", *Journal of Philosophy* 77: 157-66.
- 1981. "The Importance of Free Will", *Mind* 90: 386-405.
- 1987. "Sanity and the Metaphysics of Responsibility", en Schoeman 1987.
- 1990. *Freedom within Reason*. Nueva York: Oxford University Press.
- 1997. "Meaning and Morality", *Proceedings of the Aristotelian Society* 97: 299-315.
- 2001. "The Moral of Moral Luck", *Philosophic Exchange* 31: 4-19.
- Woolfolk, R., J. Doris y J. Darley. 2006. "Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility", *Cognition* 100: 283–301.
- Zagzebski, Linda. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press.
- Zimbardo, Phillip. 2004. "A Situationist Perspective on the Psychology of Evil: Understanding How Good People Are Transformed Into Perpetrators", en A.G. Miller, ed. *The Social Psychology of Good and Evil*. Nueva York: Guilford Press, pp. 21-50.
- Zimmerman, David. 2000. "Making Do: Troubling Stoic Tendencies in an Otherwise Compelling Theory of Autonomy", *Canadian Journal of Philosophy*, 25-53.
- 2003. "That Was Then, This Is Now: Personal History vs. Psychological Structure in Compatibilist Theories of Autonomous Agency", *Nous* 37 (4): 638-671.
- Zimmerman, Michael. 1987. "Luck and Moral Responsibility", *Ethics* 97: 374-386. Se cita la reimpresión en Statman 1993.

- 1988. *An Essay on Moral Responsibility*. Totowa, NJ: Roman and Littlefield.
- 1997. “Moral Responsibility and Ignorance”, *Ethics* 107: 410–26.
- 2002. “Taking Luck Seriously”, *Journal of Philosophy* 99: 553-576.
- 2006. “Moral Luck: A Partial Map”, *Canadian Journal of Philosophy* 36 (4): 585-608.

Índice detallado

INTRODUCCIÓN (INTRODUCTION).....	15
RESUMEN (PRÉCIS).....	I

PARTE I

CAPÍTULO 1. LAS ATRIBUCIONES DE RESPONSABILIDAD.....	23
1.1. Esquema básico, relación causal y agencia.....	23
1.2. Las actitudes reactivas y Agentes Moralmente Responsables.....	27
1.3. Voluntariedad y control.....	30
1.4. El sistema de la responsabilidad moral y sus tensiones.....	35
CAPÍTULO 2. LA SUERTE MORAL. <i>Williams y Nagel sobre la suerte moral</i>	39
2.1. ¿Qué es la suerte moral?.....	39
2.1.1. La paradoja de Nagel acerca del control y la responsabilidad.....	41
2.1.2. Williams: justificación retrospectiva y ‘agent regret’.....	50
2.2. Casos y variedades de la suerte moral.....	56
2.2.1. Tipos de casos y el corolario de PC.....	57
2.2.2. Los cuatro tipos nagelianos discutidos (y modificados).....	61
2.3. Contraste de los planteamientos de Nagel y Williams.....	68
2.4. La articulación del debate: ¿prácticas o principio?.....	73
APÉNDICE: <i>La suerte moral y el debate sobre el libre albedrío y la responsabilidad moral</i>	81

PARTE II

CAPÍTULO 3. EL CASO CONTRA LA SUERTE MORAL. <i>Articulación y respuesta</i>	93
3.1. Reconstrucción del Caso Global contra la Suerte Moral.....	94
3.1.1. Fundamentos: el Principio de Control y el Argumento Epistémico.....	94
3.1.2. Escenarios contrafácticos.....	102
3.1.3. La Estrategia Moderada y la Estrategia Radical.....	107
3.2. Réplica al Caso Global contra la Suerte Moral.....	109
3.2.1. Tipos de evaluación moral y la insuficiencia de EM.....	110
3.2.2. ER: historias posibles y merecimiento último. Dudas escépticas y revisionismo.....	116
3.2.3. ER: identidad e incoherencia.....	127
3.3. Recapitulación.....	132
3.3.1. La estructura del argumento.....	132
3.3.2. Contraste con el argumento del <i>todo-o-nada</i>	133
3.3.3. El distinto significado moral de los diversos tipos de suerte moral.....	139

CAPÍTULO 4. LA SUERTE CONSTITUTIVA. <i>Sobre necesidad y contingencia en nuestros orígenes e identidad</i>	145
4.1. Suerte antecedente y estrategias.....	146
4.2. ¿Es la noción de suerte constitutiva incoherente?.....	152
4.3. La noción de suerte.....	157
4.4. Suerte y fortuna, o la especial relevancia moral del carácter.....	166
4.5. Posibilidades formativas e indeterminación.....	171
CAPÍTULO 5. LA SUERTE FORMATIVA. <i>Responsabilidad por el carácter</i>	179
5.1. Autocontrol reflexivo.....	181
5.1.1. Reflexividad volitiva.....	181
5.1.2. Reflexividad racional.....	185
5.1.3. Dos proyectos.....	188
5.2. Competencia normativa y condiciones sociales.....	190
5.2.1. Apreciar los Verdadero y lo Bueno.....	192
5.2.2. Afectividad y relaciones interpersonales en la formación del carácter.....	198
5.3. Autocreación. Decisiones y acciones autoformativas.....	202
5.3.1. ¿Autocreación?.....	202
5.3.2. Decisiones y acciones autoformativas.....	205
5.3.3. La “decisión” de Raskólnikov.....	211
5.3.4. La repercusión futura de las decisiones autoformativas.....	213
5.4. El rastreo de la responsabilidad y la condición epistémica.....	217
5.5. Carácter y evaluación. Conclusiones.....	229
CAPÍTULO 6. SUERTE CIRCUNSTANCIAL. <i>Atrapados por las circunstancias</i>	239
6.1. El rechazo de la suerte circunstancial.....	241
6.1.1. El Principio de Control.....	241
6.1.2. Carácter y acción.....	247
6.2. Disposiciones y la fuerza de las circunstancias: el reto situacionista.....	252
6.2.1. Los experimentos de Milgram sobre la obediencia.....	253
6.2.2. Otros experimentos situacionistas.....	260
6.2.3. Repercusiones: escepticismo acerca de los rasgos de carácter.....	264
6.2.4. Resistir el eliminativismo.....	269
6.3. Circunstancias y evaluación.....	280
6.3.1. Intenciones, planes e indeterminación. El significado de la suerte circunstancial.....	280
6.3.2. Más repercusiones: la opacidad de los propios motivos e intenciones.....	288
6.3.3. Sobre los dilemas morales.....	293
CAPÍTULO 7. SUERTE RESULTANTE. <i>Justificación retrospectiva y emociones morales</i>	299
7.1. Contra la suerte moral Resultante.....	301
7.1.1. El primado de la voluntad.....	301
7.1.2. Negligencia, riesgo y previsibilidad.....	306
7.2. Justificación y retrospectividad. La decisión de Gauguin.....	313
7.2.1. Justificación racional y justificación moral. Conflicto de valores y necesidad.....	315

7.2.2. Deliberación y resultado.....	322
7.2.3. El lugar de la suerte moral: ¿qué pasó con la familia Gauguin?.....	329
7.3. Juicios, actitudes y emociones.....	334
7.3.1. El pesar-del-agente, un sentimiento moral.....	335
7.3.2. Actitudes reactivas y juicio moral.....	345
7.3.3. Agencia y daño.....	353

PARTE III

CAPÍTULO 8. EL RECHAZO DEL CONTROL Y SUS CONSECUENCIAS.....	359
8.1. El principio de control.....	360
8.1.1. Nociones de control. El problema escéptico.....	360
8.1.2. El Principio de Equitatividad y el ideal de justicia última.....	367
8.2. Contra el principio de control.....	369
8.2.1. Causación y asimetrías normativas.....	370
8.2.2. Asunción de responsabilidades, agencia impura e integridad.....	378
8.3. La insostenible fragilidad del juicio moral.....	385
8.3.1. Jurisdicciones, límites borrosos y perspectivas.....	387
8.3.2. Censura, castigo e ironía.....	394
8.4. Diferencias culturales en el concepto de responsabilidad moral.....	402
CAPÍTULO 9. TEORÍA, PRÁCTICA Y CHOQUE DE INTUICIONES.....	413
9.1. Intuiciones comunes y su papel en la construcción teórica.....	414
9.1.1. Naturaleza de las intuiciones.....	414
9.1.2. Análisis conceptual y filosofía experimental.....	416
9.1.3. Algunos desiderata metodológicos.....	420
9.2. Intuiciones acerca de la responsabilidad moral.....	423
9.2.1. La apelación a las intuiciones populares.....	423
9.2.2. Experimentos: ¿somos incompatibilistas ‘naturales’?.....	425
9.2.3. La cuestión de la compatibilidad y la suerte moral.....	430
9.3. Conflicto de intuiciones.....	433
9.3.1. Tipos de conflicto.....	433
9.3.2. Investigación de las fuentes del conflicto concreto/abstracto.....	435
9.4. La suerte moral: conflicto y revisionismo.....	440
9.4.1. Suerte moral y tipos de revisionismo.....	441
9.4.2. Revisionismo moderado: ¿resultado de un “equilibrio reflexivo”?..	445
CONCLUSIONES (CONCLUSIONS).....	451
BIBLIOGRAFÍA.....	463