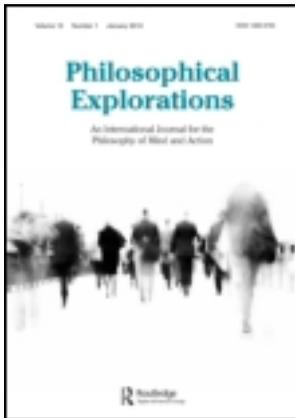


This article was downloaded by: [Brown University]

On: 01 July 2013, At: 21:35

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Philosophical Explorations: An International Journal for the Philosophy of Mind and Action

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rpex20>

### A Strawsonian look at desert

Adina L. Roskies<sup>a</sup> & Bertram F. Malle<sup>b</sup>

<sup>a</sup> Department of Philosophy , Dartmouth College , Hanover , NH , USA

<sup>b</sup> Department of Cognitive, Linguistic, and Psychological Sciences , Brown University , Providence , RI , USA

Published online: 14 May 2013.

To cite this article: Adina L. Roskies & Bertram F. Malle (2013): A Strawsonian look at desert, *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, 16:2, 133-152

To link to this article: <http://dx.doi.org/10.1080/13869795.2013.787439>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## A Strawsonian look at desert

Adina L. Roskies<sup>a\*</sup> and Bertram F. Malle<sup>b</sup>

<sup>a</sup>*Department of Philosophy, Dartmouth College, Hanover, NH, USA;* <sup>b</sup>*Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, USA*

P.F. Strawson famously argued that reactive attitudes and ordinary moral practices justify moral assessments of blame, praise, and punishment. Here we consider whether Strawson's approach can illuminate the concept of desert. After reviewing standard attempts to analyze this concept and finding them lacking, we suggest that to deserve something is to justifiably receive a moral assessment in light of certain criteria – in particular, eligibility criteria (a subject's properties that make the subject principally eligible for moral assessments) and assignment criteria (particulars about the subject, act, and circumstances that justify assessments such as blame in a particular case). Strawson's analysis of the ordinary attitudes and practices of moral assessment hints at these criteria but does not unequivocally ground a notion of desert. Following Strawson's general naturalistic approach, we show that recent psychological research on folk concepts and practices regarding freedom, moral responsibility, and blame illuminates how people actually arrive at moral assessments, thus revealing the very eligibility criteria and assignment criteria we suggest ground a concept of desert. By pushing the Strawsonian line even further than Strawson did, by empirically investigating actual moral practice and folk understandings, we can illuminate desert and lend credence to Strawson's general anti-metaphysical position.

**Keywords:** free will; determinism; responsibility; blame; folk concepts; desert

P.F. Strawson's seminal paper *Freedom and Resentment* redefined the character of philosophical arguments about freedom and responsibility, shifting the perspective from the metaphysical to the normative. In *Freedom and Resentment*, Strawson<sup>1</sup> rejects the common Libertarian view that moral responsibility presupposes a contra-causal freedom and argues instead that our reactive attitudes and ordinary practices ground and justify our moral assessments. As he sees it, the traditional debate about whether blame and punishment can be justified at all – and if so, what justifies them – is misconceived. Strawson characterizes that debate in terms of optimists and pessimists about determinism, whose monikers describe their attitudes toward our ability to justify our moral practices if determinism is true. Strawson's pessimist fears that the truth of determinism undermines or invalidates our moral practices of holding people responsible, praising and blaming them for their actions, and punishing them, because those practices presuppose a notion of desert that cannot be reconciled with determinism and the lack of freedom he believes it entails. His optimist thinks instead that even if determinism is true our practices of praise, blame, and moral assessment are justified or justifiable in virtue of their instrumental value. Although Strawson is more sympathetic to the optimist, he takes issue with both

---

\*Corresponding author. Email: [adina.roskies@dartmouth.edu](mailto:adina.roskies@dartmouth.edu)

these positions. He accedes that each captures aspects of the truth, but his own view is that both are mistaken in that each fails to recognize the import of our everyday moral practices and of the reactive attitudes such as resentment, gratitude, and guilt, for grounding those practices. Strawson's view represents a radical shift, in that it looks to the actual world and truths about human interaction, rather than to metaphysical argument, for grounding responsibility. In the 50 years since the publication of *Freedom and Resentment*, Strawson's positive story has redirected the debate about free will, responsibility, and desert in ways that continue to bear fruit.

*Freedom and Resentment* is primarily about moral responsibility and its relationship to determinism, but it has implications for the less-discussed notion of desert. Here we consider whether and how Strawson's argument and the Strawsonian approach more generally succeed in illuminating desert. The discussion will proceed as follows. First we review some philosophical attempts to analyze the concept and suggest our own way of understanding it. Next we consider Strawson's arguments regarding whether and how moral responsibility, and the practices of moral approbation and blame connected with it, can be reconciled with determinism. We ask whether Strawson's arguments can be extended to justify a notion of desert. Here, we argue, the outcome is equivocal. However, taking Strawson's cue to pay closer attention to our commonsense moral practices, we look at some recent work in psychology that investigates folk concepts and practices regarding freedom and moral responsibility. These studies illuminate how people actually arrive at judgments of blame and desert, and lend credence to Strawson's general anti-metaphysical position. We suggest that by pushing the Strawsonian line even further than Strawson did, by empirically investigating actual moral practice and folk understandings, we can dispel some of the uncertainty about desert that *Freedom and Resentment* provokes.

### 1. Standard conceptual analyses of desert

Philosophers have typically sought to clarify concepts and conceptual relations by armchair analysis. By laying out necessary and sufficient conditions for a concept's proper deployment, the concept – at first obscure or confused – can be clarified and made more precise. Among the concepts of particular interest to philosophers are those important to moral and social thought. Although the concept of desert has garnered much less attention in the philosophical literature than its conceptual relatives such as praise, blame, and responsibility, it plays an important role in social and moral transactions. Because of its centrality, it deserves a closer look.

There is considerable agreement about the conceptual structure of desert: desert is usually conceptualized as a three-place relation between a subject, an object of desert (e.g. praise, blame, reward), and a desert basis (Feinberg 1970; Kleinig 1971; McLeod 2008). The desert basis is a fact or property in virtue of which the object of desert (e.g. praise, blame, compensation) is in fact appropriately assigned to the subject. The desert basis provides a reason or justification for the desert claim. It is widely accepted that the desert basis must be a fact about the subject (Feinberg 1970): either a fact adducing a characteristic of the subject or a fact about an act that he or she has performed.<sup>2</sup> Facts that have nothing to do with the subject cannot serve as the sole basis of desert claims.

However, attempts to provide concrete conceptual content beyond this general structure have been less successful. When is the object of desert 'appropriately assigned' or properly due to the subject? Desert claims always involve evaluation, and thus they are normative, whether moral or nonmoral. What is the ground for their normativity? Is there a single concept of desert that encompasses both moral and nonmoral desert claims? Some

philosophers recognize moral and nonmoral desert as importantly different and emphasize the heterogeneity of desert bases (Feinberg 1970; Kleinig 1971; Sher 1989), while others aim for an overarching analysis (Baiaxu 2006; Kershner 2008). Although we think a general account of desert can accommodate both types, for the remainder of the paper we focus exclusively on moral desert: the basis for the fittingness of moral praise, blame, and punishment.<sup>3</sup>

In most discussions, the concepts of desert and justice are closely linked. Desert is explicated as getting what one is due, where what one is due is what considerations of justice would demand should be accorded to the deserving subject. However, in the absence of an independent theory of justice this account of desert is largely unhelpful. Indeed, some accounts of justice explicitly analyze justice in terms of desert: 'Justice is getting what one deserves' (Hospers, from Feldman 1995). Feinberg also recognizes the problematic nature of the relationship between justice and desert: 'What is it to deserve something? ... Until its peculiar perplexities are resolved, a full understanding of the nature of justice is impossible, for surely the concepts of justice and desert are closely connected' (Feinberg 1970, 55). Since the concept of justice is perhaps even more nuanced than the concept of desert, and since understanding what makes something just is arguably as obscure as understanding what is deserved and why, analyzing desert as justice does not seem like a promising strategy.

An alternative approach to understanding desert is via the concept of fittingness: A deserves X in virtue of P only if it is fitting or appropriate for A to get X because of P. However, a full analysis then requires an account of what is fitting or appropriate in terms that do not make reference to desert. Without being able to discharge this burden, the initially promising approach threatens to be circular. Alternatively, fittingness could be taken to be primitive. However, since understanding the concept of desert is arguably more accessible and straightforward than understanding fittingness, it is not clear how productive this analytical approach would be.

In an attempt to further explicate the concept of desert some philosophers distinguish it from other closely related concepts such as qualification, entitlement, rights, and duties. For instance, Feinberg (1970) explains that you may *deserve* blame and praise but you are not 'entitled' to it or 'qualified' for it. Because we have some intuitive purchase on our language these distinctions are helpful, but the approach still leaves the major question unanswered: what makes *what is deserved* different from what one is entitled to or qualified for? Feinberg argues that desert goes beyond conditions of eligibility, entitlement, and liability, desert-resembling notions that are linked to institutional frameworks. Desert implies a worthiness in a way not codified in any formal rules but instead depending on the satisfaction of conditions required by the private sensitivities of a qualified judge. However, what these conditions are and who would qualify as such a judge are questions left unanswered.

The appeal to a qualified judge in an account of desert is not without merit, for it seems to capture something intuitive, albeit vague, about the notion of desert. Perhaps this notion can be leveraged to distinguish between instances of moral and nonmoral desert: in nonmoral desert the qualified judge is one that is or could legitimately be codified by particular existent or hypothetical institutional frameworks, whereas desert of a moral kind is what remains when no institutionally sanctioned judge is identifiable.

As we shall explain in more detail later, our own view is that desert does not denote an entity (or even a property), but rather it acts as a placeholder for various criteria that if satisfied render our specific moral practices (e.g. blaming or punishing) justified. We thus prefer to focus on the relational formulation mentioned earlier: A deserves X because of P. That very formulation leads to the idea of *criteria* for X. To deserve something is thus to justifiably

receive a (moral) assignment in light of certain criteria. Elucidation of those criteria is a continuing job for philosophy and psychology, but it seems they include two types: (1) eligibility criteria, which must be met in order for a subject to even be a candidate for our moral practices to be potentially justified and (2) assignment criteria, which are particulars about the subject, act, and circumstances that determine whether assigning, say, blame or punishment, is warranted in a particular case. The view we advocate is not unlike Feinberg's, which recognizes 'qualifications' and posits certain criteria for desert bases.<sup>4</sup>

### 1.1 *Desert and the responsibility requirement*

Further clarification of the notion of desert will thus involve getting clearer on the criteria of justification for the objects of moral desert. It is often argued that in order for praise or blame to be deserved, the subject must be responsible for the desert base (Feldman 1995). The intuition is that if the desert base is just a fact about the subject due to luck or accident, if there is nothing about the subject that brings about the existence of the desert base, then desert is undermined. Therefore, one is not (morally) deserving unless one is (morally) responsible for the fact that serves as the desert base. Let us call this plausible and widely held principle that there is no desert without moral responsibility the 'Responsibility Requirement (RR)':

**RR:** No desert without responsibility.<sup>5</sup>

This principle seems to be embraced by almost all theorists of moral desert. R. Jay Wallace appears to have the RR in mind when he writes 'this is a fundamental principle of desert; it expresses an abstract moral conviction in which reflective moral judges have the highest degree of confidence . . . I will call it the principle of no blameworthiness without fault.' (Wallace 1998, 135) Wallace's principle calls attention to the relation between desert and transgression: what it is to be worthy (i.e. deserving) of blame is (moral) responsibility for violation of some moral norm. This idea goes for praise and reward as well. A similar view is put forth by Sadurski, who writes:

When we are pronouncing judgments of desert we are inevitably making judgments about persons whom we hold responsible for their actions. It makes no sense to attribute desert, positive or negative, to persons for actions or facts over which they have no control. (Sadurski 1985, 117)

Desert is here necessarily dependent on moral responsibility and moral responsibility is indexed by judgments of control.<sup>6</sup> Freiman and Nichols (2011) note the widespread reluctance among philosophers to base desert claims on luck and identify what they call 'the brute luck constraint' as a principle of moral common sense: desert should not be affected by luck, where luck is something outside of our control and for which we are therefore not morally responsible. If differential benefits are to be distributed to people on the basis of desert, then the desert base cannot be affected by luck. In other words, mere accident or luck, such as one's genetic makeup, the country or socioeconomic class one is born into, or one's talents, are not properties that can affect one's deservingness.<sup>7</sup>

By tying desert to moral responsibility, the RR seemingly connects desert to the problem of free will, for free will is often assumed to be a necessary condition for moral responsibility. In brief, the supposed connection is this: if free will does not exist then people are not morally responsible at all, for anything, ever. And if people are never morally responsible for anything, then nothing can serve as a desert base, and so nothing can be deserved. If this is true, we are left with two main options. Either our praising

and blaming and punishing practices are unjustified (and therefore unjust) or they are justified on purely instrumental or consequentialist grounds. Many people think that praise and blame, punishment and reward can only be justified if they are deserved – instrumental value is not a possible justifier for our moral practices. For example, Galen Strawson writes: ‘ultimate responsibility exists if and only if punishment and reward can be truly just and fair without having any pragmatic justification’ (Strawson 2008, 361). If so, and if no one is ultimately responsible for anything, as Galen Strawson thinks, then none of our common practices of praise, blame, punishment, nor the social and legal institutions that depend on them, are legitimate. This dilemma is precisely the problem to which Strawson’s *Freedom and Resentment* is addressed.

## 2. Theories of freedom and responsibility

Strawson’s target in *Freedom and Resentment* is the standard challenge to free will and the legitimacy of moral responsibility: the challenge from determinism. Determinism rules out Libertarian freedom and also plausibly undermines certain notions of control that some may think essential for moral responsibility. For example, some incompatibilists argue that if determinism is true, then no one has control over anything, so nothing one does or is can constitute a desert base (van Inwagen 1986).

The philosophical literature contains a number of classic examples of arguments that conclude that no one ever deserves anything. For example, Galen Strawson argues that what he calls Ultimate-Freedom (U-freedom) would be necessary in order for punishment or reward to be deserved. U-freedom is possible only if it is possible that we can be ultimately morally responsible for the way we are. Galen Strawson denies that we can be ultimately morally responsible for anything, since what he argues is a necessary condition for ultimate responsibility – being a cause of oneself – is impossible (Strawson 1986, 1994). One of the interim conclusions that forms part of his Basic Argument seems to embrace the RR and deny the possibility of moral desert:

One cannot be ultimately morally responsible for one’s decisions or actions: one cannot be ultimately morally deserving of praise or blame for one’s decisions or actions or one’s character or indeed for anything else. (Strawson 2008, 360)

If G. Strawson’s Basic Argument succeeds in showing that U-freedom and U-responsibility are impossible, and if the RR is true, then it seems that no one ever deserves praise or blame, punishment or reward.

Galen Strawson and other hard incompatibilists think that freedom and responsibility are incompatible both with determinism and with indeterminism. Some Libertarians in contrast believe that freedom and responsibility are possible, and possible because determinism is false. Although their views vary, they all maintain that agents deserve praise or blame because they are free and responsible for at least some of their actions. This responsibility is secured because of a kind of radical freedom: in Kane’s view, some actions are ‘self-forming’ (Kane 1999), a notion not unrelated to G. Strawson’s condition of self-causation. Chisholm and O’Connor, on the other hand, postulate agent-causes, another attempt to secure radical freedom, this time by positing a *sui generis* kind of uncaused cause manifested only in agency (Chisholm 1964; O’Connor 2003). Both sets of views reflect the incompatibilist intuition that many have claimed is an element of folk views of freedom (Nahmias et al. 2006), an intuition not captured by the Strawsonian picture.

Clearly, even if one accepts the RR, one's view of whether moral desert is possible will depend upon one's theories of free will and responsibility. Importantly, theories of free will and responsibility are not just theories about how these concepts are related but are also theories that elaborate the nature of these concepts. The concepts of freedom, responsibility, and control as used in these traditional incompatibilist arguments are 'strong' concepts that involve metaphysically demanding commitments for freedom – such as commitments to contra-causal agentive powers, to violations of natural law, or to a particular role for indeterminism. Only if these demanding conditions are met can people be deserving of moral praise or blame; and only if they are deserving can punishment (and presumably reward) be justified. Hard incompatibilists deny that the metaphysically demanding conditions are met and therefore give up desert or rely upon consequentialist arguments to justify our practices. Libertarians argue that these metaphysically demanding conditions *are* met, and they often maintain that the folk views of free will, responsibility, and desert embrace these conditions.

Such debates are puzzling, as they turn on abstractions such as *freedom*, *determinism*, *responsibility*, and *desert* – terms of art whose philosophical meanings are obscure and whose ordinary uses are assumed but rarely tested (Monroe and Malle 2010). Are ordinary people really expected to concede that their everyday moral practices of blaming and praising are unjustified because some metaphysical thesis about freedom and responsibility is false? If we care at all about 'our' (i.e. people's) moral practices perhaps it is more prudent to carefully explicate the conditions under which an agent, in this concrete world, 'deserves' a moral assignment such as blame. This is the direction we want to explore – first in the way Strawson pioneered it.

### 3. Strawson

In the standard views canvassed above, a desert basis must be something an agent is morally responsible for, and moral responsibility in turn depends upon a metaphysically demanding kind of freedom that is either Libertarian or impossible. P.F. Strawson, however, seems to turn this chain of reasoning on its head, looking to the nature and function of our blaming practices and the attitudes underlying them, rather than to metaphysical notions, for a justification of those practices.

Strawson does not speak explicitly about desert, but he does consider the justification for our practices of holding one another morally accountable. Chief among the questions of Strawson's paper is whether and how notions of moral responsibility and desert can be reconciled with the potential truth of determinism. Then, as now, most traditional philosophical positions saw freedom as the logical link between the issues of determinism and desert – freedom is what saves the notion of being deserving of certain moral treatments from elimination at the hands of determinism. As mentioned before, Strawson characterizes the traditional debate about justification of our moral practices in terms of the optimist and the pessimist. Strawson's pessimist might be characterized as reasoning thus:

- (1) If determinism is true, neither our wills nor our actions are free.
- (2) One can be morally responsible only for actions that are freely willed (or actions or events that stem from actions that are freely willed).
- (3) Therefore, if determinism is true, no one is morally responsible for his actions.
- (4) One can deserve praise or blame only for actions for which one is morally responsible (RR).
- (5) Therefore, if determinism is true, neither praise nor blame is deserved.

The optimist denies these conclusions, for he takes the first premise to be false: Even if determinism is true, our wills and our actions can be free.

Evidently, the main locus of disagreement between these two positions lies in their different conceptions of freedom. Whereas the pessimist thinks that the freedom required to justify our practices is a radical or Libertarian freedom (and therefore incompatible with determinism), the optimist sees free actions as actions done for reasons, done intentionally. This is a typical compatibilist gloss on free will and it is not undermined by determinism. It is also a metaphysically deflationary view of free will, in contrast to the pessimist's. Strawson is sympathetic to the optimist, but he nonetheless sees the force of the pessimist's rejoinder, which points to the notion of desert:

But why does freedom in this sense justify blame, etc.? . . . the only reason you have given for the practices of moral condemnation and punishment in cases where this freedom is present is the efficacy of these practices in regulating behaviour in socially desirable ways. But this is not a sufficient basis, it is not even the right *sort* of basis, for these practices as we understand them. (Strawson 1993, 47–48)

Strawson lets stand the pessimist's claim that justification for moral treatments cannot stem from utility or expediency. Merely showing that praise or blame is instrumentally good does not make it the case that praise and blame are deserved.<sup>8</sup> This is simply the wrong kind of reason to justify engaging in these practices; something else must provide grounds for desert.

As Strawson relates, the underlying assumption of the pessimist is that desert depends on radical freedom:

Thus he [the pessimist] may say, familiarly enough, that the man who is the subject of justified punishment, blame or moral condemnation must really *deserve* it; and then add, perhaps, that, in the case at least where he is blamed for a positive act rather than an omission, the condition of his really deserving blame is something that goes beyond the negative freedoms that the optimist concedes. It is, say, a genuinely free identification of the will with the act. And this is the condition that is incompatible with the truth of determinism. (1993, 47)

This 'genuinely free identification of the will with the act' denotes some kind of metaphysically demanding notion of freedom such as those briefly discussed above.

Strawson does not side with the pessimist and aims to give the optimist more to work with than mere expediency. His positive proposal represents a radical shift away from metaphysics, and toward the actual world and the realities of interpersonal relations in providing justification for our moral practices. Strawson's argument is that our attributions of moral responsibility to others and ourselves are (often) justified, and are justified by our reasonable expectations of one another and the reactive emotions and attitudes that we manifest, which are indispensable and essentially social. He specifically focuses on the importance of the 'non-detached attitudes' – the felt (not abstracted or theoretical) attitudes that we have or are disposed to have simply in virtue of engaging in social interaction with other human beings. These reactive attitudes, such as resentment for harms motivated by ill will and gratitude for acts motivated by beneficence, not only provide a kind of emotional scaffolding that makes interpersonal interaction possible, but they are also particularly sensitive to precisely the kinds of considerations that seem to make a difference to whether our moral assessments of others are deserved or not.

Importantly, these reactive attitudes are sensitive to the presence or absence of certain kinds of capacities in others, such as the capacity to appreciate moral norms, to conform

one's behavior to considered desires, and to rationally assess options. The absence of these capacities in the psychologically abnormal or the morally undeveloped makes these agents incapable of taking part in the normal adult web of social – moral interactions and likewise removes them as candidates for our normal practices of moral assessment: we no longer hold them responsible or blame them, nor do we think it appropriate to subject them to reactive attitudes of resentment or expect them to react in appropriate ways to moral sanctioning. These conditions for suspension of reactive attitudes constitute what we refer to as *eligibility criteria* for holding people responsible: an agent must possess these capacities in order to be a proper target of moral assessment.

Strawson views these capacities as so foundational to human interaction that he denies that we could have normal interpersonal relations without them. The indispensability of the reactive attitudes for human relations, and their internal logic or structure, form the justificatory base for Strawson's argument, but the justification is internal to the practice and not dependent on some outside factor. Critically, these abilities and capacities are neither dependent on some metaphysical notion of freedom, nor on a gloss on freedom that is threatened by determinism, such as the ability to do otherwise. As several scholars have since argued (e.g. Fara 2008; Sie 2005; Smith 2003), capacity claims are robust even in a deterministic world, made so by the 'raft of counterfactuals' true of the same subject in nearby possible worlds, which are conceptual, not nomological, constructs. Thus, if modal robustness is independent of determinism, it could ground the claim that people possess capacities that they may not exercise and they possess them even if determinism is true.<sup>9</sup>

Strawson also recognizes another kind of case in which people suspend or modify their ordinary reactive attitudes toward wrongdoers: in particular circumstances in which they recognize excuses for what would otherwise be blameworthy behavior. These are cases in which an agent harms without intent and knowledge, or under compulsion from outside forces. In such cases people are not tempted to suspend their general reactive attitudes toward the agent, but they rather perceive the injury as an inappropriate trigger for their reactive attitudes. The patterns for such perceptions in particular cases illuminate what we have called the *assignment criteria* for holding people morally responsible.<sup>10</sup>

In sum, Strawson's radical argument is that the conditions that legitimate or ground moral responsibility are just those conditions that we ordinarily take to be necessary for holding people accountable for their actions – conditions that reliably guide our reactive attitudes. The ability to both hold such attitudes and properly respond to the attitudes of others forms a foundational scaffolding for our social and moral practices, and it is this scaffolding that supplies the justificatory ground for praise and blame.

Many find it surprising that freedom is seldom mentioned in Strawson's paper, notwithstanding its title. Indeed, once he has characterized the debates of the optimist and pessimist, the words 'freedom' or 'free will' do not again occur in *Freedom and Resentment*. This is no mere oversight: it signals the surprising shift away from a justification of blame and praise in terms of metaphysics. Instead, moral responsibility as a human practice takes center stage. One is left to conclude that freedom either is irrelevant to the question of responsibility or that it is stripped of its metaphysical content. Freedom is just what is sufficient for responsibility: the satisfaction of conditions of capacity and mental state that Strawson describes.

### 3.1 Does Strawson's argument succeed?

The upshot of the Strawsonian argument is the denial that a metaphysically demanding notion of freedom is necessary for justified moral assessment (i.e. desert). However, does

Strawson succeed in his arguments? As we have argued, the RR is widely thought to be a condition of moral desert, and freedom is taken to be in turn a prerequisite for moral responsibility. Strawson's own argument is conservative regarding the RR: he does not explicitly state that justified praise and blame entail moral responsibility but he seems to assume the connection. Suppose we are persuaded by Strawson that moral responsibility is a property that humans have, not in virtue of some intrinsic property conferring a metaphysically robust notion of freedom but rather in virtue of their position as agents (with certain capacities and mental states) in a web of social interactions. Does it then follow that agents embedded in our social world are *ipso facto* (at least sometimes) deserving of praise and blame?

Here it depends on how one reads Strawson's argument. Strawson argues that the reactive attitudes we feel toward others and ourselves, and the standards of expectation they reflect, are in fact the justifiers for our moral practices. These practices, which make possible normal human interaction, form a coherent and self-contained system within which all necessary justification resides. As Strawson seems to intend it, the bases of our reactive or appraising attitudes just are the bases of desert.<sup>11</sup> By explaining why our reactive attitudes and the practices built upon them are justified in virtue of their role in mediating human relationships, we simultaneously give an explanation of and justification for moral responsibility. And with the assumed RR, we then have reason to believe that those attributions are justified, and thus deserved.<sup>12</sup>

However compelling his argument may be, there is something nonetheless troubling about Strawson's picture. We could, for instance, take Strawson's argument to show merely that the existence of our appraising attitudes or reactive attitudes functions as generally reliable *evidence* for the presence of a desert base. As such, the fact that we do praise or admire those who do good, or blame and feel retributive urges toward those who act with ill will, provides *prima facie* reason to think that people do in fact often deserve praise and blame for their actions. However, it still remains open to us to question just how reliable such evidence is and in what circumstances it is defeasible. Perhaps our attitudes are a worse guide to desert than Strawson suggests.

More worrisome still, one might question how psychological states like people's reactive attitudes or expectations of others can legitimately serve as justifiers for, say, inflicting suffering on those others. It may seem more likely that deserving blame or punishment would depend on the presence or absence of some intrinsic property of the target of that response – and that is what Strawson seems to imply as well in his discussion of the target's capacities and mental states. It is this kind of intuition that makes conditions like free will seem promising for grounding desert but at the same time opens the door for the possibility that lack of freedom (in a particular case or in all cases, as hard incompatibilists may have it) could be a defeating condition for desert claims. Indeed, even if we accept that people are morally responsible in Strawson's sense, we can still question whether they are *deserving* (in the 'basic desert' sense) of blame or punishment for wrongdoing. The pessimist may admit that our moral practices are necessarily founded in our reactive attitudes yet nonetheless maintain that this alone does not justify their role. For even if desert requires moral responsibility – as the RR suggests – moral responsibility may not be *sufficient* for moral approbation or sanction to be deserved. According to the RR, desert entails moral responsibility *and not vice versa*. For one not willing to absorb Strawson's picture from the inside, as it were, his arguments may appear less than convincing. Even if the pessimist accepts that reactive attitudes and moral practices are indispensable for normal social relations, he may still insist on a further necessary condition for desert – perhaps a metaphysically robust freedom – and deny it holds if

determinism is true. Arguments about the absence of freedom in a deterministic world then serve as a *modus tollens* for both the justification of our reactive attitudes and for the possibility of desert.

On this way of looking at it, what Strawson's argument aims to show (and succeeds in showing, according to many) is that many of our ascriptions of moral responsibility are justified: that it is simply false to claim that no one is ever morally responsible or that no one ever deserves anything. However, it seems we are not thereby licensed to conclude that anyone ever *does* deserve praise or blame, reward or punishment. Because moral responsibility is a necessary, but not sufficient, condition for desert, we cannot conclude from the fact that people can be morally responsible that there is something they therefore do deserve. To guarantee desert, Strawson would need to argue that moral responsibility entails desert (no moral responsibility without desert), or that the conditions for moral responsibility are the same as the conditions for desert. The latter seems to be implicit in Strawson's picture but he does not argue for the claim.

Table 1. Possible relationships between freedom, responsibility and desert.

	Metaphysical freedom	Moral responsibility	Desert
(1) Hard incompatibilist	×	×	×
(2) Libertarian			
(3) Strawson (charitable reading)	×		
(4) Strawson (critical reading)	×		×

To elaborate, take the following three properties: metaphysical freedom, moral responsibility (being a suitable target of moral assessments), and desert (e.g. deserving moral assessments such as blame). How are these related? As depicted in Table 1, the hard incompatibilists (1), such as van Inwagen and G. Strawson, argue that we lack metaphysical freedom and are thus not morally responsible, and that therefore no one ever deserves blame or punishment. The Libertarians (2), in contrast, argue that we have metaphysical freedom and are thus morally responsible, and that we therefore have appropriate grounds for desert. On a charitable reading, P.F. Strawson (3), in contrast, maintains that even if we lack metaphysical freedom, in general, people are nonetheless morally responsible, and he implies that the notion of desert is unproblematic. However, Strawson owes us an argument that moral responsibility alone guarantees desert. Thus, it remains possible that metaphysical freedom is a necessary condition for desert, even if it is not a condition for moral responsibility.<sup>13</sup> So one may conclude from a critical reading of Strawson (4) that his arguments go no further than showing that we are morally responsible, but in a way that does not support desert. It is this worry that leads some to think that Strawson's view, though a vindication of our practices of holding each other responsible, does not suffice to *justify* praise, blame, reward, and punishment.

It seems, then, that we have reached an impasse. Strawson implies, but does not explicitly argue, that his account of reactive attitudes and the role they play in moral practices do more than partly constitute our moral assessments – that they also render praise, blame, reward, and punishment *deserved*. The pessimist can still maintain that these practices are not justified.

However, Strawson went only part way in taking a Strawsonian approach – in offering an analysis of phenomena such as freedom, responsibility, and desert as grounded in actual

human practice. Can we do better if we take seriously the spirit of Strawson's proposal and explore the conditions for justified moral assessment in everyday moral practices? These practices reveal the actual criteria that people consider when deciding whether someone deserves blame. If these criteria for desert are not metaphysically suspect, then determinism is no threat to the possibility of desert. Any 'deeper' justification of the practices, if that is desired, may be another topic of investigation.

#### 4. Recent experimental work on free will and moral practices

Strawson's game-changing move in *Freedom and Resentment* was to pay close attention to people's actual moral practices and the beliefs and attitudes that underlie them, rather than to look for abstract metaphysical justifications. However, to really embrace Strawson's exhortation we need to know more about those practices, which have conceptual, linguistic, and social components. Of particular importance are the *conceptual* assumptions people make in their moral assessments because these assumptions may reveal whether they require robust freedom of the will for moral responsibility, whether they take responsibility to be a necessary condition of desert, or what other criteria of desert are heeded. Strawson himself was still firmly based in his armchair when analyzing these conceptual assumptions. Whether we want to vindicate or eliminate the human practice of moral assessment, we need to get out of the armchair to document the actual practice. Accordingly, we now turn to some recent empirical studies of the folk concept of free will and the criteria underlying people's social practice of blame.

Although a turn to the empirical is philosophically controversial, it is also philosophically valuable. There is an inherent tension in many philosophical arguments about such concepts as free will and responsibility. On the one hand, philosophers would like to talk about the 'things themselves' – what free will really is, what moral responsibility really is – but at the same time, in order to be relevant to the normative and practical concerns we have about regulating behavior, justifying sanctions, or bettering our institutions, we cannot simply abandon the folk conceptions underlying these terms. If, for example, the philosophical concept of freedom is too far removed from people's intuitive or pretheoretical concepts, philosophers run the risk of changing the subject and not addressing the deep philosophical questions that ordinary humans encounter when contemplating their lives. Certainly philosophers such as Strawson, van Inwagen, Pereboom, and Kane mean to speak to universal questions and intend their arguments to enter into our practical deliberations. Any kind of forceful argument about the existence or nonexistence of free will, or about the lack of justification for our practices of blaming and punishing on account of a suspect assumption of free will, must therefore link up to the folk conceptions of responsibility, desert, and free will. Otherwise, what scholars claim or deny is irrelevant to ordinary people.

The commonly held assumption that a robust conception of desert depends upon a metaphysically robust notion of freedom serves as the starting point. The Libertarian response to such an assumption is to ground freedom in the ability to choose or act in ways undetermined by prior causes, or to postulate that agents are uncaused causes, are self-caused, or to claim that actions stem from a soul that operates free of natural law. The Hard Incompatibilist, on the other hand, denies that any such metaphysically robust notion is tenable. However, are ordinary notions of desert and responsibility hostage to such strong metaphysical commitments? Recent research in psychology suggests they are not. We will discuss two main findings. First, people's concept of free will is not the metaphysically burdened notion that many assume must underlie conceptions of desert.

Second, people systematically analyze morally relevant behavior in terms of a set of causal and mental criteria (e.g. intentionality, capacity, obligation); and as a function of whether these criteria are met they judge that an agent *deserves* a certain amount of moral blame. What grounds desert is the satisfaction of those psychological criteria – including the ones that Strawson pointed to but did not systematically analyze – without the need for any metaphysical assumption of free will.

#### 4.1 *The folk concept of free will*

Monroe and Malle (2010) suggest that ordinary people hold a pragmatic rather than a metaphysical view of free will. In one study, people were asked to ‘explain in a few lines what you think it means to have free will’. The convergent folk meaning of free will was to make a *choice* in line with one’s *desires* and *free of constraints*. Notably, not a single person mentioned the involvement of a soul, uncaused causes, or any form of indeterminism. Independently, Stillman, Baumeister, and Mele (2011) found very similar conceptual components of free will. The researchers asked participants to describe behaviors they felt were either ‘of their own free will’ or ‘not the result of free will’. Participants in the ‘free will’ condition reported behaviors associated with pursuing *desired goals*, making *choices*, and acting *against external forces* (e.g. temptation or social pressures), whereas participants in the ‘no free will’ condition wrote about behaviors in the presence of a powerful authority figure (i.e. strong *constraints*).

In two recent studies, Monroe, Dillon, and Malle (2013) examined which of several properties attributable to agents predict (a) judgments of free will (probed by ‘Does the agent have free will?’) and (b) moral judgments (assessed by averaging blame judgments for seven immoral actions the agent performed, such as pulling a dog’s tail or selling drugs to high-school students). The candidate properties to predict these judgments were as follows:

- Can the agent intervene on the normal causal flow of the universe?
- Was the agent the sole cause of his action?
- Does the agent have the capacity for choice?
- Can the agent act intentionally?
- Does the agent have a soul?

Participants were randomly assigned to make judgments about these properties, free will, and blame for one of five agents:

- a normal human;
- an ‘akratic’ human (who has ‘a brain disorder called “Crick’s Syndrome”, characterized by a person’s complete inability to use his thoughts to control his actions”);
- a cyborg – a human brain in a robot body;
- an artificial intelligence in a human body; or
- an advanced robot.

Two findings are of particular interest here. First, people’s free will ascriptions were strongly and almost exclusively predicted by the capacities for choice and intentional action. Second, people’s blame judgments were also strongly and almost exclusively predicted by the capacities for choice and intentional action. No notable predictive power came from having free will, having a soul, or being able to intervene on the causal flow of the universe.

Thus, ordinary ascriptions of free will do not seem to presuppose more than basic cognitive capacities of choice and intentionality, and blame judgments rely on those very same capacities, without any unique considerations of free will. This is in stark contrast to the nature and role of freedom posited by a good number of philosophers.<sup>14</sup>

The capacities of choice and intentionality are *eligibility* criteria for an agent to be the target of moral blame. However, people have even more sophisticated *assignment* criteria: they systematically track causal relations, obligations, and mental states to determine how much blame an agent should receive. Uncovering these criteria is a paradigm case of the analysis that Strawson promotes for reactive attitudes: examining ‘the particular conditions in which they do or do not seem natural or reasonable or appropriate’ (Strawson 1993, 50) – the conditions, that is, that make blame *deserved*. We briefly describe recent work that identifies such conditions.

#### 4.2 Folk judgments of blame

Integrating decades of research on responsibility and blame attributions, a recent model of blame has emerged that specifies what concepts, information, and judgments underlie ordinary assignments of blame (Guglielmo, Monroe, and Malle 2009; Malle, Guglielmo, and Monroe 2012; Figure 1). According to this model, the initial steps in people’s assignment of blame are:

- (1) Detecting an event that violated a norm or value;
- (2) Judging that an *agent* caused this event.<sup>15</sup>

However, humans are not satisfied with merely establishing that an agent *caused* a negative event. They immediately pose a fundamental question of human social cognition (Malle 2008) as a third step:

- (3) Asking whether the agent *intentionally* brought about the event.

Once this judgment has been made, two very different paths of information processing lead up to a final assessment of blame:

- (4a) If the agent is judged to have acted intentionally, perceivers consider the agent’s reasons for acting and blame is graded depending on the justification these reasons provide.
- (4b) If, however, the agent is judged to have brought about the event unintentionally, perceivers consider whether the agent had an *obligation* to prevent the negative outcome and had the *capacity* to attempt such prevention. Only if both of those conditions hold is the agent blamed.

Such information processing will not always be conscious and can be influenced by biased assumptions and premature conclusions, but there is ample evidence that people are reliably sensitive to this kind of information in arriving at their graded blame judgments (Malle, Guglielmo, and Monroe 2012). Moreover, people are expected to take this information into account when blaming others – acts of blame that blatantly ignore such information will be considered unfair or inappropriate. Thus, the practice of blaming is itself regulated by norms whose violation can lead to social sanctions.

It should be noted that the model does not accord the concept of responsibility per se any role in the emergence of blame, even though previous social psychological models,

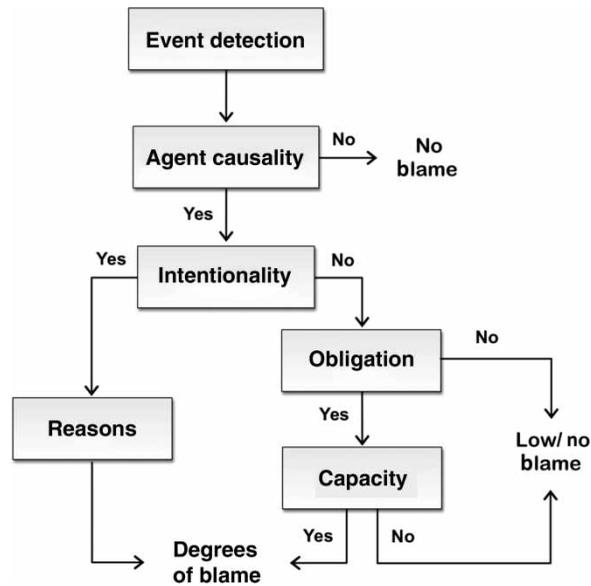


Figure 1. Concepts and processing paths in the path model of blame. Obligation = obligation to prevent the event in question; Capacity = capacity to prevent the event in question. Source: Adapted from Malle, Guglielmo, and Monroe (2012).

and certainly the philosophical literature, have focused a great deal on responsibility. The reason is simple: the term *responsibility* is hopelessly equivocal (Fincham and Jaspars 1980; Hart 1968; Sousa 2009). It collapses distinct phenomena under a single label and is often confounded with other phenomena. A recent study shows at least four distinct constructs subsumed under or co-measured with responsibility: wrongfulness, causality, foreknowledge, and intentionality (Gailey and Falk 2008). In addition, the term responsibility has been used to refer to an agent's obligation (Hamilton 1986), intentionality and justification (Fincham and Bradbury 1992), blame (e.g. Quigley and Tedeschi 1996; Shultz, Schleifer, and Altman 1981), and – most prominent in the philosophical literature – eligibility for moral assessment (Oshana 2001).<sup>16</sup>

Strawsonians may welcome these findings about ordinary people's conceptions of free will and blame as they seem to strengthen their position. First, free will does not seem to be a core concept in folk psychology but a summary label for a triad of more important elements: choice, intentional action in line with one's desires, and absence of serious constraints. None of these components is *prima facie* incompatible with determinism, though a deeper analysis of each may be needed to understand their assumptions.<sup>17</sup> Second, blame – arguably the central moral judgment – is a complex cognitive and social practice that relies on a number of constituent concepts (e.g. causality, intentionality, prevention, obligation), none of which, again, is *prima facie* incompatible with determinism. Importantly, the abstract and puzzling concepts of 'freedom' and 'responsibility' do not even have to be mentioned in an analysis of the social practice of blaming. One might suspect that the deep questions surrounding freedom and responsibility are introduced primarily by philosophical investigations. Ordinary people are concerned with finding out whether someone caused a negative outcome and had an obligation to prevent the outcome, or intentionally violated a norm but perhaps had justified reasons. The joint consideration of these criteria guides *deserved* blame.

**5. Desert without freedom and without ‘moral responsibility’**

What does all this tell us about desert? Feinberg (1970) laid important groundwork by distinguishing desert from related concepts such as *entitlement* and *qualification*. Genuinely *moral* desert, as we see it, refers to the decision whether a person deserves<sup>18</sup> blame or praise and, if so, how much. The concept of *deserving* plays no special role here over and above the criteria that people consider when deciding on a proper moral response – heeding norms, causality, intentionality, mental states, and so on. Ordinary people seem to assume that if blame (or praise) is appropriate, given the abovementioned conditions, it is *ipso facto* deserved. Perhaps there is no justification for the entire practice, because justification normally occurs *as part of* the practice – which is what Strawson appears to have argued. If further justification is desired, the practice would have to be measured against deeper values such as fairness and justice (and perhaps some cultures have more just practices of blame and punishment than others). But even then, the spirit of *Freedom and Resentment* reminds us, we are not best served by worrying about freedom of the will and determinism.

On this analysis, the RR quietly drops out of the picture as a central principle governing desert. The criteria people attend to for attributing moral praise and blame resemble those that scholars have cited as necessary for ‘moral responsibility’, so that when people judge that an agent deserves praise or blame they would also judge the agent as morally responsible. However, because of its polysemy the concept of responsibility makes no unique conceptual contribution. People can express some key moral claims in the language of responsibility, but they can also substitute more precise, unequivocal terms:

- It is your responsibility [→ obligation] to protect the environment.
- The mentally ill cannot be considered morally responsible [→ eligible for ordinary moral assessment].
- He isn’t responsible for [→ did not cause] the mess.
- If the company fails to survive, the upper management is going to be responsible [→ is to be blamed].

We suggest, then, that moral desert, as understood and practiced by ordinary people, gets by without freedom (at least without a metaphysically burdened kind of freedom) and without relying upon responsibility as a necessary conceptual link. Thus, we can add the last line to our Table 1:

Table 1 (revised). Possible relationships between freedom, responsibility, and desert.

	Metaphysical freedom	Moral responsibility	Desert
(1) Hard incompatibilist	×	×	×
(2) Libertarian	✓	✓	✓
(3) Strawson (charitable reading)	×	✓	✓
(4) Strawson (critical reading)	×	✓ <sup>a</sup>	×
+ (5) Ordinary human	×		✓

<sup>a</sup>We avoid the check mark here to indicate that desert is not predicated on an abstract moral responsibility but rather on a collection of criteria. People may also attribute moral responsibility whenever these criteria are fulfilled, so that judgments of desert are paralleled by judgments of responsibility; but we do not view desert as dependent on the equivocal concept of moral responsibility, as the RR does.

Downloaded by [Brown University] at 21:35 01 July 2013

Is our analysis deflationary? Yes and no. It is in the sense of deflating bloated concepts of ‘free will’ and ‘responsibility’ that appear not to play any significant role in ordinary moral understanding and practices. However, it is not in the sense of deflating ordinary moral practices themselves. Strawson’s program was to elevate reactive attitudes to be the justifiers of moral assessments – to render praise, blame, reward, and punishment *deserved*. We suggest that the justifiers of moral assessments are more than just reactive and they are more than just attitudes. They are sophisticated judgments in light of characteristic criteria, and if the criteria are met, the ensuing judgments are considered justified – justified within the social community that subscribes to those criteria. This means that people sometimes get it wrong. They can assign undeserving blame or praise when, for example, they do not properly heed the criteria or rely on false information when assessing the criteria. However, in general, communities can tell deserving from undeserving moral assessments, because the criteria for desert are socially learned, socially shared, and socially enforced.

To sum up, then, if we look both to people’s actual practices of morally assessing others’ behavior and to the nature of the concepts that they invoke to ground such assessments, we see that these practices do not rely on an impossible kind of freedom that leaves us responsible for nothing, but instead on basic capacities for intentional, rational action. Desert itself is a placeholder for a kind of calculation of an input-output balance, whereby the input is typically an agent’s behavior and the output is a social (sometimes worldly) response. ‘Deserving’ are matches between input and output, and what creates matches is governed by norms, expectations, common practices, and considerations of agentic capacities; no metaphysical assumptions of free will appear to be involved. This is very much in the spirit of what Strawson argued half a century ago. Both Strawson’s optimist and his pessimist are mistaken, for desert neither requires radical freedom or any other metaphysically strong assumptions about agency, nor does it require instrumental justification. For moral assessments like blame to be deserved (in this real world) people merely presuppose a collection of everyday capacities, such as the capacity to reason and to choose and to act intentionally, and the presence of various causal–mental relations between an agent and the morally relevant event. And those capacities and relations are just certain types of causal event chains that a determinist could describe, if so desired, as ‘following from the laws of nature and the preceding events’.

### Acknowledgements

The authors would like to thank the Big Questions in Free Will project, supported by the John Templeton Foundation/FSU Research Foundation (Prime Award No. 15462/Subaward SCI05 and Subaward TU03). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. We also thank the Princeton University Center for Human Values, George Sher, Derk Pereboom, and anonymous reviewers for helpful comments on an earlier draft.

### Notes

1. In this paper, ‘Strawson’ will always refer to P.F. Strawson. When we mention Galen Strawson, we will always indicate this by explicitly referring to him as such or as G. Strawson.
2. The act that is the desert base is usually thought by many to precede desert, but see Feldman (1995) for an alternative view.
3. With respect to moral responsibility, it is useful to distinguish between what Pereboom (2001) calls ‘basic’ and ‘non-basic’ desert, where desert is basic in the sense that an agent would deserve blame just in virtue of having performed the action in the way in which he did and not just in virtue of consequentialist or contractualist considerations. Only the former is truly

at issue in the context of determinism and the free will debate, and it is this notion with which we concern ourselves here.

4. The idea of criteria can be pushed in interesting directions. For example, in addition to moral desert, we also talk of people deserving respect, admiration, compensation, etc. – all of which vary in their criteria of justification and certainly differ from criteria of moral assignments of blame, praise, etc. Framing desert in this way allows a unified account of moral and nonmoral desert.
5. In other words, for a moral assessment (e.g. praise, blame) of an agent to be deserved, the object of the assessment (e.g. a behavior, an outcome) must be something that the agent was morally responsible for.
6. There are complications in determining exactly what is involved in being morally responsible for an event – with control being only one, and an imperfect, index of responsibility. We will take up this thorny issue further below.
7. Not everyone agrees with the RR (Feldman 1995, 1996). For example, Feldman and Vilhauer (Feldman 1995; Vilhauer 2009) have argued that personhood is a desert base that is independent of responsibility. People are not responsible for being people; that is just something that happened to them. However, the argument goes, just in virtue of being persons, individuals deserve respect, autonomy, and basic human rights. Thus, some things are deserved even if the subject has no control over the desert base. However, a careful analysis of desert suggests that human rights are properly seen as entitlements, not objects of moral desert. Freiman and Nichols (2011), after introducing the brute luck constraint, used findings from a survey to question whether moral common sense really is committed to this constraint. They found that when participants were asked in the abstract to ‘suppose that some people make more money than others solely because of genetic advantages’, participants judged this as unfair and that those people did not deserve the extra money, consistent with the brute luck constraint. By contrast, when asked about a concrete case, in which two specific individuals were described as having differential monetary success because of their different talents or capacities, which they had solely because of differences in their genetics, participants made quite different judgments: They judged that the more talented individual did deserve greater financial rewards, thus seemingly rejecting the brute luck constraint. However, the concrete case in this survey was described as a rich causal chain: from genetics to dispositions to effort to action to popularity to monetary reward, whereas the abstract case described monetary rewards directly because of genetics. Thus, the abstract case met none of the candidate criteria of responsibility whereas the concrete case met several. Moreover, monetary reward for talent-based performance is arguably not a moral desert but a societal entitlement. Thus, it is doubtful whether these results call into question the RR.
8. Some may disagree with this anti-pragmatic conception of desert. If one believes that social utility has an evolutionary dimension, a species that has these kinds of concepts and practices might survive more successfully in increasingly large communities than a species that does not have those concepts and practices. Our moral concepts and practices may have allowed the growth of culture from small to large communities. While this does not demonstrate that the practice is just or fair within a broader normative framework, it is a plausible naturalistic account that clarifies the genesis and utility of the practices.
9. More recently, there has been some debate regarding whether the kinds of abilities that are compatible with determinism are the ones necessary for free will and/or moral responsibility (see Clarke 2009).
10. Thus, although Strawson does not frame his argument in terms of eligibility and assignment criteria, his discussion of capacities on the one hand, and of the excusing conditions for suspending our reactive attitudes on the other, nicely illustrates the two classes of criteria.
11. Others have followed Strawson in this belief. Miller, for instance, claims that the basis for appraising attitudes (a class of attitudes that include the reactive attitudes Strawson discusses, but may extend more widely) coincides with bases for desert (Miller 1976). His view, like Strawson’s, is that the triggers for our appraising attitudes are sufficient justification for holding them.
12. That Strawson considers this internal justification sufficient is evident when, for example, he writes:

It might be said that all this leaves the real question unanswered . . . It is a question about the rational justification of ordinary inter-personal attitudes in general. To this I shall

reply, first, that such a question could seem real only to one who had utterly failed to grasp the purport of the preceding answer, the fact of our natural human commitment to ordinary inter-personal attitudes. This commitment is part of the general framework of human life, not something that can come up for review as particular cases can come up for review within this general framework. (Strawson 1993)

13. See, for example, Parfit (2011), who argues that even if moral responsibility exists, we cannot deserve to suffer. Scanlon (2010) makes a similar point.
14. Some experiments seem inconsistent with this finding. For example, Nichols and Knobe (2007) suggest that, under certain circumstances, ordinary people give up their commitment to moral attributions if freedom is threatened by a deterministic universe. These circumstances, however, are entirely removed from ordinary moral practice. When pushed to *imagine* a universe in which ‘each decision has to happen the way that it does’ and asked whether it is ‘possible for a person to be fully morally responsible for their actions’, they are willing to say *No*. However, when their normal moral thinking is engaged by a concrete case of a man who ‘stabs his wife and children to death so that he can be with his secretary’, then even in a deterministic universe, 50–72% of respondents insist that the man is ‘fully morally responsible’. (See also Roskies and Nichols 2008, in which people were willing to ascribe responsibility and blame when the deterministic universe was our own.) Many questions can be raised about such studies, including the considerable experimental demands, ambiguous terms (e.g. ‘fully morally responsible?’), limited response options, and confounds of transgression severity. However, we certainly consider it the right approach to clarify people’s assumptions and interpretations empirically, rather than ‘from the armchair’. Philosophical manipulation arguments (e.g. Pereboom 2001) have also been used to challenge compatibilist views; however, they tend to complicate the picture by adding other intentional agents.
15. If the event is a behavior, the agent judgment comes for free.
16. Eligibility considerations are located in the model in at least two places: as part of the causal agency node (‘Was the event caused by an *eligible* agent?’) and as part of the capacity component. See Malle, Guglielmo, and Monroe (2012, 316, 318), for more details.
17. Is the folk conception of *choice* consistent or inconsistent with determinism? Some of emerging data (Monroe, Dillon, and Malle 2013) suggest that ascriptions of choice presuppose having options or reasonable alternatives, which then raises questions about the status of such ‘alternative possibilities’. Choice among options may involve the intelligibility of counterfactuals, and counterfactuals (such as ‘he could have avoided the harmful outcome’) are reconcilable with determinism. Moreover, counterfactuals may psychologically function as mental simulations, serving as much the understanding of past behaviors as the guidance for similar future behaviors. However, no empirical data are currently available on these issues.
18. In many cases in which we ordinarily use the term ‘deserve’ (e.g. deserving a day off, a raise, an answer), we in fact mean ‘is entitled to’, which is a distinct concept (Feinberg 1970). However, there is a second, derivative meaning of *deserve* that characterizes responses other than social blame or praise that an agent faces – such as being struck by misfortune after a wrongdoing or unexpected fortune after a life of selfless dedication. Such cases of ‘poetic justice’ are best analyzed as substituting a worldly event of a certain degree of valence for a social – moral blame or praise response of that degree of valence. Presumably people go through very similar information processing of the agent’s (im)moral act (in terms of causal – mental criteria) as in the case of human moral responses, and they apply some calculus of ‘proportional’ worldly outcomes to determine whether the agent ‘deserved’ a given outcome. Proportional means nothing but a reliable correlation (holding constant causal – mental criteria) between the rank-ordered severity of moral norm violations and the rank-ordered severity of worldly outcomes (obviously, these rank orderings will be influenced by historical, cultural, and idiosyncratic factors). This secondary meaning of *deserving* thus exhibits the same structure as the primary meaning: an agent’s (im)moral behavior is analyzed in terms of causal – mental criteria that dictate an appropriate, justified, and thus *deserving* (here, worldly) response.

### Notes on contributors

Adina L. Roskies is Professor of Philosophy at Dartmouth College. She received a PhD in neuroscience and cognitive science from the University of California, San Diego, in 1995, and a PhD in

philosophy from MIT in 2004. Dr Roskies' philosophical research interests lie at the intersection of philosophy and neuroscience and include philosophy of mind, philosophy of science, and ethics.

Bertram F. Malle was trained in psychology, philosophy, and linguistics at the University of Graz, Austria, before receiving his PhD in psychology at Stanford University in 1995. He is Professor of Cognitive, Linguistic, and Psychological Sciences at Brown University. His research focuses on theory of mind, behavior explanations, and moral judgments.

## References

- Baiasu, S. 2006. The normative pluralism of desert. *Journal of Political Philosophy* 14: 226–37.
- Chisholm, R.M. 1964. *Human freedom and the self. The Lindley lecture*. Lawrence, KS: Department of Philosophy, University of Kansas.
- Clarke, R. 2009. Dispositions, abilities to act, and free will: The new dispositionalism. *Mind* 118: 323–51.
- Fara, M. 2008. Masked abilities and compatibilism. *Mind* 117: 843–65. doi: 10.1093/mind/fzn078
- Feinberg, J. 1970. Justice and personal desert. In *Doing and deserving: Essays in the theory of responsibility*, 55–94. Princeton, NJ: Princeton University Press.
- Feldman, F. 1995. Desert: Reconsideration of some received wisdom. *Mind* 104: 63–77.
- Feldman, F. 1996. Responsibility as a condition for desert. *Mind* 105: 165–8.
- Fincham, F.D., and T.N. Bradbury. 1992. Assessing attributions in marriage: The relationship attribution measure. *Journal of Personality and Social Psychology* 62: 457–68. doi: 10.1037/0022-3514.62.3.457
- Fincham, F.D., and J.M. Jaspars. 1980. Attribution of responsibility: From man the scientist to man as lawyer. *Advances in Experimental Social Psychology* 13: 81–138.
- Freiman, C., and S. Nichols. 2011. Is desert in the details? *Philosophy and Phenomenological Research* 82: 121–33.
- Gailey, J.A., and R.F. Falk. 2008. Attribution of responsibility as a multidimensional concept. *Sociological Spectrum* 28: 659–80. doi: 10.1080/02732170802342958
- Guglielmo, S., A.E. Monroe, and B.F. Malle. 2009. At the heart of morality lies folk psychology. *Inquiry: An Interdisciplinary Journal of Philosophy* 52: 449–66.
- Hamilton, V.L. 1986. Chains of command: Responsibility attribution in hierarchies. *Journal of Applied Social Psychology* 16: 118–38. doi: 10.1111/j.1559-1816.1986.tb02283.x
- Hart, H.L.A. 1968. *Punishment and responsibility*. New York: Oxford University Press.
- van Inwagen, P. 1986. *An essay on free will*. New York: Oxford University Press.
- Kane, R. 1999. Responsibility, luck, and chance: Reflections on free will and indeterminism. *Journal of Philosophy* 96: 217–40.
- Kershner, S. 2008. Desert tracks character alone. *International Journal of Applied Philosophy* 22: 71–88.
- Kleinig, J. 1971. The concept of desert. *American Philosophical Quarterly* 8: 71–8.
- Malle, B.F. 2008. The fundamental tools, and possibly universals, of social cognition. In *Handbook of motivation and cognition across cultures*, ed. R.M. Sorrentino and S. Yamaguchi, 267–96. New York: Elsevier/Academic Press.
- Malle, B.F., S. Guglielmo, and A.E. Monroe. 2012. Moral, cognitive, and social: The nature of blame. In *Social thinking and interpersonal behaviour*, ed. J. Forgas, K. Fiedler and C. Sedikides, 313–31. Philadelphia, PA: Psychology Press.
- McLeod, O. 2008. Desert. In *The Stanford encyclopedia of philosophy*, ed. E.N. Zalta. <http://plato.stanford.edu/archives/win2008/entries/desert/%3E>.
- Miller, D. 1976. *Social justice*. Oxford: The Clarendon Press.
- Monroe, A.E., K.D. Dillon, and B.F. Malle. 2013. Developing a model of the folk concept of free will and its impact on moral judgment. Paper presented at the symposium on big questions in Free Will, Jan. 13, 2013, in Florida State University, Tallahassee, FL.
- Monroe, A.E., and B.F. Malle. 2010. From uncaused will to conscious choice: The need to study, not speculate about, people's folk concept of free will. *Review of Philosophy and Psychology* 1: 211–24.
- Nahmias, E., S. Morris, T. Nadelhoffer, and J. Turner. 2006. Is incompatibilism intuitive? *Philosophy and Phenomenological Research* 73: 28–53.
- Nichols, S., and J. Knobe. 2007. Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous* 41: 663–85.

- O'Connor, T. 2003. Agent causation. In *Free will*, 2nd ed. ed. G. Watson, 257–84. Oxford: Oxford University Press.
- Oshana, M. 2001. Responsibility: Philosophical aspects. In *International encyclopedia of the social & behavioral sciences*, ed. N.J. Smelser and P.B. Baltes, 13279–83. Oxford: Pergamon.
- Parfit, D. 2011. *On what matters*. New York: Oxford University Press.
- Pereboom, D. 2001. *Living without free will*. Cambridge: Cambridge University Press.
- Quigley, B.M., and J.T. Tedeschi. 1996. Mediating effects of blame attributions on feelings of anger. *Personality and Social Psychology Bulletin* 22: 1280–8. doi: 10.1177/01461672962212008
- Roskies, A.L., and S. Nichols. 2008. Bringing moral responsibility down to earth. *Journal of Philosophy* 105: 371–88.
- Sadurski, W. 1985. *Giving desert its due: Social justice and legal theory*. Dordrecht: Reidel.
- Scanlon, T.M. 2010. *Moral dimensions: Permissibility, meaning, blame*. Cambridge, MA: Harvard University Press.
- Sher, G. 1989. *Desert*. Princeton, NJ: Princeton University Press.
- Shultz, T.R., M. Schleifer, and I. Altman. 1981. Judgments of causation, responsibility, and punishment in cases of harm-doing. *Canadian Journal of Behavioural Science* 13: 238–53. doi: 10.1037/h0081183
- Sie, M. 2005. *Justifying blame: Why free will matters and why it does not*. Amsterdam: Rodopi.
- Smith, M. 2003. Rational capacities, or: How to distinguish recklessness, weakness, and compulsion. In *Weakness of will and practical irrationality*, ed. S. Stroud and C. Tappolet, 17–38. New York: Clarendon Press.
- Sousa, P. 2009. A cognitive approach to moral responsibility: The case of a failed attempt to kill. *Journal of Cognition and Culture* 9: 171–94. doi: 10.1163/156770909x12489459066183
- Stillman, T.F., R.F. Baumeister, and A.R. Mele. 2011. Free will in everyday life: Autobiographical accounts of free and unfree actions. *Philosophical Psychology* 24: 381–94. doi: 10.1080/09515089.2011.556607
- Strawson, G. 1986. *Freedom and belief*. Oxford: Clarendon.
- Strawson, G. 1994. The impossibility of moral responsibility. *Philosophical Studies* 75: 5–24.
- Strawson, G. 2008. Free agents In *Real Materialism and Other Essays*, 359–386. New York: Oxford University Press.
- Strawson, P.F. 1993. Freedom and resentment. In *Perspectives on moral responsibility*, ed. J.M. Fischer and M. Ravizza, 45–66. Ithaca, NY: Cornell University Press.
- Vilhauer, B. 2009. Free will skepticism and personhood as a desert base. *Canadian Journal of Philosophy* 39: 489–512.
- Wallace, R.J. 1998. *Responsibility and the moral sentiments*. Boston, MA: Harvard University Press.