



Stanford Encyclopedia of Philosophy Archive

Summer 2014 Edition

Moral Responsibility

First published Sat Jan 6, 2001; substantive revision Wed Mar 26, 2014

When a person performs or fails to perform a morally significant action, we sometimes think that a particular kind of response is warranted. Praise and blame are perhaps the most obvious forms this reaction might take. For example, one who encounters a car accident may be regarded as worthy of praise for having saved a child from inside the burning car, or alternatively, one may be regarded as worthy of blame for not having used one's mobile phone to call for help. To regard such agents as worthy of one of these reactions is to regard them as responsible for what they have done or left undone. (These are examples of other-directed ascriptions of responsibility. The reaction might also be self-directed, e.g., one can recognize oneself to be blameworthy). Thus, to be morally responsible for something, say an action, is to be worthy of a particular kind of reaction—praise, blame, or something akin to these—for having performed it.^[1]

Though further elaboration and qualification of the above characterization of moral responsibility is called for and will be provided below, this is enough to distinguish concern about this form of responsibility from some others commonly referred to through use of the terms 'responsibility' or 'responsible.' To illustrate, we might say that higher than normal rainfall in the spring is responsible for an increase in the amount of vegetation or that it is the judge's responsibility to give instructions to the jury before they begin deliberating. In the first case, we mean to identify a causal connection between the earlier amount of rain and the later increased vegetation. In the second, we mean to say that when one assumes the role of judge, certain duties, or obligations, follow. Although these concepts are connected with the concept of moral responsibility discussed here, they are not the same, for in neither case are we directly concerned about whether it would be appropriate to react to some candidate (here, the rainfall or a particular judge) with something like praise or blame.^[2]

Philosophical reflection on moral responsibility has a long history. One reason for this persistent interest is the way the topic seems connected with our conception of ourselves as 'persons.'^[3] Many have held that one distinct feature of persons is their status as morally responsible agents, a status resting—some have proposed—on a special kind of control that only they can exercise. Many who view persons in this way have wondered whether their special status is threatened if certain other claims about our universe are true. For example, can a person be morally responsible for her behavior if that behavior can be explained solely by reference to physical states of the universe and the laws governing changes in those physical states, or solely by reference to the existence of a sovereign God who guides the world along a divinely ordained path? Concerns of this sort have often motivated reflection on moral responsibility.

A comprehensive theory of moral responsibility would elucidate the following: (1) the concept, or idea, of moral responsibility itself; (2) the criteria for being a moral agent, i.e., one who qualifies generally as an agent open to responsibility ascriptions (e.g., only beings possessing the general capacity to evaluate reasons for acting can be moral agents); (3) the conditions under which the concept of moral responsibility is properly applied, i.e., those conditions under which a moral agent is responsible for a particular something (e.g., a moral agent can be responsible for an action she has performed only if she performed it freely, where acting freely entails the ability to have done otherwise at the time of action); and finally 4) possible objects of responsibility ascriptions (e.g., actions, omissions, consequences, character traits, etc.). Although each of these will be touched upon in the discussion below (see, e.g., the brief sketch of Aristotle's account in the next section), the primary focus of this entry is on the first component—i.e., the concept of moral responsibility. The section immediately following this

introduction is a discussion of the origin and history of Western reflection on moral responsibility. This is followed by an overview of recent work on the concept of moral responsibility. For further discussion of issues associated with moral responsibility, see the related entries below.

- [1. Some Historical Background](#)
 - [2. Recent Work on the Concept of Responsibility](#)
 - [2.1 Strawson and the Reactive Attitudes](#)
 - [2.2 Developments After Strawson](#)
 - [Bibliography](#)
 - [Academic Tools](#)
 - [Other Internet Resources](#)
 - [Related Entries](#)
-

1. Some Historical Background

What follows in this section is a brief outline of the origins and trajectory of reflection on moral responsibility in the Western philosophical tradition. Against this background, a distinction will be drawn between two conceptions of moral responsibility that have exerted considerable influence on subsequent thinkers.

An understanding of the concept of moral responsibility and its application is present implicitly in some of the earliest surviving Greek texts, i.e., the Homeric epics (circa 8th century BCE but no doubt informed by a much earlier oral tradition).^[4] In these texts, both human and superhuman agents are often regarded as fair targets of praise and blame on the basis of how they have behaved, and at other times, an agent's behavior is excused because of the presence of some factor that has undermined his/her control (Irwin 1999: 225). Reflection on these factors gave rise to *fatalism*—the view that one's future or some aspect of it is predetermined, e.g., by the gods, or the stars, or simply some facts about truth and time—in such a way as to make one's particular deliberations, choices and actions irrelevant to whether that particular future is realized (recall, e.g., the plight of Oedipus). If some particular outcome is fated, then it seems that the agent concerned could not be morally responsible for that outcome. Likewise, if fatalism were true with respect to all human futures, then it would seem that no human agent could be morally responsible for anything. Though this brand of fatalism has sometimes exerted significant historical influence, most philosophers have rejected it on the grounds that there is no good reason to think that our futures are fated in the sense that they will unfold no matter what particular deliberations we engage in, choices we make, or actions we perform.

Aristotle (384–323 BCE) seems to have been the first to construct a theory of moral responsibility.^[5] In the course of discussing human virtues and their corresponding vices, Aristotle pauses in *Nicomachean Ethics* III.1–5 to explore their underpinnings. He begins with a brief statement of the concept of moral responsibility—that it is sometimes appropriate to respond to an agent with praise or blame on the basis of her actions and/or dispositional traits of character (1109b30–35). A bit later, he clarifies that only a certain kind of agent qualifies as a moral agent and is thus properly subject to ascriptions of responsibility, namely, one who possess a capacity for decision. For Aristotle, a decision is a particular kind of desire resulting from deliberation, one that expresses the agent's conception of what is good (1111b5–1113b3). The remainder of Aristotle's discussion is devoted to spelling out the conditions under which it is appropriate to hold a moral agent blameworthy or praiseworthy for some particular action or trait. His general proposal is that one is an apt candidate for praise or blame if and only if the action and/or disposition is voluntary. According to Aristotle, a voluntary action or trait has two distinctive features. First, there is a control condition: the action or trait must have its origin in the agent. That is, it must be up to the agent whether to perform that action or possess the trait—it cannot be compelled externally. Second, Aristotle proposes an epistemic condition: the agent must be aware of what it is she is doing or bringing about (1110a–1111b4).^[6]

There is an instructive ambiguity in Aristotle's account of responsibility, an ambiguity that has led to competing interpretations of his view. Aristotle aims to identify the conditions under which it is appropriate to praise or

blame an agent, but it is not entirely clear how to understand the pivotal notion of appropriateness in his conception of responsibility. There are at least two possibilities: a) praise or blame is appropriate in the sense that the agent *deserves* such a response, given his behavior and/or traits of character; or b) praise or blame is appropriate in the sense that such a reaction is likely to bring about a desired consequence, namely an improvement in the agent's behavior and/or character. These two possibilities may be characterized in terms of two competing interpretations of the concept of moral responsibility: 1) the *merit-based view*, according to which praise or blame would be an appropriate reaction toward the candidate if and only if she merits—in the sense of ‘deserves’—such a reaction; vs. 2) the *consequentialist view*, according to which praise or blame would be appropriate if and only if a reaction of this sort would likely lead to a desired change in the agent and/or her behavior.^[7]

Scholars disagree about which of the above views Aristotle endorsed, but the importance of distinguishing between them grew as philosophers began to focus on a newly conceived threat to moral responsibility. While Aristotle argued against a version of fatalism (*On Interpretation*, ch. 9), he may not have recognized the difference between it and the related possible threat of *causal determinism*. Causal determinism is the view that everything that happens or exists is caused by sufficient antecedent conditions, making it impossible for anything to happen or be other than it does or is. One variety of causal determinism, *scientific determinism*, identifies the relevant antecedent conditions as a combination of prior states of the universe and the laws of nature. Another, *theological determinism*, identifies those conditions as being the nature and will of God. It seems likely that theological determinism evolved out of the shift, both in Greek religion and in Ancient Mesopotamian religions, from polytheism to belief in one sovereign God, or at least one god who reigned over all others. The doctrine of scientific determinism can be traced back as far as the Presocratic Atomists (5th cent. BCE), but the difference between it and the earlier fatalistic view seems not to be clearly recognized until the development of Stoic philosophy (3rd. cent. BCE). Though fatalism, like causal determinism, might seem to threaten moral responsibility by threatening an agent's control, the two differ on the significance of human deliberation, choice, and action. If fatalism is true, then human deliberation, choice, and action are completely otiose, for what is fated will transpire no matter what one chooses to do. According to causal determinism, however, one's deliberations, choices, and actions will often be necessary links in the causal chain that brings something about. In other words, even though our deliberations, choices, and actions are themselves determined like everything else, it is still the case, according to causal determinism, that the occurrence or existence of yet other things depends upon our deliberating, choosing and acting in a certain way (Irwin 1999: 243–249; Meyer 1998: 225–227; and Pereboom 1997: ch. 2).

Since the Stoics, the thesis of causal determinism, if true, and its ramifications, have taken center stage in theorizing about moral responsibility. During the Medieval period, especially in the work of Augustine (354–430) and Aquinas (1225–1274), reflection on freedom and responsibility was often generated by questions concerning versions of theological determinism, including most prominently: a) Does God's sovereignty entail that God is responsible for evil?; and b) Does God's foreknowledge entail that we are not free and morally responsible since it would seem that we cannot do anything other than what God foreknows we will do? During the Modern period, there was renewed interest in scientific determinism—a change attributable to the development of increasingly sophisticated mechanistic models of the universe culminating in the success of Newtonian physics. The possibility of giving a comprehensive explanation of every aspect of the universe—including human action—in terms of physical causes became much more plausible. Many thought that persons could not be free and morally responsible if such an explanation of human action turned out to be true. Others argued that freedom and responsibility would not be undermined by the truth of scientific determinism. In keeping with this focus on the ramifications of causal determinism for moral responsibility, thinkers may be classified as being one of two types: 1) an *incompatibilist* about causal determinism and moral responsibility—one who maintains that if causal determinism is true, then there is nothing for which one can be morally responsible; or 2) a *compatibilist*—one who holds that a person can be morally responsible for some things, even if both who she is and what she does is causally determined.^[8] In Ancient Greece, these positions were exemplified in the thought of Epicurus (341–270 BCE) and the Stoics, respectively.

Above, an ambiguity in Aristotle's conception of moral responsibility was highlighted—that it was not clear whether he endorsed a merit-based vs. a consequentialist conception of moral responsibility. The history of

reflection on moral responsibility demonstrates that how one interprets the concept of moral responsibility strongly influences one's overall account of moral responsibility. For example, those who accept the merit-based conception of moral responsibility have tended to be incompatibilists. That is, most have thought that if an agent were to genuinely merit praise or blame for something, then he would need to exercise a special form of control over that thing (e.g., the ability at the time of action to both perform or not perform the action) that is incompatible with one's being causally determined. In addition to Epicurus, we can cite early Augustine, Thomas Reid (1710–1796), and Immanuel Kant (1724–1804) as historical examples here. Those accepting the consequentialist conception of moral responsibility, on the other hand, have traditionally contended that determinism poses no threat to moral responsibility since praising and blaming could still be an effective means of influencing another's behavior, even in a deterministic world. Thomas Hobbes (1588–1679), David Hume (1711–1776), and John Stuart Mill (1806–1873) are, along with the Stoics, representatives of this view. This general trend of linking the consequentialist conception of moral responsibility with compatibilism about causal determinism and moral responsibility and the merit-based conception with incompatibilism continued to persist through the first half of the twentieth century.

2. Recent Work on the Concept of Responsibility

The issue of how best to understand the concept of moral responsibility is important, for it can strongly influence one's view of what, if any, philosophical problems might be associated with the notion, and further, if there are problems, what might count as a solution. As discussed above, philosophical reflection on moral responsibility has historically relied upon one of two broad interpretations of the concept: 1) the *merit-based view*, according to which praise or blame would be an appropriate reaction toward the candidate if and only if she merits—in the sense of ‘deserves’—such a reaction; or 2) the *consequentialist view*, according to which praise or blame would be appropriate if and only if a reaction of this sort would likely lead to a desired change in the agent and/or her behavior. Though versions of the consequentialist view have continued to garner support (Smart; Frankena 1963: ch. 4; Schlick 1966; Brandt 1992; Dennett 1984: ch. 7; Kupperman 1991: ch. 3; and Vargas 2013: ch. 6), work in the last 50 years on the concept of moral responsibility has increasingly focused on: a) offering alternative versions of the merit-based view; and b) questioning whether there is but one concept of moral responsibility.

Increased attention to the stance of regarding and holding persons morally responsible has generated much of the recent work on the concept of moral responsibility. All theorists have recognized features of this practice—inner attitudes and emotions, their outward expression in censure or praise, and the imposition of corresponding sanctions or rewards. However, most understood the inner attitudes and emotions involved to rest on a more fundamental theoretical judgment about the agent's *being* responsible. In other words, it was typically assumed that blame and praise depended upon a judgment, or belief (pre-reflective in most cases), that the agent in question had satisfied the objective conditions on being responsible. These judgments were presumed to be independent of the inner attitudinal/emotive states involved in holding responsible in the sense that reaching such judgments and evaluating them required no essential reference to the attitudes and emotions of the one making the judgment. For the holder of the consequentialist view, this is a judgment that the agent exercised a form of control that could be influenced through outward expressions of praise and blame in order to curb or promote certain behaviors. For those holding the merit view, it is a judgment that the agent has exercised the requisite form of metaphysical control, e.g., that she could have done otherwise at the time of action (Watson 1987: 258).

If holding responsible is best understood as resting on an independent judgment about being responsible, then it is legitimate to inquire whether such underlying judgments and their associated outward expressions can be justified, as a whole, in the face of our best current understanding of the world, e.g., in the face of evidence that our world is possibly deterministic. According to incompatibilists, a judgment that someone is morally responsible could never be true if the world were deterministic; thus praising and blaming in the merit-based sense would be beside the point. Compatibilists, on the other hand, contend that the truth of determinism would not undermine the relevant underlying judgments concerning the efficacy of praising and blaming practices, thereby leaving the rationale of such practices intact.

2.1 Strawson and the Reactive Attitudes

In his landmark essay, 'Freedom and Resentment,' P. F. Strawson (1962) sets out to adjudicate the dispute between those compatibilists who hold a consequentialist view of responsibility and those incompatibilists who hold the merit-based view.^[9] Both are wrong, Strawson believes, because they distort the concept of moral responsibility by sharing the prevailing assumption sketched above — the assumption that holding persons responsible rests upon a theoretical judgment of their being responsible. According to Strawson, the attitudes expressed in holding persons morally responsible are varieties of a wide range of attitudes deriving from our participation in personal relationships, e.g., resentment, indignation, hurt feelings, anger, gratitude, reciprocal love, and forgiveness. The function of these attitudes is to express "...how much we actually mind, how much it matters to us, whether the actions of other people—and particularly *some* other people—reflect attitudes towards us of good will, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other." (p. 5, author's emphasis) These attitudes are thus *participant reactive attitudes*, because they are: a) natural attitudinal reactions to the perception of another's good will, ill will, or indifference (pp. 4–6), and b) expressed from the stance of one who is immersed in interpersonal relationships and who regards the candidate held responsible as a participant in such relationships as well (p. 10).^[10]

The reactive attitudes can be suspended or modified in at least two kinds of circumstances, corresponding to the two features just mentioned. In the first, one might conclude that, contrary to first appearances, the candidate did not violate the demand for a reasonable degree of good will. For example, a person's behavior may be *excused* when one determines that it was an accident, or one may determine that the behavior was *justified*, say, in the case of an emergency when some greater good is being pursued. In the second kind of circumstance, one may abandon the participant perspective in relation to the candidate. In these cases, one adopts the *objective standpoint*, one from which one ceases to regard the individual as capable of participating in genuine personal relations (either for some limited time or permanently). Instead, one regards the individual as psychologically/morally abnormal or undeveloped and thereby a candidate, not for the full range of reactive attitudes, but primarily for those objective attitudes associated with treatment or simply instrumental control. Such individuals lie, in some sense or to some varying extent, outside the boundaries of the moral community. For example, we may regard a very young child as initially exempt from the reactive attitudes (but increasingly less so in cases of normal development) or adopt the objective standpoint in relation to an individual we determine to be suffering from severe mental illness (P. F. Strawson 1962: 6–10; Bennett: 40; Watson 1987: 259–260; R. Jay Wallace: chs. 5–6).

The central criticism Strawson directs at both consequentialist and traditional merit views is that both have over-intellectualized the issue of moral responsibility—a criticism with which many subsequent thinkers have wrestled.^[11] The charge of over intellectualization stems from the traditional tendency to presume that the rationality of holding a person responsible depends upon a judgment that the person in question has satisfied some set of objective requirements on being responsible (conditions on efficacy or metaphysical freedom) and that these requirements themselves are justifiable. Strawson, by contrast, maintains that the reactive attitudes are a natural expression of an essential feature of our form of life, in particular, the interpersonal nature of our way of life. The practice, then, of holding responsible—embedded as it is in our way of life—"neither calls for nor permits, an external 'rational' justification" (p. 23). Though judgments about the appropriateness of particular responses may arise (i.e., answers to questions like: Was the candidate's behavior really an expression of ill will?; or Is the candidate involved a genuine participant in the moral sphere of human relations?), these judgments are based on principles internal to the practice. That is, their justification refers back to an account of the reactive attitudes and their role in personal relationships, *not to some independent theoretical account of the conditions on being responsible*.

Given the above, Strawson contends that it is pointless to ask whether the practice of holding responsible can be rationally justified if determinism is true. This is either because it is not psychologically possible to divest ourselves of these reactions and so continually inhabit the objective standpoint, or even if that were possible, because it is not clear that rationality could ever demand that we give up the reactive attitudes, given the loss in quality of life should we do so. In sum, Strawson attempts to turn the traditional debate on its head, for now

judgments about being responsible are understood in relation to the role reactive attitudes play in the *practice of holding responsible*, rather than the other way around. Whereas judgments are true or false and thereby can generate the need for justification, the desire for good will and those attitudes generated by it possess no truth value themselves, thereby eliminating any need for an external justification (Magill 1887: 21; Double 1996b: 848).

Strawson's concept of moral responsibility yields a compatibilist account of being responsible but one that departs significantly from earlier such accounts in two respects. First, Strawson's is a compatibilist view by default only. That is, on Strawson's view, the problem of determinism and freedom/responsibility is not so much *resolved* by showing that the objective conditions on being responsible are consistent with one's being determined but rather *dissolved* by showing that the practice of holding people responsible relies on no such conditions and therefore needs no external justification in the face of determinism. Second, Strawson's is a merit-based form of compatibilism. That is, unlike most former consequentialist forms of compatibilism, it helps to explain why we feel that some agents deserve our censure or merit our praise. They do so because they have violated, met, or exceeded our demand for a reasonable degree of good will.

2.2 Developments After Strawson

Most agree that Strawson's discussion of the reactive attitudes is a valuable contribution to our understanding of the practice of holding responsible, but many have taken issue with his contentions about the insular nature of that practice, namely that a) since propriety judgments about the reactive attitudes are strictly internal to the practice (i.e., being responsible is defined in relation to the practice of holding responsible), their justification cannot be considered from a standpoint outside that practice; and b) since the reactive attitudes are natural responses deriving from our psychological constitution, they cannot be dislodged by theoretical considerations. Responding to the first of these, some have argued that it does seem possible to critique existing practices of holding responsible from standpoints outside them. For example, one might judge that either one's own existing community practice or some other community's practice of holding responsible ought to be modified (Fischer and Ravizza 1993: 18; Ekstrom 2000: 148–149). If such evaluations are legitimate, then, contrary to what Strawson suggested, it seems that an existing practice can be questioned from a standpoint external to it. In other words, being responsible cannot be explicated strictly in terms of an existing practice of holding responsible. This then, would suggest a possible role to be played by independent theoretical conditions on being responsible, conditions which could prove to be compatibilist or incompatibilist in nature.

Objecting to the second of Strawson's anti-theory contentions, some have argued that incompatibilist intuitions are embedded in the reactive attitudes themselves so that these attitudes cannot persist unless some justification can be given of them, or more weakly, that they cannot but be disturbed if something like determinism is true. Here, cases are often cited where negative reactive attitudes seem to be dispelled or mitigated upon learning that an agent's past includes severe deprivation and/or abuse. There is a strong pull to think that our reactive attitudes are altered in such cases because we perceive such a background to be deterministic. If this is the proper interpretation of the phenomenon, then it is evidence that theoretical considerations, like the truth of determinism, could in fact dislodge the reactive attitudes (Nagel 1986: 125; Kane 1996: 84–89; Galen Strawson 1986: 88; Honderich 1988: vol. 2, ch. 1; and replies by Watson 1987: 279–286 and 1996: 240; and McKenna 1998).

Versions of Strawson's view continue to be very ably defended, and shortly, more will be said about the significant way in which his work continues to shape contemporary discussion of the concept of responsibility. However, many have taken objections of the above sort to be decisive in undermining the most radical of Strawson's anti-theory claims. Incompatibilists, in particular, seem largely unpersuaded and so have continued to assume a more or less traditional merit-based conception of moral responsibility as the basis for their theorizing. A number of compatibilists also remain unconvinced that Strawson has successfully shown independent theoretical considerations to be irrelevant to ascriptions of responsibility. It is noteworthy that some of these have accorded the reactive attitudes a central role in their discussions of the concept of responsibility, resulting in new merit-based versions of compatibilism (see e.g., Fischer and Ravizza 1998, and McKenna 2012).

Until recently philosophers have assumed that they were concerned about a shared concept of moral responsibility. Even when controversy increasingly arose over how best to characterize it, the assumption seems to have been that it was a controversy over the one correct way of characterizing the concept of responsibility. Strawson was certainly amongst those who made this assumption in trying to adjudicate the dispute between those compatibilists who held the consequentialist view of responsibility and incompatibilists who held the merit-based view. However, a number of authors have suggested of late that at least some disagreements about the most plausible overall theory might be based on a failure to distinguish between different but related concepts of responsibility.

Broadly speaking, a distinction has been made between responsibility as accountability and responsibility as attributability.^[12] Drawing on Strawson's work, many contemporary accountability theorists maintain that to be responsible is to be an apt candidate for the reactive attitudes (Bennett 1980; Wallace 1994; Watson 1996; Fischer & Ravizza 1998; and Darwall 2006). In other words, an agent is responsible, if and only if it is appropriate for us to hold her responsible, or accountable, via the reactive attitudes. This highlights a main theme in Strawson--namely, that our responsibility practices are inherently social. Through the reactive attitudes (e.g., resentment) we communicate to fellow members of the moral community our interpersonal expectation for a reasonable degree of goodwill (Stern 1974; Watson 1987/1996; McKenna 1998/2012; Darwall 2006; Shoemaker 2007).^[13]

Since the reactive attitudes--when expressed and accompanied by their associated practices--may have consequences for the well-being of an agent (especially in the case of those blaming attitudes and practices involved in holding someone accountable for wrong-doing), they would seem to be appropriate only if it is fair that the agent be subject to them in the sense that s/he deserve them.^[14] This concern about fairness may be the original source of the merit-based view of responsibility. Relatedly, this line of thought may help explain the historical preoccupation with whether responsibility for an action requires the ability to have done otherwise. That is, the normative concern for a fair opportunity to avoid blame and sanction may lie behind the felt need to have access to alternatives. (Zimmerman: ch. 5; Wallace: 103–117; Watson 1996: 238–9; Magill 1997: 42–53; Nelkin 2012:31–50).

Notably, some accounts of responsibility make no essential reference to the reactive attitudes or their accompanying practices. Perhaps the clearest example of these are so-called “ledger” views of moral responsibility. According to such views, the practice of ascribing responsibility involves assigning a credit or debit to a metaphorical ledger associated with each agent (Feinberg: 30–1; Glover: 64; Zimmerman: 38–9; and discussion of such views in Watson 1987: 261–2; and Fischer & Ravizza 1998: 8–10, nt. 12). In other words, an agent is responsible if a fault or credit is properly attributable to her.

Ledger views belong to a broader class of views which regard responsibility to be a matter of proper attributability. As Gary Watson has highlighted, the central concern in such views is whether the agent's action or attitude discloses her evaluative judgments or commitments (1996). Satisfying some baseline conditions of responsibility as attributability would appear to be necessary in order to be responsible in the sense of accountable. For example, it would seem unfair to hold someone accountable for an action via reactive attitudes such as resentment or indignation, if the action was not properly attributable to the agent--say, because she succumbed to a genuinely coercive psychological compulsion. Yet being responsible in the attributability sense is not sufficient for being responsible in the accountability sense. As Watson points out, it may make no sense to hold the agent responsible for the action in question, since it may not be the sort of thing for which they are accountable to us. For example, one may think that in making a career decision, an acquaintance failed to give due consideration to what would most fully develop and exercise his talents. Though this is not a moral judgment in the narrow sense favored by accountability theorists (that is, it is unconnected to any interpersonal demand, or mutual expectation, of the sort presupposed by the reactive attitudes) it is a case of finding fault in the way an agent has exercised his judgment. If responsibility as accountability and attributability can come apart in this way, then there appear to be at least two distinct concepts of responsibility.^[15]

Some views fit uneasily in either of these two categories. For example, according to another influential view, someone is responsible for an action or attitude just in case it is connected to her capacity for evaluative

judgment in a way that opens her up, in principle, to demands for justification from others (Oshana 1997; Scanlon 1998; and Smith 2005/2008/2012). Such a view—call it the "answerability" model—appears to combine aspects of the attributability and accountability models (see discussion by Watson 2011 and Shoemaker 2012). The self-disclosure aspect of the attributability model is reflected in emphasizing that the target of appraisal must be judgment-sensitive. The interpersonal emphasis characteristic of Strawson-inspired accountability models is reflected in the demand for justification (though answerability theorists tend to reject a necessary connection between these demands and the reactive attitudes). In this way, the answerability model offers the possibility of re-unifying discussions of responsibility (Smith 2012), but some see further grounds for distinguishing an additional sense of responsibility (Shoemaker 2012).

The recognition of diversity within the concept or amongst concepts of moral responsibility has generated new reflection on whether the conditions on being morally responsible are in tension with one another (Nagel 1986; G. Strawson 1986, 105–117, 307–317; Honderich 1988: vol. 2, ch. 1; Double 1996a: chs. 6–7; Bok 1998: ch. 1; Smilansky 2000: ch. 6); For example, some have argued that while a compatibilist sense of freedom is sufficient for attributability, genuine accountability would require that agents be capable of exercising libertarian freedom. A rapidly expanding body of empirical data on folk intuitions about freedom and responsibility has added fuel to this debate (Nahmias et. al. 2005 and 2007; Vargas 2006; Nichols & Knobe 2007; Nelkin 2007; Roskies & Nichols 2008; and Knobe & Doris 2010).

If there are irreconcilable tensions within the concept of responsibility, then the conditions of its application cannot be jointly satisfied. Of course, there have always been those—e.g., hard determinists—who have concluded that the conditions on being morally responsible cannot be met and thus that no one is ever morally responsible. However, a noteworthy new trend amongst both contemporary hard determinists and others who conclude that the conditions for the applicability of our folk concept cannot be jointly satisfied has been the move to offer a revisionist conception of moral responsibility (or something analogous to moral responsibility) and its associated practices rather than to reject talk about being responsible outright. Revisionism about moral responsibility is a matter of degree. Some revisionists seek to salvage much if not most of what they take to be linked to the folk concept (Dennett 1984: 19; Honderich 1988: vol. 2, ch. 1; Scanlon 1998: 274–277; Vargas 2004/2005/2013), while others offer more radical reconstructions of the concept and associated practices (Smart 1961; Pereboom 2001: 199–212; Smilansky 2000: chps. 7–8; Kelly 2002).^[16]

Conclusion

It is difficult to overestimate the influence of Strawson's work on the topic of moral responsibility. The resurgence of interest in metaphysical treatments of freedom and moral responsibility in recent years is a sign that most have not been persuaded by his most radical critique of such approaches. Nevertheless, his enduring influence is reflected in the ongoing rich discussion of the place and role of the reactive attitudes in human life and in the way contemporary theorists situate their models of responsibility in relation to the accountability model, which he helped to define.

Bibliography

- Adkins, A.W.H., 1960. *Merit and Responsibility: A Study in Greek Values*, Oxford: Clarendon Press.
- Aristotle, 1985. *The Nicomachean Ethics*, trans. by Terence Irwin, Indianapolis: Hackett Publishing Co..
- , 1984. *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. Jonathan Barnes, 2 vols., Princeton: Princeton University Press.
- Arpaly, Nomy, 2003. *Unprincipled Virtue: An Inquiry into Moral Agency*, New York: Oxford University Press.
- , 2006. *Merit, Meaning, and Human Bondage: An Essay on Free Will*, Princeton: Princeton University Press.
- Austin, J.L., 1979. "A Plea for Excuses" in *Philosophical Papers*, J.O. Urmson and G.J. Warnock, eds., New York: Oxford University Press.
- Ayer, A.J., 1980. "Free Will and Rationality" in van Stratten (ed.) 1980.
- Baier, Kurt, 1991. "Types of Responsibility." in *The Spectrum of Responsibility*, Peter French, ed., New York: St. Martin's Press.

- Berofsky, Bernard, ed., 1966. *Free Will and Determinism*, New York: Harper & Row.
- Bennett, Jonathan, 1980. "Accountability" in van Stratten 1980.
- Bobsien, Susanne, 2001. *Determinism and Freedom in Stoic Philosophy*, Oxford: Clarendon Press.
- Bok, Hilary, 1998. *Freedom and Responsibility*, Princeton: Princeton University Press.
- Brandt, Richard, 1969. "A Utilitarian Theory of Excuses" *The Philosophical Review*, 78: 337–361. Reprinted in *Morality, Utility, and Rights*, New York: Cambridge University Press, 1992.
- , 1959. *Ethical Theory*, Englewood Cliffs, NJ: Prentice Hall, Inc..
- , 1958. "Blameworthiness and Obligation," in Melden 1958.
- Brickhouse, T.C. 1991. "Roberts on Responsibility for Action and Character in the *Nicomachean Ethics*" *Ancient Philosophy* 11: 137–48.
- Brink, David and Nelkin, Dana, 2013. "Fairness and the Architecture of Responsibility," in Shoemaker 2013.
- Broadie, Sarah, 1991. *Ethics with Aristotle*, New York: Oxford University Press.
- Coates, Justin D. and Tognazzini, Neal, 2012. "The Nature and Ethics of Blame," *Philosophy Compass*, 7: 197–207.
- eds., 2013. *Blame: Its Nature and Norms*, New York: Oxford University Press.
- Curren, Randall, 2000. *Aristotle on the Necessity of Public Education*, New York: Roman & Littlefield.
- , 1989. "The Contribution of *Nicomachean Ethics* iii.5 to Aristotle's Theory of Responsibility," *History of Philosophy Quarterly*, 6: 261–277.
- Dennett, Daniel, 2003. *Freedom Evolves*, New York: Viking Press.
- , 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*, Cambridge, MA: MIT Press.
- Darwall, Stephen, 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, MA: Harvard University Press.
- Doris, John M., 2002. *Lack of Character: Personality and Moral Behavior*, New York: Cambridge University Press.
- Double, Richard, 2000. "Metaethics, Metaphilosophy, and Free Will Subjectivism," in Kane 2002.
- , 1996a. *Metaphilosophy and Free Will*, New York: Oxford University Press.
- , 1996b. "Honderich on the Consequences of Determinism," *Philosophy and Phenomenological Research*, 66 (December): 847–854.
- , 1991. *The Non-reality of Free Will*, New York: Oxford University Press.
- Echenique, Javier, 2012. *Aristotle's Ethics and Moral Responsibility*, New York: Cambridge University Press.
- Ekstrom, Laura Waddell, 2000. *Free Will: A Philosophical Study*, Boulder, CO: Westview Press.
- Eshleman, Andrew, forthcoming. "Worthy of Praise: Responsibility and Better-than-Minimally-Decent Agency" in Shoemaker and Tognazzini forthcoming.
- Everson, Stephen, ed., 1998. *Companions to Ancient Thought 4: Ethics*, New York: Cambridge University Press.
- , 1990. "Aristotle's Compatibilism in the *Nicomachean Ethics*," *Ancient Philosophy* 10: 81–103.
- Feinberg, Joel, 1970. *Doing and Deserving: Essays in the Theory of Responsibility*, Princeton: Princeton University Press.
- Feldman, Fred, 1995. "Desert: Reconsideration of Some Received Wisdom" *Mind*, 104 (January): 63–77.
- Fingarette, Herbert, 1967. *On Responsibility*, New York: Basic Books, Inc.
- Fischer, John Martin, 1999. "Recent Work on Moral Responsibility" *Ethics*, 110 (October): 93–139.
- , 1994. *The Metaphysics of Free Will: An Essay on Control*, Cambridge, MA: Blackwell.
- , ed., 1986. *Moral Responsibility*, Ithaca: Cornell University Press.
- Fischer, John Martin and Ravizza, Mark, 1998. *Responsibility and Control: A Theory of Moral Responsibility*, New York: Cambridge University Press.
- , eds., 1993. *Perspectives on Moral Responsibility*, Ithaca: Cornell University Press.
- Fischer, John Martin and Tognazzini, Neal, 2011. "The Physiognomy of Responsibility" *Philosophy and Phenomenological Research*, 82: 381–417.
- Fischer, J.M., Kane, R., Pereboom, D., and Vargas, M. 2007. *Four Views on Free Will*, Oxford: Blackwell Publishers.
- Frankfurt, Harry, 1969. "Alternate Possibilities and Moral Responsibility," *The Journal of Philosophy*, 66: 828–839.
- Glover, Jonathan, 1970. *Responsibility*, New York: Humanities Press.
- Haji, Ishtiyaque, 2002. "Compatibilist Views of Freedom and Responsibility" in Kane 2002.
- , 1998. *Moral Appraisability: Puzzles, Proposals, and Perplexities*, New York: Oxford University Press.

- Hart, H. L., 1968. *Punishment and Responsibility*, New York: Oxford University Press.
- Helm, Bennett, 2012. "Accountability and Some Social Dimensions of Human Agency," *Philosophical Issues: A Supplement to Noûs*, 22: 217–32.
- Hieronymi, Pamela, 2004. "The Force and Fairness of Blame," *Philosophical Perspectives*, 18: 115–148.
- Honderich, Ted, 2002. "Determinism as True, Both Compatibilism and Incompatibilism as False, and the Real Problem," in Kane 2002.
- , 1996. "Compatibilism, Incompatibilism, and the Smart Aleck," *Philosophy and Phenomenological Research*, 66 (December): 855–862.
- , 1988. *A Theory of Determinism: The Mind, Neuroscience, and Life Hopes*, 2 vols., Oxford: Clarendon Press
- Irwin, Terrance, ed., 1999. *Classical Philosophy*, New York: Oxford University Press.
- , 1980. "Reason and Responsibility in Aristotle," in Rorty 1980.
- Kane, Robert, ed., 2002. *The Oxford Handbook of Free Will*, New York: Oxford University Press.
- , 1996. *The Significance of Free Will*, New York: Oxford University Press.
- Kelly, Erin, 2002. "Doing Without Desert," *Pacific Philosophical Quarterly*, 83: 180–205.
- Kennett, Jeanette, 2001. *Agency and Responsibility: A Common-sense Moral Psychology*, New York: Oxford University Press.
- Knobe, J. and Doris, J., 2010. "Responsibility," in *The Handbook of Moral Psychology*, eds. John Doris, et al., New York: Oxford University Press.
- Kupperman, Joel, 1991. *Character*, New York: Oxford University Press.
- Levy, Neil, 2005. "The Good, the Bad, and the Blameworthy," *Journal of Ethics and Social Philosophy*, 2(1): 2–16.
- Mackie, John L., 1985. "Morality and the Retributive Emotions," in *Persons and Values* (Volume 2), Oxford: Clarendon Press.
- Macnamara, Coleen, 2011. "Holding Others Responsible," *Philosophical Studies*, 152: 81–102.
- Magill, Kevin, 2000. "Blaming, Understanding, and Justification," in T. van den Beld 2000.
- , 1997. *Freedom and Experience: Self-Determination without Illusions*, New York: St. Martins Press.
- McKenna, Michael, 2012. *Conversation and Responsibility*, New York: Oxford University Press.
- , 1998. "The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism," *Journal of Ethics*, 2: 123–142.
- McKenna, Michael and Russell, Paul, eds., 2008. *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson's "Freedom and Resentment"*, Burlington, VT: Ashgate Publishing.
- Melden, A.I., ed., 1958. *Essays in Moral Philosophy*, Seattle: University of Washington Press.
- Meyer, Susan Suave, 1988. "Moral Responsibility: Aristotle and After," in Everson 1998.
- , 1993. *Aristotle on Moral Responsibility*, Cambridge, MA: Blackwell Pub..
- Nagel, Thomas, 1986. *The View From Nowhere*, New York: Oxford University Press.
- Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J. 2005. "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility," *Philosophical Psychology*, 18: 561–584.
- Nahmias, E., Coates, D. Justin, Kvaran, Trevor, 2007. "Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions," *Midwest Studies in Philosophy*, 31: 214–242.
- Nelkin, Dana, 2011. *Making Sense of Freedom and Responsibility*, New York: Oxford University Press.
- , 2007. "Do We Have a Coherent Set of Intuitions About Moral Responsibility?" *Midwest Studies in Philosophy*, 31: 243–259.
- Nichols, Shaun and Knobe, Joshua, 2007. "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions," *Noûs*, 41(4): 663–685.
- Nozick, Robert, 1981. *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- Oshana, Marina, 1997. "Ascriptions of Responsibility," *American Philosophical Quarterly*, 34: 71–83.
- Pereboom, Derk, 2001, *Living Without Free Will*, New York: Cambridge University Press.
- , 2000. "Living Without Free Will: The Case for Hard Compatibilism" in Kane 2000.
- , ed., 1997. *Free Will*, Indianapolis: Hackett Publishing Co..
- Roberts, Jean, 1984. "Aristotle on Responsibility for Action and Character," *Ancient Philosophy*, 9: 23–36.
- Rorty, Amelie Oksenberg, ed., 1980. *Essays on Aristotle's Ethics*, Los Angeles: University of California Press.
- Roskies, A.L., and Nichols, S. 2008. "Bringing Responsibility Down to Earth" *Journal of Philosophy*, 105(7): 371–388.
- Russell, Paul, 2013. "Responsibility, Naturalism, and 'The Morality System'," in Shoemaker 2013.

- , 1995. *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility*, New York: Oxford University Press.
- , 1992. "Strawson's Way of Naturalizing Responsibility," *Ethics*, 102: 287–302.
- Scanlon, T. M., 2008. *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge, MA: Harvard University Press.
- , 1998. *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
- , 1988. "The Significance of Choice," in *The Tanner Lectures on Human Values* (Volume 8), Salt Lake City: University of Utah Press.
- Schlick, Moritz, 1966. "When is a Man Responsible," in Berofsky, 1966.
- Schoeman, Ferdinand, ed., 1987. *Responsibility, Character, and the Emotions*, New York: Cambridge University Press.
- Sher, George, 2006. *In Praise of Blame*, New York: Oxford University Press.
- Shoemaker, David, 2011. "Attributability, Answerability, and Accountability," *Ethics*, 121: 602–32.
- , 2009. "Responsibility and Disability," *Metaphilosophy*, 40: 438–61.
- , 2007. "Moral Address, Moral Responsibility, and the Boundaries of the Moral Community," *Ethics*, 118: 70–108.
- Shoemaker, David, ed., 2013. *Oxford Studies in Agency and Responsibility* (Volume 1), New York: Oxford University Press.
- Shoemaker, David and Tognazzini, Neal, eds., forthcoming. *Oxford Studies in Agency and Responsibility*, Volume 2, New York: Oxford University Press.
- Slote, Michael, 1990. "Ethics Without Free Will," *Social Theory and Practice*, 16: 369–383.
- Smart, J.J.C., 1961. "Free Will, Praise, and Blame," *Mind*, 70: 291–306.
- Smilansky, Saul, 2000. *Free Will and Illusion*, New York: Oxford University Press.
- , 1996. "Responsibility and Desert: Defending the Connection," *Mind*, 105: 157–163.
- Smiley, Marion, 1992. *Moral Responsibility and the Boundaries of Community*, Chicago: University of Chicago Press.
- Smith, Angela M., 2012. "Attributability, Answerability, and Accountability: In Defense of a Unified Account," *Ethics*, 122: 575–589.
- , 2008. "Control, Responsibility, and Moral Assessment," *Philosophical Studies*, 138: 367–392.
- , 2007. "On Being Responsible and Holding Responsible," *The Journal of Ethics*, 11: 465–484.
- , 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics*, 115: 236–71.
- Sommers, Tamler, 2012. *Relative Justice: Cultural Diversity, Free Will, and Moral Responsibility*, Princeton, NJ: Princeton University Press.
- Sorabji, Richard, 1980. *Necessity, Cause, and Blame*, Ithaca: Cornell University Press.
- Stern, Lawrence, 1974. "Freedom, Blame, and the Moral Community," *The Journal of Philosophy*, 71: 72–84.
- Strawson, Galen, 1994. "The Impossibility of Moral Responsibility," *Philosophical Studies*, 75: 5–24.
- , 1986. *Freedom and Belief*, New York: Oxford University Press.
- Strawson, P. F., 1980. "Reply to Ayer and Bennett," in van Stratten 1980.
- , 1962. "Freedom and Resentment," *Proceedings of the British Academy*, 48: 1–25. Reprinted in Fischer and Ravizza, 1993.
- Talbert, Matthew, 2012. "Blame and Responsiveness to Moral Reasons: Are Psychopaths Blameworthy?" *Pacific Philosophical Quarterly*, 89: 516–35.
- Tognazzini, Neal, forthcoming. "Blameworthiness and the Affective Account of Blame," *Philosophia*.
- , 2013. "Responsibility," *The International Encyclopedia of Ethics*, ed. Hugh LaFollette, Malden, MA: Wiley Blackwell:4592–4602.
- van den Beld, T., 2000. *Moral Responsibility and Ontology*, Dordrecht: Kluwer.
- van Inwagen, Peter, 1978. *An Essay on Free Will*, New York: Oxford University Press.
- van Stratten, Z., ed., 1980. *Philosophical Subjects: Essays Presented to P.F. Strawson*, New York: Oxford University Press.
- Vargas, Manuel, 2013. *Building Better Beings: A Theory of Moral Responsibility*, New York: Oxford University Press.
- , 2004. "Responsibility and the Aims of Theory: Strawson and Revisionism," *Pacific Philosophical Quarterly*, 85: 218–241.
- , 2005. "The Revisionist's Guide to Responsibility," *Philosophical Studies*, 125: 399–429.


- , 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.
- Wallace, James, 1974. “Excellences and Merit,” *Philosophical Review*, 83: 182–199.
- Wallace, R.J., 2011. “Dispassionate Opprobrium: On Blame and the Reactive Sentiments,” in Wallace, Kumar, and Freeman 2011.
- , 1994. *Responsibility and the Moral Sentiments*, Cambridge, MA: Harvard University Press.
- Wallace, R.J., Kumar, R., and Freeman, S., eds., 2011. *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, New York: Oxford University Press.
- Waller, Bruce, 2011. *Against Moral Responsibility*, Cambridge, MA: MIT Press.
- Watson, Gary, 2011. “The Trouble with Psychopaths,” in Wallace, Kumar, and Freeman 2011.
- , 1996. “Two Faces of Responsibility,” *Philosophical Topics*, 24: 227–248.
- , 1987. “Responsibility and the Limits of Evil,” in Schoeman 1987.
- Williams, Bernard, 1993. *Shame and Necessity*, Los Angeles: University of California Press.
- Wolf, Susan, 1990. *Freedom Within Reason*, New York: Oxford University Press.
- , 1981. “The Importance of Free Will,” *Mind*, 90: 386–405.
- Zimmerman, Michael, 2010. “Responsibility, Reaction, and Value,” *Journal of Ethics*, 14: 103–115.
- , 1988. *An Essay on Moral Responsibility*, Totowa, NJ: Roman and Littlefield.

Academic Tools

 [How to cite this entry.](#)

 [Preview the PDF version of this entry](#) at the [Friends of the SEP Society](#).

 [Look up this entry topic](#) at the [Indiana Philosophy Ontology Project](#) (InPhO).

 [Enhanced bibliography for this entry](#) at [PhilPapers](#), with links to its database.

Other Internet Resources

- [The Determinism and Freedom Philosophy Website](#) edited by Ted Honderich, University College London.
- [Flickers of Freedom](#) (multiple contributors, coordinated by Thomas Nadelhoffer)

[Please contact the author with other suggestions.]

Related Entries

[Aristotle, General Topics: ethics](#) | [compatibilism](#) | [determinism: causal](#) | [fatalism](#) | [free will](#) | [incompatibilism: \(nondeterministic\) theories of free will](#) | [luck: moral](#) | [responsibility: collective](#)

[Copyright © 2014](#) by
Andrew Eshleman <eshleman@up.edu>

This is a file in the archives of the Stanford Encyclopedia of Philosophy.
Please note that some links may no longer be functional.

Stanford | Center for the Study of
Language and Information

The Stanford Encyclopedia of Philosophy is [copyright © 2014](#) by [The Metaphysics Research Lab](#), Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054