

PDF version of the entry  
Moral Psychology: Empirical Approaches  
<https://plato.stanford.edu/archives/spr2020/entries/moral-psych-emp/>  
from the SPRING 2020 EDITION of the

## STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Edward N. Zalta   Uri Nodelman   Colin Allen   R. Lanier Anderson  
Principal Editor   Senior Editor   Associate Editor   Faculty Sponsor

Editorial Board  
<https://plato.stanford.edu/board.html>

Library of Congress Catalog Data  
ISSN: 1095-5054

**Notice:** This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

*Stanford Encyclopedia of Philosophy*  
Copyright © 2020 by the publisher  
The Metaphysics Research Lab  
Center for the Study of Language and Information  
Stanford University, Stanford, CA 94305

Moral Psychology: Empirical Approaches  
Copyright © 2020 by the authors  
John Doris, Stephen Stich, Jonathan Phillips, and Lachlan Walmsley

All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

## Moral Psychology: Empirical Approaches

*First published Wed Apr 19, 2006; substantive revision Mon Jan 6, 2020*

Moral psychology investigates human functioning in moral contexts, and asks how these results may impact debate in ethical theory. This work is necessarily interdisciplinary, drawing on both the empirical resources of the human sciences and the conceptual resources of philosophical ethics. The present article discusses several topics that illustrate this type of inquiry: thought experiments, responsibility, character, egoism v. altruism, and moral disagreement.

- 1. Introduction: What is Moral Psychology?
- 2. Thought Experiments and the Methods of Ethics
- 3. Moral Responsibility
- 4. Virtue Ethics and Skepticism About Character
- 5. Egoism vs. Altruism
- 6. Moral Disagreement
- 7. Conclusion
- Bibliography
- Academic Tools
- Other Internet Resources
- Related Entries

## 1. Introduction: What is Moral Psychology?

Contemporary moral psychology—the study of human thought and behavior in ethical contexts—is resolutely interdisciplinary: psychologists freely draw on philosophical theories to help structure their empirical research, while philosophers freely draw on empirical findings from psychology to help structure their theories.<sup>[1]</sup>

While this extensive interdisciplinarity is a fairly recent development (with few exceptions, most of the relevant work dates from the past quarter century), it should not be a surprising development. From antiquity to the present, philosophers have not been bashful about making empirical claims, and many of these empirical claims have been claims about human psychology (Doris & Stich 2005). It is therefore unremarkable that, with the emergence of scientific psychology over the past century and a half, some of these philosophers would think to check their work against the systematic findings of psychologists (hopefully, while taking special care to avoid being misled by scientific controversy; see Doris 2015, Chapter 3; Machery & Doris forthcoming).

Similarly, at least since the demise of behaviorism, psychologists have been keenly interested in normative phenomena in general and ethical phenomena in particular. It is therefore unremarkable that some of these psychologists would seek to enrich their theoretical frameworks with the conceptual resources of a field intensively focused on normative phenomena: philosophical ethics. As a result, the field demarcated by “moral psychology”, routinely involves an admixture of empirical and normative inquiry, pursued by both philosophers and psychologists—increasingly, in the form of collaborative efforts involving practitioners from both fields.

For philosophers, the special interest of this interdisciplinary inquiry lies in the ways moral psychology may help adjudicate between competing ethical theories. The plausibility of its associated moral psychology is not, of course, the only dimension on which an ethical theory may be evaluated; equally important are *normative* questions having to do with how well a theory fares when compared to important convictions about such things as justice, fairness, and the good life. Such questions have been, and will continue to be, of central importance for philosophical ethics. Nonetheless, it is commonly supposed that an ethical theory committed to an impoverished or inaccurate conception of moral psychology is at a serious competitive disadvantage. As Bernard Williams (1973, 1985; cf. Flanagan 1991) forcefully argued, an ethical conception that commends relationships, commitments, or life projects that are at odds with the sorts of attachments that can be reasonably expected to take root in and vivify actual human lives is an ethical conception with—at best—a very tenuous claim to our assent.

With this in mind, problems in ethical theory choice making reference to moral psychology can be framed by two related inquiries:

1. What empirical claims about human psychology do advocates of competing perspectives on ethical theory assert or presuppose?
2. How empirically well supported are these claims?

The first question is one of philosophical scholarship: what are the psychological commitments of various positions in philosophical ethics? The second question takes us beyond the corridors of philosophy departments and to the sorts of questions asked, and sometimes answered, by the human sciences, including psychology, anthropology, sociology, history, cognitive science, linguistics and neuroscience. Thus, contemporary moral psychology is *methodologically pluralistic*: it aims to answer philosophical questions, but in an empirically responsible way.

However, it will sometimes be difficult to tell which claims in philosophical ethics require empirical substantiation. Partly, this is because it is sometimes unclear whether, and to what extent, a contention counts as empirically assessable. Consider questions regarding “normal functioning” in mental health care: are the answers to these questions statistical, or evaluative (Boorse 1975; Fulford 1989; Murphy 2006)? For example, is “normal” mental health simply the psychological condition of most people, or is it *good* mental health? If the former, the issue is, at least in principle, empirically decidable. If the latter, the issues must be decided, if they can be decided, by arguments about value.

Additionally, philosophers have not always been explicit about whether, and to what extent, they are making empirical claims. For example, are their depictions of moral character meant to identify psychological features of actual persons, or to articulate ideals that need not be instantiated in actual human psychologies? Such questions will of course be complicated by the inevitable diversity of philosophical opinion.

In every instance, therefore, the first task is to carefully document a theory’s empirically assessable claims, whether they are explicit or, as may often be the case, tacit. Once claims apt for empirical assessment have been located, the question becomes one of identifying any relevant empirical literatures. The next job is to assess those literatures, in an attempt to determine what conclusions can be responsibly drawn from them. Science, particularly social science, being what it is, many conclusions will be provisional; the philosophical moral psychologist must be prepared to adjudicate controversies in other fields, or offer informed conjecture regarding future findings. Often, the empirical record will be crucially incomplete. In such cases, philosophers may be forced to engage in empirically disciplined conjecture, or even to engage in their own empirical work, as some philosophers are beginning to do.<sup>[2]</sup>

When the philosophical positions have been isolated, and putatively relevant empirical literatures assessed, we can begin to evaluate the plausibility of the philosophical moral psychology: Is the speculative picture of psychological functioning that informs some region of ethical theory compatible with the empirical picture that emerges from systematic observation? In short, is the philosophical picture *empirically adequate*? If it is determined that the philosophical conception is empirically adequate, the result is *vindictory*. Conversely, if the philosophical moral psychology in question is found to be empirically *inadequate*, the result is *revisionary*, compelling alteration, or even rejection, of those elements of the philosophical theory presupposing the problematic moral psychology. The process will often be *comparative*. Theory choice in moral psychology, like other theory choice, involves tradeoffs, and while an empirically undersupported approach may not be decisively eliminated from contention on empirical grounds alone, it may come to be seen as less attractive than theoretical options with firmer empirical foundations.

The winds driving the sort of disciplinary cross-pollination we describe do not blow in one direction. As philosophers writing for an encyclopedia of philosophy, we are naturally concerned with the ways empirical research might shape, or re-shape, philosophical ethics. But philosophical reflection may likewise influence empirical research, since such research is often driven by philosophical suppositions that may be more or less philosophically sound. The best interdisciplinary conversations, then, should benefit both parties. To illustrate the dialectical process we have described, we will consider a variety of topics in moral psychology. Our primary concerns will be philosophical: What are some of the most central problems in philosophical moral psychology, and how might they be resolved? However, as the hybrid nature of our topic invites us to do, we will pursue these questions in an interdisciplinary spirit, and are hopeful that our remarks will also engage interested scientists. Hopefully, the

result will be a broad sense of the problems and methods that will structure research on moral psychology during the 21<sup>st</sup> century.

## 2. Thought Experiments and the Methods of Ethics

“Intuition pumps” or “thought experiments” have long been well-used items in the philosopher’s toolbox (Dennett 1984: 17–18; Stuart et al. 2018). Typically, a thought experiment presents an example, often a hypothetical example, in order to elicit some philosophically telling response. If a thought experiment is successful, it may be concluded that competing theories must account for the resulting response. These responses are supposed to serve an *evidential role* in philosophical theory choice; if you like, they can be understood as *data* competing theories must accommodate.<sup>[3]</sup> If an appropriate audience’s ethical responses to a thought experiment conflict with the response a theory prescribes for the case, the theory has suffered a counterexample.

The question of whose responses “count” philosophically (or, who is the “appropriate” audience) has been answered in a variety of ways, but for many philosophers, the intended audience for thought experiments seems to be some species of “ordinary folk” (see Jackson 1998: 118, 129; Jackson & Pettit 1995: 22–9; Lewis 1989: 126–9). Of course, the relevant folk must possess such cognitive attainments as are required to understand the case at issue; very young children are probably not an ideal audience for thought experiments. Accordingly, some philosophers may insist that the relevant responses are the considered judgments of people with the training required to see “what is at stake philosophically”. But if the responses are to help adjudicate between competing theories, the responders must be more or less theoretically neutral, and this sort of neutrality is rather likely to be vitiated by philosophical education. A dilemma emerges. On the one hand, philosophically naïve subjects may be thought to lack the erudition required to grasp the philosophical stakes. On

the other, with increasing philosophical sophistication comes, very likely, philosophical partiality; one audience is naïve, and the other prejudiced.<sup>[4]</sup>

However exactly the philosophically relevant audience is specified, there are empirical questions that must be addressed in determining the philosophical potency of a thought experiment. In particular, when deciding what philosophical weight to give a response, philosophers need to determine its *origins*. What features of the *example* are implicated in a given judgment—are people reacting to the substance of the case, or the style of exposition? What features of the *audience* are implicated in their reaction—do different demographic groups respond to the example differently? Are there factors in the environment that are affecting people’s intuitive judgments? Does the order in which people consider examples affect their judgments? Such questions raise the following concern: judgments about thought experiments dealing with moral issues might be strongly influenced by *ethically irrelevant* characteristics of the example or the audience or the environment or the order of presentation. Whether a characteristic is ethically relevant is a matter for philosophical discussion, but determining the status of a particular thought experiment also requires empirical investigation of its causally relevant characteristics. We’ll now describe some examples of such investigation.

As part of their famous research on the “heuristics and biases” that underlie human reasoning, Tversky and Kahneman (1981) presented subjects with the following problem:

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

- If Program A is adopted, 200 people will be saved.
- If Program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved.

A second group of subjects was given an identical problem, except that the programs were described as follows:

- If Program C is adopted, 400 people will die.
- If Program D is adopted, there is a 1/3 probability that nobody will die and a 2/3 probability that 600 people will die.

On the first version of the problem, most subjects thought that Program A should be adopted. But on the second version, most chose Program D, despite the fact that the outcome described in A is identical to the one described in C. The disconcerting implication of this study is that ethical responses may be strongly influenced by the manner in which cases are described or *framed*. It seems that such framing sensitivities constitute ethically irrelevant influences on ethical responses. Unless this sort of possibility can be confidently eliminated, one should hesitate to rely on responses to a thought experiment for adjudicating theoretical controversies. Such possibilities can only be eliminated through systematic empirical work.<sup>[5]</sup>

While a relatively small percentage of empirical work on “heuristics and biases” directly addresses moral reasoning, numerous philosophers who have addressed the issue (Horowitz 1998; Doris & Stich 2005; Sinnott-Armstrong 2005; Sunstein 2005) agree that phenomena like framing effects are likely to be pervasively implicated in responses to ethically freighted examples, and argue that this state of affairs should cause philosophers to view the thought-experimental method with considerable concern.

We turn now to order effects. In a pioneering study, Petrinovich and O’Neill (1996) found that participants’ moral intuitions varied with the order in which the thought experiments were presented. Similar findings have been reported by Liao et al. (2012), Wiegman et al. (2012), and Schwitzgebel & Cushman (2011, 2015). The Schwitzgebel and Cushman studies are particularly striking, since they set out to explore whether order effects in moral intuitions were smaller or non-existent in professional philosophers. Surprisingly, they found that professional philosophers were also subject to order effects, even though the thought experiments used are well known in the field. Schwitzgebel and Cushman also report that in some cases philosophers intuitions show substantial order effects when the intuitions of non-philosophers don’t.

Audience characteristics may also affect the outcome of thought experiments. Haidt and associates (1993: 613) presented stories about “harmless yet offensive violations of strong social norms” to men and women of high and low socioeconomic status (SES) in Philadelphia (USA), Porto Alegre, and Recife (both in Brazil). For example:

A man goes to the supermarket once a week and buys a dead chicken. But before cooking the chicken, he has sexual intercourse with it. Then he cooks it and eats it. (Haidt et al. 1993: 617)

Lower SES subjects tended to “moralize” harmless and offensive behaviors like that in the chicken story. These subjects were more inclined than their high SES counterparts to say that the actor should be “stopped or punished”, and more inclined to deny that such behaviors would be “OK” if customary in a given country (Haidt et al. 1993: 618–19). The point is not that lower SES subjects are mistaken in their moralization of such behaviors while the urbanity of higher SES subjects represents a more rationally defensible response. The difficulty is deciding which—if any—of the conflicting responses is fit to serve as a constraint on ethical

theory, when both may equally be the result of more or less arbitrary cultural factors.

Philosophical audiences typically decline to moralize the offensive behaviors, and we ourselves share their tolerant attitude. But of course these audiences—by virtue of educational attainments, if not stock portfolios—are overwhelmingly high SES. Haidt’s work suggests that it is a mistake for a philosopher to say, as Jackson (1998: 32n4; cf. 37) does, that “my intuitions reveal the folk conception in as much as I am reasonably entitled, as I usually am, to regard myself as typical”. The question is: typical of what demographic? Are philosophers’ ethical responses determined by the philosophical substance of the examples, or by cultural idiosyncrasies that are very plausibly thought to be ethically irrelevant? Once again, until such possibilities are ruled out by systematic empirical investigation, the philosophical heft of a thought experiment is open to question.

In recent years there has been a growing body of research reporting that judgments evoked by moral thought experiments are affected by environmental factors that look to be completely irrelevant to the moral issue at hand. The presence of dirty pizza boxes and a whiff of fart spray (Schnall et al. 2008a), the use of soap (Schnall et al. 2008b) or an antiseptic handwipe (Zhong et al. 2010), or even the proximity of a hand sanitizer dispenser (Helzer & Pizarro 2011) have all been reported to influence moral intuitions. Tobia et al. (2013) found that the moral intuitions of both students and professional philosophers are affected by spraying the questionnaire with a disinfectant spray. Valdesolo and DeSteno (2006) reported that viewing a humorous video clip can have a substantial impact on participant’s moral intuitions. And Strohminger et al. (2011) have shown that hearing different kinds of audio clips (stand-up comedy or inspirational stories from a volume called *Chicken Soup for the Soul*) has divergent effects on moral intuitions.

How should moral theorists react to findings like these? One might, of course, eschew thought experiments in ethical theorizing. While this methodological austerity is not without appeal, it comes at a cost. Despite the difficulties, thought experiments are a window, in some cases the only accessible window, into important regions of ethical experience. In so far as it is disconnected from the thoughts and feels of the lived ethical life, ethical theory risks being “motivationally inaccessible”, or incapable of engaging the ethical concern of agents who are supposed to live in accordance with the normative standards of the theory.<sup>[6]</sup> Fortunately, there is another possibility: continue pursuing the research program that systematically investigates responses to intuition pumps. In effect, the idea is to subject philosophical thought experiments to the critical methods of experimental social psychology. If investigations employing different experimental scenarios and subject populations reveal a clear trend in responses, we can begin to have some confidence that we are identifying a deeply and widely shared moral conviction. Philosophical discussion may establish that convictions of this sort should serve as a constraint on moral theory, while responses to thought experiments that empirical research determines to lack such solidity, such as those susceptible to order, framing or environmental effects, or those admitting of strong cultural variation, may be ones that ethical theorists can safely disregard.

### 3. Moral Responsibility

A philosophically informed empirical research program akin to the one just described is more than a methodological fantasy. This approach accurately describes a number of research programs aimed at informing philosophical debates through interdisciplinary research.

One of the earliest examples of this kind of work was inspired in large part by the work of Knobe (2003a,b, 2006) and addressed questions surrounding “folk morality” on issues ranging from intentional action to

causal responsibility (see Knobe 2010 for review and discussion). This early work helped to spur the development of a truly interdisciplinary research program with both philosophers and psychologists investigating the folk morality of everyday life. (See the *Stanford Encyclopedia of Philosophy* article on Experimental Moral Philosophy for a more complete treatment of this research.)

Another related philosophical debate concerns the compatibility of free will and moral responsibility with determinism. On the one hand, incompatibilists insist that determinism (the view that all events are jointly determined by antecedent events as governed by laws of nature), is *incompatible* with moral responsibility. Typically, these accounts also go on to specify what particular capacity is required to be responsible for one's own behavior (e.g., that agents have alternate possibilities for behavior, or are the "ultimate" source of their behavior, or both (Kane 2002: 5; Haji 2002: 202–3).<sup>[7]</sup> On the other hand, compatibilists argue that determinism and responsibility are *compatible*, often by denying that responsible agency requires that the actor have genuinely open alternatives, or rejecting the ultimacy condition that requires indeterminism (or impossible demands for self-creation). In short, compatibilists hold that people may legitimately be held responsible even though there is some sense in which they "could not have done otherwise" or are not the "ultimate source" of their behavior. Incompatibilists deny that this is the case. Proponents of these two opposing positions have remained relatively entrenched, and some participants have raised fears of a "dialectical stalemate" (Fischer 1994: 83–5).

A critical issue in these debates has been the claim that the incompatibilist position better captures folk moral judgments about agents whose actions have been completely determined (e.g., G. Strawson 1986: 88; Smilansky 2003: 259; Pereboom 2001: xvi; O'Connor 2000: 4; Nagel 1986: 113, 125; Campbell 1951: 451; Pink 2004: 12). For example, Robert Kane (1999:

218; cf. 1996: 83–5), a leading incompatibilist, reports that in his experience "most ordinary persons start out as natural incompatibilists", and "have to be talked out of this natural incompatibilism by the clever arguments of philosophers".

Unsurprisingly, some compatibilists have been quick to assert the contrary. For example, Peter Strawson (1982) famously argued that in the context of "ordinary interpersonal relationships", people are not haunted by the specter of determinism; such metaphysical concerns are irrelevant to their experience and expression of the "reactive attitudes"—anger, resentment, gratitude, forgiveness, and the like—associated with responsibility assessment. Any anxiety about determinism, Strawson insisted, is due to the "panicky metaphysics" of philosophers, not incompatibilist convictions on the part of ordinary people. However, incompatibilists have historically been thought to have ordinary intuitions on their side; even some philosophers with compatibilist leanings are prepared to concede the incompatibilist point about "typical" response tendencies (e.g., Vargas 2005a,b).

Neither side, so far as we are aware, has offered much in the way of systematic evidence of actual patterns of folk moral judgments. Recently however, a now substantial research program has begun to offer empirical evidence on the relationship between determinism and moral responsibility in folk moral judgments.

Inspired by the work of Frankfurt (1988) and others, Woolfolk, Doris, and Darley (2006) hypothesized that observers may hold actors responsible even when the observers judge that the actors could not have done otherwise, if the actors appear to "identify" with their behavior. Roughly, the idea is that the actor identifies with a behavior—and is therefore responsible for it—to the extent that she "embraces" the behavior, or performs it "wholeheartedly" regardless of whether genuine alternatives

for behavior are possible.<sup>[8]</sup> Woolfolk et al.'s suspicion was, in effect, that people's (presumably tacit) theory of responsibility is compatibilist.

To test this, subjects were asked to read a story about an agent who was forced by a group of armed hijackers to kill a man who had been having an affair with his wife. In the "low identification" condition, the man was described as being horrified at being forced to kill his wife's lover, and as not wanting to do so. In the "high identification" condition, the man is instead described as welcoming the opportunity and wanting to kill his wife's lover. In both cases, the man is not given a choice, and does kill his wife's lover.

Consistent with Woolfolk and colleagues' hypothesis, subjects judged that the highly identifying actor was more responsible, more appropriately blamed, and more properly subject to guilt than the low identification actor.<sup>[9]</sup> This pattern in folk moral judgments seems to suggest that participants were not consistently incompatibilist in their responsibility attributions, because the lack of alternatives available to the actor was not alone sufficient to rule out such attributions.

In response to these results, those who believe that folk morality is incompatibilist may be quick to object that the study merely suggests that responsibility attributions are influenced by identification, but says nothing about incompatibilist commitments or the lack thereof. Subjects still may have believed that the actor could have done otherwise. To address this concern, Woolfolk and colleagues also conducted a version of the study in which the man acted under the influence of a "compliance drug". In this case, participants were markedly less likely to agree that the man "was free to behave other than he did" and yet they still held the agent who identified with the action as more responsible than the agent who did not. These results look to pose a clear challenge to the view that ordinary folk are typically incompatibilists.

A related pattern of responses was obtained by Nahmias, Morris, Nadelhoffer and Turner (2009) who instead described agents performing immoral behaviors in a "deterministic world" of the sort often described in philosophy classrooms. One variation read as follows:

Imagine that in the next century we discover all the laws of nature, and we build a supercomputer which can deduce from these laws of nature and from the current state of everything in the world exactly what will be happening in the world at any future time. It can look at everything about the way the world is and predict everything about how it will be with 100% accuracy. Suppose that such a supercomputer existed, and it looks at the state of the universe at a certain time on March 25th, 2150 C.E., twenty years before Jeremy Hall is born. The computer then deduces from this information and the laws of nature that Jeremy will definitely rob Fidelity Bank at 6:00 PM on January 26th, 2195. As always, the supercomputer's prediction is correct; Jeremy robs Fidelity Bank at 6:00 PM on January 26th, 2195.

Subjects were then asked whether Jeremy was morally blameworthy. Most said yes, indicating that they thought an agent could be morally blameworthy even if his behaviors were entirely determined by natural laws. Consistent with the Woolfolk et al. results, it appears that the subjects' judgments, at least those having to do with moral blameworthiness, were not governed by a commitment to incompatibilism.

This emerging picture was complicated, however, by Nichols and Knobe (2007), which argued that the ostensibly compatibilist responses were performance errors driven by an affective response to the agents' immoral actions. To demonstrate this, all subjects were asked to imagine two universes—a universe completely governed by deterministic laws (Universe A) and a universe (Universe B) in which everything is



determined except for human decisions which are not completely determined by deterministic laws and what has happened in the past. In Universe B, but not Universe A, “each human decision *does not have to happen* the way it does”. Some subjects were assigned to a concrete condition, and asked to make a judgment about a specific individual in specific circumstances, while others were assigned to an abstract condition, and asked to make a more general judgment, divorced from any particular individual. The hypothesis was that the difference between these two conditions would generate different responses regarding the relationship between determinism and moral responsibility. Subjects in the concrete condition read a story about a man, “Bill”, in the deterministic universe who murders his wife and children in a particularly ghastly manner, and were asked whether Bill was morally responsible for what he had done. By contrast, subjects in the abstract condition were asked “In Universe A, is it possible for a person to be fully morally responsible for their actions?” Seventy-two percent of subjects in the concrete condition gave a compatibilist response, holding Bill responsible in Universe A, whereas less than fifteen percent of subjects in the abstract condition gave a compatibilist response, allowing that people could be fully morally responsible in the deterministic Universe A.

In line with previous experimental work demonstrating that increased affective arousal amplified punitive responses to wrongdoing (Lerner, Goldberg, & Tetlock 1998), Nichols and Knobe hypothesized that previously observed compatibilist responses were the result of the affectively laden nature of the stimulus materials. When this affective element was eliminated from the materials (as in the abstract condition), participants instead exhibited an incompatibilist pattern of responses.

More recently, Nichols and Knobe’s line of reasoning has come under fire from two directions. First, a number of studies have now tried to systematically manipulate how affectively arousing the immoral behavior

performed is, but have not found that these changes significantly alter participants’ judgments of moral responsibility in deterministic scenarios. Rather, the differences seem to be best explained simply by whether the case was described abstractly or concretely (see Cova et al. 2012 for work with patients who have frontotemporal dementia, and see Feltz & Cova 2014 for a meta-analysis). Second, a separate line of studies from Murray and Nahmias (2014) argued that participants who exhibited the apparently incompatibilist pattern of responses were making a critical error in how they understood the deterministic scenario. In particular, they argued these participants mistakenly took the agents, or their mental states, in these deterministic scenarios to be “bypassed” in the causal chain leading up to their behavior. In support of their argument, Murray and Nahmias (2014) demonstrated that when analyses were restricted to the participants who clearly did not take the agent to be bypassed, these participants judged the agent to be morally responsible (blameworthy, etc.) despite being in a deterministic universe. Unsurprisingly, this line of argument has, in turn, inspired a number of further counter-responses, both empirical (Rose & Nichols 2013) and theoretical (Björnsson & Pereboom 2016), which caution against the conclusions of Murray and Nahmias.

While the debate continues over whether the compatibilist or incompatibilist position better captures folk moral judgments of agents in deterministic universes, a related line of research has sprung up around what is widely taken to be the most convincing contemporary form of argument for incompatibilism: manipulation arguments (e.g., Mele 2006, 2013, Pereboom 2001, 2014). Pereboom’s Four-Case version, for example, begins with the case of an agent named Plum who is manipulated by neuroscientists who use a radio-like technology to change Plum’s neural states, which results in him wanting and then deciding to kill a man named White. In this case, it seems clear that Plum did not freely decide to kill White. Compare this case to a second one, in which the team of neuroscientists programmed Plum at the beginning of his life

in a way that resulted in him developing the desire (and making the decision) to kill White. The incompatibilist argues that these two cases do not differ in a way that is relevant for whether Plum acted freely, and so, once again, it seems that Plum did not freely decide to kill White. Now compare this to a third case, in which Plum's desire and decision to kill White were instead determined by his cultural and social milieu, rather than by a team of neuroscientists. Since the only difference between the second and third case is the particular technological process through which Plum's mental states were determined, he would again seem to not have freely decided to kill White. Finally, in a fourth and final case, Plum's desire and decision to kill White was determined jointly by the past states and the laws of nature in our own deterministic universe. Regarding these four cases, Pereboom argues that, since there is no difference between any of the four cases that is relevant to free will, if Plum was not morally responsible in the first case, then he was not morally responsible in the fourth.

In response to this kind of manipulation-based argument for incompatibilism, a number of researchers have taken aim at painting a better empirical picture of ordinary moral judgments concerning manipulated agents. This line of inquiry has been productive on two levels. First, a growing number of empirical studies have investigated moral responsibility judgments about cases of manipulation, and now provide a clearer psychological picture for why manipulated agents are judged to lack free will and moral responsibility. Second, continuing theoretical work, informed by this empirical picture, has provided new reasons for doubting that manipulation based arguments actually provide evidence against compatibilism.

One line of empirical research, led by Chandra Sripada (2012) has asked whether manipulated agents are perceived to be unfree because (a) they lack ultimate control over their actions (a capacity incompatibilists take to

be essential for moral responsibility) or instead because (b) their psychological or volitional capacities (the capacities focused on by compatibilists) have been damaged. Using a statistical approach called Structural Equation Modeling (or SEM), Sripada found that participants' moral responsibility judgments were best explained by whether they believed the psychological and volitional capacities of the agent were damaged by manipulation and not whether the agent lacked control over her actions. This finding suggests that patterns of judgment in cases of manipulation are more consistent with the predictions of compatibilism than with incompatibilism.

Taking a different approach, Phillips and Shaw (2014) demonstrated that the reduction of moral responsibility that is typically observed in cases of manipulation depends critically on the role of an *intentional* manipulator. In particular, ordinary people were shown to distinguish between (1) the moral responsibility of agents who are made to do a particular act by features of the situation they are in (i.e., situational determinism), and (2) the moral responsibility of agents who are made to do that same act by another intentional agent (i.e., manipulation). This work suggests that the ordinary practice of assessing freedom and responsibility is likely to clearly distinguish between cases that do and do not involve a manipulator who intervenes with the intention of causing the manipulated agent to do the immoral action. A series of studies by Murray and Lombrozo (2016) further elaborates these findings by providing evidence that the specific reduction of moral responsibility that results from being manipulated arises from the perception that the agent's mental states are *bypassed*.

Collectively, two lessons have come out of this work on the ordinary practice of assessing the moral responsibility of manipulated agents: (1) folk morality provides a natural way of distinguishing between the different cases used in manipulation-based arguments (those that do involve the intentional intervention of a manipulator vs. those that don't)

and (2) folk morality draws an intimate link between the moral responsibility of an agent and that agent's mental and volitional capacities. Building on this increasingly clear empirical picture, Deery and Nahmias (2017) formalized these basic principles in theoretical work that argues for a principled way of distinguishing between the moral responsibility of determined and manipulated agents.

While the majority of evidence may currently be in favor of the view that folk morality adheres to a kind of "natural compatibilism" (Cova & Kitano 2013), this remains a contentious topic, and new work is continually emerging on both sides of the debate (Andow & Cova 2016; Bear & Knobe 2016; Björnsson 2014; Feltz & Millan 2013; Figdor & Phelan 2015; Knobe 2014). One thing that has now been agreed on by parties on both sides of this debate, however, is a critical role for careful empirical studies (Björnsson & Pereboom 2016; Knobe 2014; Nahmias 2011).

#### 4. Virtue Ethics and Skepticism About Character

To date, empirically informed approaches to moral psychology have been most prominent in discussions of moral character and virtue. The focus is decades of experimentation in "situationist" social psychology: unobtrusive features of situations have repeatedly been shown to impact behavior in seemingly arbitrary, and sometimes alarming, ways. Among the findings that have most interested philosophers:

- *The Phone Booth Study* (Isen & Levin (1972: 387): people who had just found a dime in a payphone's coin return were 22 times more likely than those who did not find a dime to help a woman who had dropped some papers (88% v. 4%).
- *The Good Samaritan Study* (Darley & Batson 1973: 105): unhurried passersby were 6 times more likely than hurried passersby to help an unfortunate who appeared to be in significant distress (63% v. 10%).

- *The Obedience Experiments* (Milgram 1974) subjects repeatedly punished a screaming victim with realistic (but simulated) electric shocks at the polite request of an experimenter.
- *The Stanford Prison Study* (Zimbardo 2007): college students role-playing as "guards" in a simulated prison subjected student "prisoners" to grotesque verbal and emotional abuse.

These experiments are *part of an extensive empirical literature*, where social psychologists have time and again found that disappointing omissions and appalling actions are readily induced by apparently minor situational features.<sup>[10]</sup> The striking fact is not that people fail standards for good conduct, but that they can be so *easily* induced to do so.

Exploiting this observation, "character skeptics" contend that if moral conduct varies so sharply, often for the worse, with minor perturbations in circumstance, ostensibly good character provides very limited assurance of good conduct. In addition to this claim in *descriptive psychology*, concerning the fragility of moral character, some character skeptics also forward a thesis in *normative ethics*, to the effect that character merits less attention in ethical thought than it traditionally gets.<sup>[11]</sup>

Character skepticism contravenes the influential program of contemporary *virtue ethics*, which maintains that advancing ethical theory requires *more* attention to character, and virtue ethicists offer vigorous resistance.<sup>[12]</sup> Discussion has sometimes been overheated, but it has resulted in a large literature in a vibrantly interdisciplinary field of "character studies" (e.g., Miller et al. 2015).<sup>[13]</sup> The literature is too extensive for the confines of this entry, but we will endeavor to outline some of the main issues.

The first thing to observe is that the science which inspires the character skeptics may itself be subject to skepticism. Given the uneven history of the human sciences, it might be argued that the relevant findings are too

uncertain to stand as a constraint on philosophical theorizing. This contention is potentially buttressed by recent prominent replication failures in social psychology.

The psychology at issue is, like much of science, unfinished business. But the replication controversy, and the attendant suspicion of science, is insufficient grounds for dismissing the psychology out of hand. Philosophical conclusions should not be based on a few studies; the task of the philosophical consumer of science is to identify *trends* in *convergent* strands of evidence (Doris 2015: 49, 56; Machery & Doris forthcoming). The observation that motivates character skepticism—the surprising situational sensitivity of behavior—is supported by a wide range of scientific findings, as well as by recurring themes in history and biography (Doris 2002, 2005). The strong situational discriminativeness of behavior is accepted as fact by high proportion of involved scientists; accordingly, it is not much contested in debates about character skepticism.

But the philosophical implications of this fact remain, after considerable debate, a contentious issue. The various responses to character skepticism need not be forwarded in isolation, and some of them may be combined as part of a multi-pronged defense. Different rejoinders have differing strengths and weaknesses, particularly with respect to the differing pieces of evidence on which character skeptics rely; the phenomena are not unitary, and accommodating them all may preclude a unitary response.

One way of defusing empirically motivated skepticism—dubbed by Alfano (2013) “the dodge”—is simply to deny that virtue ethics makes empirical claims. On this understanding, virtue ethics is cast as a “purely normative” endeavor aiming at erecting ethical ideals in complete absence of empirical commitments regarding actual human psychologies. This sort of purity is perhaps less honored than honored in the breach: historically, virtue ethics has been typified by an interest in how actual people *become*

good. Aristotle (*Nicomachean Ethics*, 1099b18–19) thought that anyone not “maimed” with regard to the capacity for virtue may acquire it “by a certain kind of study and care”, and contemporary Aristotelians have emphasized the importance of moral education and development (e.g., Annas 2011). More generally, virtue-based approaches have been claimed to have an advantage over major Kantian and consequentialist competitors with respect to “psychological realism”—the advantage of a more lifelike moral psychology (see Anscombe 1958: 1, 15; Williams 1985; Flanagan 1991: 182; Hursthouse 1999: 19–20).

To be sure, eschewing empirical commitment allows virtue ethics to escape empirical threat: obviously, empirical evidence cannot be used to undermine a theory that makes *no* empirical claims. However, it is not clear such theories could claim advantages traditionally claimed for virtue theories with regard to moral development and psychological realism. In any event, they are not contributions to empirical moral psychology, and needn’t be further discussed here.

Before seeing how the debate in moral psychology might be advanced, it is necessary to correct a mischaracterization that serves to arrest progress. It is too often said, particularly in reference to Doris (1998, 2002) and Harman (1999, 2000), that character skepticism comes to the view that character traits “do not exist” (e.g., Flanagan 2009: 55). Frequently, this attribution is made without documentation, but when documentation is provided, it is typically in reference to some early, characteristically pointed, remarks of Harman (e.g., 1999). Yet in his most recent contribution, Harman (2009: 241) says, “I do not think that social psychology demonstrates there are no character traits”. For his part, Doris has repeatedly asserted that traits exist, and has repeatedly drawn attention to such assertions (Doris 1998: 507–509; 2002: 62–6; 2005: 667; 2010: 138–141; Doris & Stich 2005: 119–20; Doris & Prinz 2009).

With good reason, to say “traits do not exist” is tantamount to denying that there are individual dispositional differences, an unlikely view that character skeptics and antiskeptics are united in rejecting. Quite unsurprisingly, this unlikely view is seriously undersubscribed in both philosophy and psychology. It is endorsed by neither the most aggressive critics of personality, situationists in social psychology such as Ross and Nisbett (1991), nor by the patron saint of situationism in personality psychology: Mischel (1999: 45). Mischel disavows a trait-based approach, but his skepticism concerns a *particular approach to traits*, not individual dispositional differences more generally.

Then the question of whether or not traits exist is *emphatically not the issue* dividing more and less skeptical approaches to character. Today, all mainstream parties to the debate are “interactionist”, treating behavioral outcomes as the function of a (complex) person by situation interaction (Mehl et al. 2015)—and it’s likely most participants have always been so (Doris 2002: 25–6). Contemporary research programs in personality and social psychology freely deploy *both* personal and situational variables (e.g., Cameron, Payne, & Doris 2013; Leikas, Lönnqvist, & Verkasalo 2012; Sherman, Nave, & Funder 2010). The issue worth discussing is not whether individual dispositional differences exist, but *how these differences should be characterized*, and how (or whether) these individual differences, when appropriately characterized, *should inform ethical thought*.

An important feature of early forays into character skepticism was that skeptics tended to focus on *behavioral implications* of traits rather than the *psychological antecedents* of behavior (Doris 2015: 15). Defenders of virtue ethics observe that character skeptics have had much to say about situational variation in behavior and little to say about the psychological processes underlying it, with the result that they overlook the rational order in people’s lives (Adams 2006: 115–232). These virtue ethicists

maintain that the behavioral variation provoking character skepticism evinces not unreliability, but rationally appropriate sensitivity to differing situations (Adams 2006; Kamtekar 2004). The virtuous person, such as Aristotle’s exemplary *phronimos* (“man of practical wisdom”) may sometimes come clean, and sometimes dissemble, or sometimes fight, and sometimes flee, depending on the particular ethical demands of his circumstances.

For example, in the Good Samaritan Study, the hurried passersby was on the way to an appointment where they had agreed to give a presentation; perhaps these people made a rational determination—perhaps even an ethically defensible determination—to weigh the demands of punctuality and professionalism over ethical requirement to check on the welfare of a stranger in apparent distress. However attractive one finds such accounting for this case (note that some of Darley and Batson’s [1973] hurried passersby failed to notice the victim, which strains explanations in terms of their rational discriminations), there are other cases where the “rationality response” seems plainly unattractive. These are cases of ethically irrelevant influences (Sec. 2 above; Doris & Stich 2005), where it seems unlikely the influence could be cited as part of a rationalizing explanation of the behavior: it’s odd to cite failing to find a dime as *justification* for failing to help—or for that matter, finding a dime as *justification* for doing so.

It is certainly appropriate for virtue ethicists to emphasize practical rationality in their accounts of character. This is a central theme in the tradition going back to Aristotle himself, who is probably the most oft-cited canonical philosopher in contemporary virtue ethics. But while the rationality response may initially accommodate some of the troubling behavioral evidence, it encounters further empirical difficulty. There is an extensive empirical literature problematizing familiar conceptions of rationality: psychologists have endlessly documented a dispiriting range of

reasoning errors (Baron 1994, 2001; Gilovich et al. 2002; Kahneman et al. 1982; Tversky & Kahneman 1973; Kruger & Dunning 1999; Nisbett & Borgida 1975; Nisbett & Ross 1980; Stich 1990; Tversky & Kahneman 1981). In light of this evidence, character skeptics claim that the vagaries afflicting behavior also afflict reasoning (Alfano 2013; Olin & Doris 2014).

Research supporting this discouraging assessment of human rationality is controversial, and not all psychologists think things are so bleak (Gigerenzer 2000; Gigerenzer et al. 1999; for philosophical commentary see Samuels & Stich 2002). Nevertheless, if virtue ethics is to have an empirically credible moral psychology, it needs to account for the empirical challenges to practical reasoning: how can the relevant excellence in practical reasoning be developed?

Faced with the challenge to practical rationality, virtue ethicists may respond that their theories concern *excellent* reasoning, not the *ordinary* reasoning studied in psychology. Practical wisdom, and the ethical virtue it supports, are expected to be *rare*, and not widely instantiated. This state of affairs, it is said, is quite compatible with the disturbing, but not exceptionlessly disturbing, behavior in experiments like Milgram's (see Athanassoulis 1999: 217–219; DePaul 1999; Kupperman 2001: 242–3). If this account is supposed to be part of an empirically contentful moral psychology, rather than unverified speculation, we require a detailed and empirically substantiated account of how the virtuous few get that way—remember that an emphasis on moral development is central to the virtue ethics tradition. Moreover, if virtue ethics is supposed to have widespread practical implications—as opposed to being merely a celebration of a tiny “virtue elite”—it should have an account of how the less-than-virtuous-many may at least tolerably *approximate* virtue.

This point is underscored by the fact that for some of the troubling evidence, as in the Stanford Prison Study, the worry is not so much that people fail standards of virtue, but that they fail standards of *minimal decency*. Surely an approach to ethics that celebrates moral development, even one that acknowledges (or rather, insists) that most people will not attain its ideal, might be expected to have an account of how people can become minimally decent.

Recently, proponents of virtue ethics have been increasingly proposing a suggestive solution to this problem: virtue is a skill acquired through effortful practice, so virtue is a kind of expertise (Annas 2011; Bloomfield 2000, 2001, 2014; Jacobson 2005; Russell 2015; Snow 2010; Sosa 2009; Stichter 2007, 2011; for reservations, see Doris, in preparation). The virtuous are expert at morality and—given the Aristotelian association of virtue and happiness—expert at life.

An extensive scientific literature indicates that developing expert skill requires extensive preparation, whether the practitioner is a novelist, doctor, or chess master—around 10,000 hours of “deliberate practice”, according to a popular generalization (Ericsson 2014; Ericsson et al. 1993). The “10,000-hour rule” is likely an oversimplification, but there is no doubt that attaining expertise requires intensive training. Because of this, people rarely achieve eminence in more than one area; for instance, “baseball trivia” experts display superior recall for baseball-related material, but not for non-baseball material (Chiesi et al. 1979). Conversely, becoming expert at morality, or (even more ambitiously) expert at the whole of life, would apparently require a *highly* generalized form of expertise: to be good, there's a *lot* to be good at. Moreover, it's quite unclear what deliberate practice at life involves; how exactly does one get better at being good?

One obvious problem concerns specifying the “good” in question. Expertises like chess have been effectively studied in part because there are accepted standards of excellence (the “ELO” score used for ranking chess players; Glickman 1995). To put it blithely, there aren’t any chess skeptics. But there have, historically, been *lots* of moral skeptics. And if there’s not moral knowledge, how could there be moral experts? And even if there are moral experts, there’s the problem of how are they to be identified, since it is not clear we are possessed of standard independent of expert opinion itself (like winning chess matches) for doing so (for the “metaethics of expertise”, see McGrath 2008, 2011).

Even if these notorious philosophical difficulties can be resolved—as defenders of expertise approaches to virtue must think they can—matters remain complicated, because if moral expertise is like other expertises, practice alone—assuming we have a clear notion of what “moral practice” entails—will be insufficient. While practice matters in attaining expertise, other factors, such as talent, also matter (Hambrick et al. 2014; Macnamara et al. 2014). And some of the required endowments may be quite unequally distributed across populations: practice cannot make a jockey into an NFL lineman, or an NFL lineman into a jockey.

What are the natural endowments required for moral expertise, and how widely are they distributed in the population? If they are rare, like the skill of a chess master or the strength of an NFL lineman, virtue will also be rare. Some virtue ethicists believe virtue should be widely attainable, and they will resist this result (Adams 2006: 119–123, and arguably Aristotle *Nicomachean Ethics* 1099b15–20). But even virtue ethicists who embrace the rarity of virtue require an account of what the necessary natural endowments are, and if they wish to also have an account of how the less well-endowed may achieve at least minimal decency, they should have something to say about how moral development will proceed across a population with widely varying endowments.

What is needed, for the study of moral character research to advance, is an account of the biological, psychological, and social factors requisite for successful moral development—on the expertise model, the conditions conducive to developing “moral skill”. This, quite obviously, is a tall order, and the research needed to systematically address these issue is in comparative infancy. Yet the expertise model, in exploiting connections with areas in which skill acquisition has been well studied, such as music and sport, provides a framework for moving discussion of character beyond the empirically under-informed conjectures and assumptions about “habituation” that have been too frequent in previous literature (Doris 2015: 128).

## 5. Egoism vs. Altruism

People often behave in ways that benefit others, and they sometimes do this knowing that it will be costly, unpleasant or dangerous. But at least since Plato’s classic discussion in the second Book of the *Republic*, debate has raged over *why* people behave in this way. Are their motives altruistic, or is their behavior ultimately motivated by self-interest? Famously, Hobbes gave this answer:

No man giveth but with intention of good to himself, because gift is voluntary; and of all voluntary acts, the object is to every man his own good; of which, if men see they shall be frustrated, there will be no beginning of benevolence or trust, nor consequently of mutual help. (1651 [1981: Ch. 15])

Views like Hobbes’ have come to be called *egoism*,<sup>[14]</sup> and this rather depressing conception of human motivation has any number of eminent philosophical advocates, including Bentham, J.S. Mill and Nietzsche.<sup>[15]</sup> Dissenting voices, though perhaps fewer in number, have been no less

eminent. Butler, Hume, Rousseau and Adam Smith have all argued that, sometimes at least, human motivation is genuinely altruistic.

Though the issue that divides egoistic and altruistic accounts of human motivation is largely empirical, it is easy to see why philosophers have thought that the competing answers will have important consequences for moral theory. For example, Kant famously argued that a person should act “not from inclination but from duty, and by this would his conduct first acquire true moral worth” (1785 [1949: Sec. 1, parag. 12]). But egoism maintains that *all* human motivation is ultimately self-interested, and thus people *can’t* act “from duty” in the way that Kant urged. Thus if egoism is true, Kant’s account would entail that no conduct has “true moral worth”. Additionally, if egoism is true, it would appear to impose a strong constraint on how a moral theory can answer the venerable question “Why should I be moral?” since, as Hobbes clearly saw, the answer will have to ground the motivation to be moral in the agent’s self-interest.<sup>[16]</sup>

While the egoism vs. altruism debate has historically been of great philosophical interest, the issue centrally concerns psychological questions about the nature of human motivation, so it’s not surprise that psychologists have done a great deal of empirical research aimed at determining which view is correct. Some of the most influential and philosophically sophisticated empirical work on this issue has been done by Daniel Batson and his associates. The conclusion Batson draws from this work is that people *do* sometimes behave altruistically, and that the emotion of empathy plays an important role in generating altruistic motivation.<sup>[17]</sup> Others are not convinced. For a discussion of Batson’s experiments, the conclusion he draws from them, and some reasons for skepticism about that conclusion, see sections 5 and 6 of the entry “Empirical Approaches to Altruism” in this encyclopedia. In this section, we’ll focus on some of the philosophical spadework that is necessary before plunging into the empirical literature.

A crucial question that needs to be addressed is: What, exactly, is the debate about; what *is* altruism? Unfortunately, there is no uncontroversial answer to this question, since researchers in many disciplines, including philosophy, biology, psychology, sociology, economics, anthropology and primatology, have written about altruism, and authors in different disciplines tend to use the term “altruism” in quite different ways. Even among philosophers the term has been used with importantly different meanings. There is, however, one account of altruism—actually a cluster closely related accounts—that plays a central role both in philosophy and in a great deal of psychology, including Batson’s work. We’ll call it “the standard account”. That will be our focus in the remainder of this section.<sup>[18]</sup>

According to the standard account, an action is altruistic if it is motivated by an ultimate desire for the well-being of another person. This formulation invites questions about (1) what it is for a behavior to be *motivated by an ultimate desire*, and (2) the distinction between *desires that are self-interested* and *desires that are for the well-being of others*.

Although the second question will need careful consideration in any comprehensive treatment, a few rough and ready examples of the distinction will suffice here.<sup>[19]</sup> Desires to save someone else’s life, to alleviate someone else’s suffering, or to make someone else happy are paradigm cases of desires for the well-being of others, while desires to experience pleasure, get rich, and become famous are typical examples of self-interested desires. The self-interested desires to experience pleasure and to avoid pain have played an especially prominent role in the debate, since one version of egoism, often called *hedonism*, maintains that these are our *only* ultimate desires.

The first question, regarding ultimate desires, requires a fuller exposition; it can be usefully explicated with the help of a familiar account of



*practical reasoning*.<sup>[20]</sup> On this account, practical reasoning is a causal process via which a desire and a belief give rise to or sustain another desire. For example, a desire to drink an espresso and a belief that the best place to get an espresso is at the espresso bar on Main Street may cause a desire to go to the espresso bar on Main Street. This desire can then join forces with another belief to generate a third desire, and so on. Sometimes this process will lead to a desire to perform a relatively simple or “basic” action, and that desire, in turn, will cause the agent to perform the basic action without the intervention of any further desires. Desires produced or sustained by this process of practical reasoning are *instrumental* desires—the agent has them because she thinks that satisfying them will lead to something else that she desires. But not *all* desires can be instrumental desires. If we are to avoid circularity or an infinite regress there must be some desires that are *not* produced because the agent thinks that satisfying them will facilitate satisfying some other desire. These desires that are not produced or sustained by practical reasoning are the agent’s *ultimate* desires, and the objects of ultimate desires, the states of affairs desired, are desired for their own sake. A behavior is *motivated* by a specific ultimate desire when that desire is part of the practical reasoning process that leads to the behavior.

If people do sometimes have ultimate desires for the well-being of others, and these desires motivate behavior, then altruism is the correct view, and egoism is false. However, if *all* ultimate desires are self-interested, then egoism is the correct view, and altruism is false. The effort to establish one or the other of these options has given rise to a vast and enormously sophisticated empirical literature. For an overview of that literature, see the empirical approaches to altruism entry.

## 6. Moral Disagreement

Given that moral disagreement—about abortion, say, or capital punishment—so often seems intractable, is there any reason to think that moral problems admit objective resolutions? While this difficulty is of ancient coinage, contemporary philosophical discussion was spurred by Mackie’s (1977: 36–8) “argument from relativity” or, as it is called by later writers, the “argument from disagreement” (Brink 1989: 197; Loeb 1998). Such “radical” differences in moral judgment as are frequently observed, Mackie (1977: 36) argued, “make it difficult to treat those judgments as apprehensions of objective truths”.

Mackie supposed that his argument undermines *moral realism*, the view that, as Smith (1994: 9, cf. 13) puts it,

moral questions have correct answers, that the correct answers are made correct by objective moral facts ... and ... by engaging in moral argument, we can discover what these objective moral facts are.<sup>[21]</sup>

This notion of objectivity, as Smith recognizes, requires *convergence* in moral views—the right sort of argument, reflection and discussion is expected to result in very substantial moral agreement (Smith 1994: 6).<sup>[22]</sup>

While moral realists have often taken pretty optimistic positions on the extent of actual moral agreement (e.g., Sturgeon 1988: 229; Smith 1994: 188), there is no denying that there is an abundance of persistent moral disagreement; on many moral issues there is a striking *failure of convergence* even after protracted argument. Anti-realists like Mackie have a ready explanation for this phenomenon: Moral judgment is not objective in Smith’s sense, and moral argument cannot be expected to accomplish what Smith and other realists think it can.<sup>[23]</sup> Conversely, the

realist's task is to *explain away* failures of convergence; she must provide an explanation of the phenomena consistent with it being the case that moral judgment is objective and moral argument is rationally resolvable. Doris and Plakias (2008) call these “defusing explanations”. The realist's strategy is to insist that the preponderance of actual moral disagreement is due to limitations of disputants or their circumstances, and insist that (very substantial, if not unanimous)<sup>[24]</sup> moral agreement *would* emerge in *ideal* conditions, when, for example, disputants are fully rational and fully informed of the relevant non-moral facts.

It is immediately evident that the relative merits of these competing explanations cannot be fairly determined without close discussion of the factors implicated in actual moral disagreements. Indeed, as acute commentators with both realist (Sturgeon 1988: 230) and anti-realist (Loeb 1998: 284) sympathies have noted, the argument from disagreement cannot be evaluated by *a priori* philosophical means alone; what's needed, as Loeb observes, is “a great deal of further empirical research into the circumstances and beliefs of various cultures”. This research is required not only to accurately assess the extent of actual disagreement, but also to determine *why* disagreement persists or dissolves. Only then can realists' attempts to “explain away” moral disagreement be fairly assessed.

Richard Brandt, who was a pioneer in the effort to integrate ethical theory and the social sciences, looked primarily to anthropology to help determine whether moral attitudes can be expected to converge under idealized circumstances. It is of course well known that anthropology includes a substantial body of work, such as the classic studies of Westermarck (1906) and Sumner (1908 [1934]), detailing the radically divergent moral outlooks found in cultures around the world. But as Brandt (1959: 283–4) recognized, typical ethnographies do not support confident inferences about the convergence of attitudes under ideal conditions, in large measure because they often give limited guidance

regarding how much of the moral disagreement can be traced to disagreement about factual matters that are not moral in nature, such as those having to do with religious or cosmological views.

With this sort of difficulty in mind, Brandt (1954) undertook his own anthropological study of Hopi people in the American southwest, and found issues for which there appeared to be serious moral disagreement between typical Hopi and white American attitudes that could not plausibly be attributed to differences in belief about nonmoral facts.<sup>[25]</sup> A notable example is the Hopi attitude toward animal suffering, an attitude that might be expected to disturb many non-Hopis:

[Hopi children] sometimes catch birds and make “pets” of them. They may be tied to a string, to be taken out and “played” with. This play is rough, and birds seldom survive long. [According to one informant:] “Sometimes they get tired and die. Nobody objects to this”. (Brandt 1954: 213)

Brandt (1959: 103) made a concerted effort to determine whether this difference in moral outlook could be traced to disagreement about nonmoral facts, but he could find no plausible explanation of this kind; his Hopi informants didn't believe that animals lack the capacity to feel pain, for example, nor did they have cosmological beliefs that would explain away the apparent cruelty of the practice, such as beliefs to the effect that animals are rewarded for martyrdom in the afterlife. The best explanation of the divergent moral judgments, Brandt (1954: 245, 284) concluded, is a “basic difference of attitude”, since “groups do sometimes make divergent appraisals when they have identical beliefs about the objects”.

Moody-Adams argues that little of philosophical import can be concluded from Brandt's—and indeed from much—ethnographic work. Deploying Gestalt psychology's doctrine of “situational meaning” (e.g., Dunker

1939), Moody-Adams (1997: 34–43) contends that all institutions, utterances, and behaviors have meanings that are peculiar to their cultural milieu, so that we cannot be certain that participants in cross-cultural disagreements are talking about the same thing.<sup>[26]</sup> The problem of situational meaning, she thinks, threatens “insuperable” methodological difficulty for those asserting the existence of intractable intercultural disagreement (1997: 36). Advocates of ethnographic projects will likely respond—not unreasonably, we think—that judicious observation and interview, such as that to which Brandt aspired, *can* motivate confident assessments of evaluative diversity. Suppose, however, that Moody-Adams is right, and the methodological difficulties are insurmountable. Now, there’s an equitable distribution of the difficulty: if observation and interview are really as problematic as Moody-Adams suggests, *neither* the realists’ *nor* the anti-realists’ take on disagreement can be supported by appeal to empirical evidence. We do not think that such a stalemate obtains, because we think the implicated methodological pessimism excessive. Serious empirical work can, we think, tell us a lot about cultures and the differences between them. The appropriate way of proceeding is with close attention to particular studies, and what they show and fail to show.<sup>[27]</sup>

As Brandt (1959: 101–2) acknowledged, the anthropological literature of his day did not always provide as much information on the exact contours and origins of moral attitudes and beliefs as philosophers wondering about the prospects for convergence might like. However, social psychology and cognitive science have recently produced research which promises to further discussion; during the last 35 years, there has been an explosion of “cultural psychology” investigating the cognitive and emotional processes of different cultures (Shweder & Bourne 1982; Markus & Kitayama 1991; Ellsworth 1994; Nisbett & Cohen 1996; Nisbett 1998, 2003; Kitayama & Markus 1999; Heine 2008; Kitayama & Cohen 2010; Henrich 2015). Here we will focus on some cultural differences found close to (our) home,

differences discovered by Nisbett and his colleagues while investigating regional patterns of violence in the American North and South. We argue that these findings support Brandt’s pessimistic conclusions regarding the likelihood of convergence in moral judgment.

The Nisbett group’s research can be seen as applying the tools of cognitive social psychology to the “culture of honor”, a phenomenon that anthropologists have documented in a variety of groups around the world. Although these groups differ in many respects, they manifest important commonalities:

A key aspect of the culture of honor is the importance placed on the insult and the necessity to respond to it. An insult implies that the target is weak enough to be bullied. Since a reputation for strength is of the essence in the culture of honor, the individual who insults someone must be forced to retract; if the instigator refuses, he must be punished—with violence or even death. (Nisbett & Cohen 1996: 5)

According to Nisbett and Cohen (1996: 5–9), an important factor in the genesis of southern honor culture was the presence of a herding economy. Honor cultures are particularly likely to develop where resources are liable to theft, and where the state’s coercive apparatus cannot be relied upon to prevent or punish thievery. These conditions often occur in relatively remote areas where herding is a main form of subsistence; the “portability” of herd animals makes them prone to theft. In areas where farming rather than herding dominates, cooperation among neighbors is more important, stronger government infrastructures are more common, and resources—like decidedly unportable farmland—are harder to steal. In such agrarian social economies, cultures of honor tend not to develop. The American South was originally settled primarily by peoples from remote areas of Britain. Since their homelands were generally unsuitable for

farming, these peoples have historically been herders; when they emigrated from Britain to the American South, they initially sought out remote regions suitable for herding, and in such regions, the culture of honor flourished.

In the contemporary South, police and other government services are widely available and herding has all but disappeared as a way of life, but certain sorts of violence continue to be more common than they are in the North. Nisbett and Cohen (1996) maintain that patterns of violence in the South, as well as attitudes toward violence, insults, and affronts to honor, are best explained by the hypothesis that a culture of honor persists among contemporary white non-Hispanic southerners. In support of this hypothesis, they offer a compelling array of evidence, including:

- demographic data indicating that (1) among southern whites, homicides rates are higher in regions more suited to herding than agriculture, and (2) white males in the South are much more likely than white males in other regions to be involved in homicides resulting from arguments, although they are *not* more likely to be involved in homicides that occur in the course of a robbery or other felony (Nisbett & Cohen 1996: Ch. 2)
- survey data indicating that white southerners are more likely than northerners to believe that violence would be “extremely justified” in response to a variety of affronts, and that if a man failed to respond violently to such affronts, he was “not much of a man” (Nisbett & Cohen 1996: Ch. 3)
- legal scholarship indicating that southern states “give citizens more freedom to use violence in defending themselves, their homes, and their property” than do northern states (Nisbett & Cohen 1996: Ch. 5, p. 63)

Two experimental studies—one in the field, the other in the laboratory—are especially striking.

In the field study (Nisbett & Cohen 1996: 73–5), letters of inquiry were sent to hundreds of employers around the United States. The letters purported to be from a hardworking 27-year-old Michigan man who had a single blemish on his otherwise solid record. In one version, the “applicant” revealed that he had been convicted for manslaughter. The applicant explained that he had been in a fight with a man who confronted him in a bar and told onlookers that “he and my fiancée were sleeping together. He laughed at me to my face and asked me to step outside if I was man enough”. According to the letter, the applicant’s nemesis was killed in the ensuing fray. In the other version of the letter, the applicant revealed that he had been convicted of motor vehicle theft, perpetrated at a time when he needed money for his family. Nisbett and his colleagues assessed 112 letters of response, and found that southern employers were significantly more likely to be cooperative and sympathetic in response to the manslaughter letter than were northern employers, while no regional differences were found in responses to the theft letter. One southern employer responded to the manslaughter letter as follows:

As for your problems of the past, anyone could probably be in the situation you were in. It was just an unfortunate incident that shouldn’t be held against you. Your honesty shows that you are sincere.... I wish you the best of luck for your future. You have a positive attitude and a willingness to work. These are qualities that businesses look for in employees. Once you are settled, if you are near here, please stop in and see us. (Nisbett & Cohen 1996: 75)

No letters from northern employers were comparably sympathetic.

In the laboratory study (Nisbett & Cohen 1996: 45–8) subjects—white males from both northern and southern states attending the University of Michigan—were told that saliva samples would be collected to measure blood sugar as they performed various tasks. After an initial sample was collected, the unsuspecting subject walked down a narrow corridor where an experimental confederate was pretending to work on some filing. The confederate bumped the subject and, feigning annoyance, called him an “asshole”. A few minutes after the incident, saliva samples were collected and analyzed to determine the level of cortisol—a hormone associated with high levels of stress, anxiety and arousal, and testosterone—a hormone associated with aggression and dominance behavior. As Figure 1 indicates, southern subjects showed dramatic increases in cortisol and testosterone levels, while northerners exhibited much smaller changes.

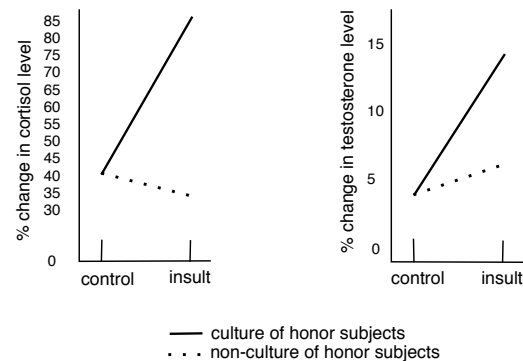


FIGURE 1

The two studies just described suggest that southerners respond more strongly to insult than northerners, and take a more sympathetic view of others who do so, manifesting just the sort of attitudes that are supposed to typify honor cultures. We think that the data assembled by Nisbett and his colleagues make a persuasive case that a culture of honor persists in the American South. Apparently, this culture affects people’s judgments,

attitudes, emotion, behavior, and even their physiological responses. Additionally, there is evidence that child rearing practices play a significant role in passing the culture of honor on from one generation to the next, and also that relatively permissive laws regarding gun ownership, self-defense, and corporal punishment in the schools both reflect and reinforce southern honor culture (Nisbett & Cohen 1996: 60–63, 67–9). In short, it seems to us that the culture of honor is deeply entrenched in contemporary southern culture, despite the fact that many of the material and economic conditions giving rise to it no longer widely obtain.<sup>[28]</sup>

We believe that the North/South cultural differences adduced by Nisbett and colleagues support Brandt’s conclusion that moral attitudes will often fail to converge, even under ideal conditions. The data should be especially troubling for the realist, for despite the differences that we have been recounting, contemporary northern and southern Americans might be expected to have rather more in common—from circumstance to language to belief to ideology—than do, say, Yanomamö and Parisians. So if there is little ground for expecting convergence in the case at hand, there is probably little ground in a good many others.

Fraser and Hauser (2010) are not convinced by our interpretation of Nisbett and Cohen’s data. They maintain that while those data do indicate that northerners and southerners differ in the strength of their disapproval of insult-provoked violence, they do not show that northerners and southerners have a real moral disagreement. They go on to argue that the work of Abarbanell and Hauser (2010) provides a much more persuasive example of a systematic moral disagreement between people in different cultural groups. Abarbanell and Hauser focused on the moral judgments of rural Mayan people in the Mexican state of Chiapas. They found that people in that community do not judge *actions* causing harms to be worse than *omissions* (failures to act) which cause identical harms, while nearby

urban Mayan people and Western internet users judge actions to be substantially worse than omissions.

Though we are not convinced by Fraser and Hauser's interpretation of the Nisbett and Cohen data, we agree that the Abarbanell and Hauser study provides a compelling example of a systematic cultural difference in moral judgement. Barrett et al. (2016) provides another example. That study looked at the extent to which an agent's intention affected the moral judgments of people in eight traditional small-scale societies and two Western societies, one urban, one rural. They found that in some of these societies, notably including both Western groups, the agent's intention had a major effect, while in other societies agent intention had little or no effect.

As we said at the outset, realists defending conjectures about convergence may attempt to *explain away* evaluative diversity by arguing that the diversity is to be attributed to shortcomings of discussants or their circumstances. If this strategy can be made good, moral realism may survive an empirically informed argument from disagreement: so much the worse for the instance of moral reflection and discussion in question, not so much the worse for the objectivity of morality. While we cannot here canvass all the varieties of this suggestion, we will briefly remark on some of the more common forms. For concreteness, we will focus on Nisbett and Cohen's study.

*Impartiality.* One strategy favored by moral realists concerned to explain away moral disagreement is to say that such disagreement stems from the distorting effects of individual interest (see Sturgeon 1988: 229–230; Enoch 2009: 24–29); perhaps persistent disagreement doesn't so much betray deep features of moral argument and judgment as it does the doggedness with which individuals pursue their perceived advantage. For instance, seemingly moral disputes over the distribution of wealth may be

due to perceptions—perhaps mostly inchoate—of individual and class interests rather than to principled disagreement about justice; persisting moral disagreement in such circumstances fails the impartiality condition, and is therefore untroubling to the moral realist. But it is rather implausible to suggest that North/South disagreements as to when violence is justified will fail the impartiality condition. There is no reason to think that southerners would be unwilling to universalize their judgments across relevantly similar individuals in relevantly similar circumstances, as indeed Nisbett and Cohen's "letter study" suggests. One can advocate a violent honor code without going in for special pleading. [29] We do not intend to denigrate southern values; our point is that while there may be good reasons for criticizing the honor-bound southerner, it is not obvious that the reason can be failure of impartiality, if impartiality is (roughly) to be understood along the lines of a willingness to universalize one's moral judgments.

*Full and vivid awareness of relevant nonmoral facts.* Moral realists have argued that moral disagreements very often derive from disagreement about nonmoral issues. According to Boyd (1988: 213; cf. Brink 1989: 202–3; Sturgeon 1988: 229),

careful philosophical examination will reveal ... that agreement on nonmoral issues would eliminate *almost all* disagreement about the sorts of moral issues which arise in ordinary moral practice.

Is this a plausible conjecture for the data we have just considered? We find it hard to imagine what agreement on nonmoral facts could do the trick, for we can readily imagine that northerners and southerners might be in full agreement on the relevant nonmoral facts in the cases described. Members of both groups would presumably agree that the job applicant was cuckolded, for example, or that calling someone an "asshole" is an insult. We think it much more plausible to suppose that the disagreement

resides in differing and deeply entrenched evaluative attitudes regarding appropriate responses to cuckolding, challenge, and insult.

Savvy philosophical readers will be quick to observe that terms like “challenge” and “insult” look like “thick” ethical terms, where the evaluative and descriptive are commingled (see Williams 1985: 128–30); therefore, it is very difficult to say what the extent of the factual disagreement is. But this is of little help for the expedient under consideration, since the disagreement-in-nonmoral-fact response apparently *requires* that one *can* disentangle factual and moral disagreement.

It is of course possible that full and vivid awareness of the nonmoral facts might motivate the sort of change in southern attitudes envisaged by the (at least the northern) moral realist. Were southerners to become vividly aware that their culture of honor was implicated in violence, they might be moved to change their moral outlook. (We take this way of putting the example to be the most natural one, but nothing philosophical turns on it. If you like, substitute the possibility of northerners endorsing honor values after exposure to the facts.) On the other hand, southerners might insist that the values of honor should be nurtured even at the cost of promoting violence; the motto “death before dishonor”, after all, has a long and honorable history. The burden of argument, we think, lies with the realist who asserts—*culture and history notwithstanding*—that southerners would change their mind if vividly aware of the pertinent facts.

*Freedom from “Abnormality”.* Realists may contend that much moral disagreement may result from failures of rationality on the part of discussants (Brink 1989: 199–200). Obviously, disagreement stemming from cognitive impairments is no embarrassment for moral realism; at the limit, that a disagreement persists when some or all disputing parties are quite insane shows nothing deep about morality. But it doesn’t seem

plausible that southerners’ more lenient attitudes towards certain forms of violence are readily attributed to widespread cognitive disability. Of course, this is an empirical issue, but we don’t know of any evidence suggesting that southerners suffer some cognitive impairment that prevents them from understanding demographic and attitudinal factors in the genesis of violence, or any other matter of fact. What is needed to press home a charge of irrationality is evidence of cognitive impairment *independent* of the attitudinal differences, and further evidence that this impairment is implicated in adherence to the disputed values. In this instance, as in many others, we have difficulty seeing how charges of abnormality or irrationality can be made without one side begging the question against the other.

Nisbett and colleagues’ work may represent a potent counterexample to any theory maintaining that rational argument tends to convergence on important moral issues; the evidence suggests that the North/South differences in attitudes towards violence and honor might well persist even under the sort of ideal conditions under consideration. Admittedly, such conclusions must be tentative. On the philosophical side, not every plausible strategy for “explaining away” moral disagreement and grounding expectations of convergence has been considered.<sup>[30]</sup> On the empirical side, this entry has reported on but a few studies, and those considered, like any empirical work, might be criticized on either conceptual or methodological grounds.<sup>[31]</sup> Finally, it should be clear what this entry is *not* claiming: any conclusions here—even if fairly earned—are not a “refutation” of all versions of moral realism, since there are versions of moral realism that do not require convergence (Bloomfield 2001; Shafer-Landau 2003). Rather, this discussion should give an idea of the empirical work philosophers must encounter, if they are to make defensible conjectures regarding moral disagreement.

## 7. Conclusion

Progress in ethical theorizing often requires progress on difficult psychological questions about how human beings can be expected to function in moral contexts. It is no surprise, then, that moral psychology is a central area of inquiry in philosophical ethics. It should also come as no surprise that empirical research, such as that conducted in psychology departments, may substantially abet such inquiry. Nor then, should it surprise that research in moral psychology has become *methodologically pluralistic*, exploiting the resources of, and endeavoring to contribute to, various disciplines. Here, we have illustrated how such interdisciplinary inquiry may proceed with regard to central problems in philosophical ethics.

## Bibliography

- Abarbanell, Linda and Marc D. Hauser, 2010, “Mayan Morality: An Exploration of Permissible Harms”, *Cognition*, 115(2): 207–224. doi:10.1016/j.cognition.2009.12.007
- Adams, Robert Merrihew, 2006, *A Theory of Virtue: Excellence in Being for the Good*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199207510.001.0001
- Alfano, Mark, 2013, *Character as Moral Fiction*, Cambridge: Cambridge University Press. doi:10.1017/CBO9781139208536
- , 2016, *Moral Psychology: An Introduction*, Cambridge: Polity Press.
- Andow, James and Florian Cova, 2016, “Why Compatibilist Intuitions Are Not Mistaken: A Reply to Feltz and Millan”, *Philosophical Psychology*, 29(4): 550–566. doi:10.1080/09515089.2015.1082542
- Annas, Julia, 2005, “Comments on John Doris’ *Lack of Character*”, *Philosophy and Phenomenological Research*, 71(3): 636–642. doi:10.1111/j.1933-1592.2005.tb00476.x
- , 2011, *Intelligent Virtue*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199228782.001.0001
- Anscombe, G.E.M., 1958, “Modern Moral Philosophy”, *Philosophy*, 33(124): 1–19. doi:10.1017/S0031819100037943
- Appiah, Kwame Anthony, 2008, *Experiments in Ethics*, Cambridge, MA: Harvard University Press.
- Aristotle, *Nicomachean Ethics*, in *The Complete Works of Aristotle*, edited by J. Barnes, Princeton: Princeton University Press, 1984.
- Arpaly, Nomy, 2005, “Comments on *Lack of Character* by John Doris”, *Philosophy and Phenomenological Research*, 71(3): 643–647. doi:10.1111/j.1933-1592.2005.tb00477.x
- Athanassoulis, Nafsika, 1999, “A Response to Harman: Virtue Ethics and Character Traits”, *Proceedings of the Aristotelian Society*, 100(1): 215–222. doi:10.1111/j.0066-7372.2003.00012.x
- Badhwar, Neera K., 2009, “The Milgram Experiments, Learned Helplessness, and Character Traits”, *The Journal of Ethics*, 13(2–3): 257–289. doi:10.1007/s10892-009-9052-4
- Baron, Jonathan, 1994, “Nonconsequentialist Decisions”, *Behavioral and Brain Sciences*, 17(1): 1–42. doi:10.1017/S0140525X0003301X
- , 2001, *Thinking and Deciding*, 3<sup>rd</sup> edition, Cambridge: Cambridge University Press.
- Barrett, H.C., A. Bolyanatz, A. Crittenden, D.M.T. Fessler, S. Fitzpatrick, M. Gurven, J. Henrich, M. Kanovsky, G. Kushnick, A. Pisor, B. Scelza, S. Stich, C. von Rueden, W. Zhao and S. Laurence, 2016, “Small-Scale Societies Exhibit Fundamental Variation in the Role of Intentions in Moral Judgment”, *Proceedings of the National Academy of Sciences*, 113(17): 4688–4693. doi:10.1073/pnas.1522070113
- Batson, C. Daniel, 1991, *The Altruism Question: Toward a Social-Psychological Answer*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- , 2011, *Altruism in Humans*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780195341065.001.0001



- Bear, Adam and Joshua Knobe, 2016, “What Do People Find Incompatible With Causal Determinism?” *Cognitive Science*, 40(8): 2025–2049. doi:10.1111/cogs.12314
- Besser-Jones, Lorraine, 2008, “Social Psychology, Moral Character, and Moral Fallibility”, *Philosophy and Phenomenological Research*, 76(2): 310–332. doi:10.1111/j.1933-1592.2007.00134.x
- Björnsson, Gunnar, 2014, “Incompatibilism and ‘Bypassed’ Agency”, in Alfred R. Mele (ed.), *Surrounding Free Will*, Oxford: Oxford University Press, pp. 95–112. doi:10.1093/acprof:oso/9780199333950.003.0006
- Björnsson, Gunnar and Derk Pereboom, 2016, “Traditional and Experimental Approaches to Free Will and Moral Responsibility”, in Sytsma and Buckwalter 2016: 142–157. doi:10.1002/9781118661666.ch9
- Bloomfield, Paul, 2000, “Virtue Epistemology and the Epistemology of Virtue”, *Philosophy and Phenomenological Research*, 60(1): 23–43. doi:10.2307/2653426
- , 2001, *Moral Reality*, New York: Oxford University Press. doi:10.1093/0195137132.001.0001
- , 2014, “Some Intellectual Aspects of the Cardinal Virtues”, in *Oxford Studies in Normative Ethics*, volume 3, Mark Timmons (ed.), pp. 287–313. doi:10.1093/acprof:oso/9780199685905.003.0013
- Boorse, Christopher, 1975, “On the Distinction between Disease and Illness”, *Philosophy and Public Affairs*, 5(1): 49–68.
- Boyd, Richard, 1988, “How to Be a Moral Realist”, in Sayre-McCord 1988b: 181–228.
- Brandt, Richard B., 1954, *Hopi Ethics: A Theoretical Analysis*, Chicago: The University of Chicago Press.
- , 1959, *Ethical Theory: The Problems of Normative and Critical Ethics*, Englewood Cliff, NJ: Prentice-Hall.
- Bratman, Michael E., 1996, “Identification, Decision, and Treating as a

- Reason”, *Philosophical Topics*, 24(2): 1–18. doi:10.5840/philtopics19962429
- Brink, David Owen, 1989, *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511624612
- Broad, C.D., 1930, *Five Types of Ethical Theory*, New York: Harcourt, Brace.
- , 1950, “Egoism as a Theory of Human Motives”, *The Hibbert Journal*, 48: 105–114. Reprinted in his *Ethics and the History of Philosophy: Selected Essays*, London: Routledge and Kegan Paul, 1952, 218–231.
- Cameron, C. Daryl, B. Keith Payne, and John M. Doris, 2013, “Morality in High Definition: Emotion Differentiation Calibrates the Influence of Incidental Disgust on Moral Judgments”, *Journal of Experimental Social Psychology*, 49(4): 719–725. doi:10.1016/j.jesp.2013.02.014
- Campbell, C.A., 1951, “Is ‘Freewill’ a Pseudo-problem?” *Mind*, 60(240): 441–465. doi:10.1093/mind/LX.240.441
- Cappelen, Herman, 2012, *Philosophy Without Intuitions*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199644865.001.0001
- Cervone, Daniel and Yuichi Shoda (eds.), 1999, *The Coherence of Personality: Social-Cognitive Bases of Consistency, Variability, and Organization*, New York and London: Guilford Press.
- Chiesi, Harry L., George J. Spilich, and James F. Voss, 1979, “Acquisition of domain-related information in relation to high and low domain knowledge”, *Journal of verbal learning and verbal behavior*, 18(3): 257–273. doi:10.1016/S0022-5371(79)90146-4
- Cialdini, Robert B., Stephanie L. Brown, Brian P. Lewis, Carol Luce and Stephen L. Neuberg, 1997, “Reinterpreting the Empathy-Altruism Relationship: When One into One Equals Oneness”, *Journal of Personality and Social Psychology*, 73(3), 481–494. doi:10.1037/0022-3514.73.3.481

- Cova, Florian and Yasuko Kitano, 2013, “Experimental Philosophy and the Compatibility of Free Will and Determinism: A Survey”, *Annals of the Japan Association for Philosophy of Science*, 22: 17–37. doi:10.4288/jafpos.22.0\_17
- Cova, Florian, Maxime Bertoux, Sacha Bourgeois-Gironde, and Bruno Dubois, 2012, “Judgments about Moral Responsibility and Determinism in Patients with Behavioural Variant of Frontotemporal Dementia: Still Compatibilists”, *Consciousness and Cognition*, 21(2): 851–864. doi:10.1016/j.concog.2012.02.004
- Cuneo, Terence, 2014, *Speech and Morality: On the Metaethical Implications of Speaking*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198712725.001.0001
- Darley, John M. and C. Daniel Batson, 1973, “‘From Jerusalem to Jericho’: A Study of Situational and Dispositional Variables In Helping Behavior”, *Journal of Personality and Social Psychology*, 27(1): 100–108. doi:10.1037/h0034449
- Decety, Jean and Thalia Wheatley (eds.), 2015, *The Moral Brain: A Multidisciplinary Perspective*, Cambridge, MA: MIT Press.
- Deery, Oisín and Eddy Nahmias, 2017, “Defeating Manipulation Arguments: Interventionist Causation and Compatibilist Sourcehood”, *Philosophical Studies*, 174(5): 1255–1276. doi:10.1007/s11098-016-0754-8.
- Dennett, Daniel C., 1984, *Elbow Room: The Varieties of Free Will Worth Wanting*, Cambridge, MA: MIT Press.
- DePaul, Michael, 1999, “Character Traits, Virtues, and Vices: Are There None?” in *Proceedings of the 20th World Congress of Philosophy*, v. 1, Bowling Green, OH: Philosophy Documentation Center, pp. 141–157.
- Deutsch, Max, 2015, *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*, Cambridge, MA: MIT Press. doi:10.7551/mitpress/9780262028950.001.0001

- Dixon, Thomas, 2008, *The Invention of Altruism: Making Moral Meanings in Victorian Britain*, Oxford: Oxford University Press. doi:10.5871/bacad/9780197264263.001.0001
- Donnellan, M. Brent, Richard E. Lucas, and William Fleeson (eds.), 2009, “Personality and Assessment at Age 40: Reflections on the Past Person-Situation Debate and Emerging Directions of Future Person-Situation Integration and Assessment at Age 40”, *Journal of Research in Personality*, special issue, 43(2): 117–290.
- Doris, John M., 1998, “Persons, Situations, and Virtue Ethics”, *Noûs*, 32(4): 504–530. doi:10.1111/0029-4624.00136
- , 2002, *Lack of Character: Personality and Moral Behavior*, New York: Cambridge University Press. doi:10.1017/CBO9781139878364
- , 2005, “Précis” and “Replies: Evidence and Sensibility”, *Philosophy and Phenomenological Research*, 71(3): 632–5, 656–77. doi:10.1111/j.1933-1592.2005.tb00479.x
- , 2006, “Out of Character: On the Psychology of Excuses in the Criminal Law”, in H. LaFollette (ed.), *Ethics in Practice*, third edition, Oxford: Blackwell Publishing.
- , 2010, “Heated Agreement: *Lack of Character* as *Being for the Good*”, *Philosophical Studies*, 148(1): 135–46. doi:10.1007/s11098-010-9507-2
- , 2015, *Talking to Our Selves: Reflection, Ignorance, and Agency*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199570393.001.0001
- , forthcoming, *Character Trouble: Undisciplined Essays on Personality and Agency*, Oxford: Oxford University Press.
- , in preparation, “Making Good: In Search of Moral Expertise”.
- Doris, John M. and Alexandra Plakias, 2008, “How to Argue about Disagreement: Evaluative Diversity and Moral Realism”, in Sinnott-Armstrong 2008b: 303–353.
- Doris, John M. and Jesse J. Prinz, 2009, “Review of K. Anthony Appiah,

- Experiments in Ethics*”, *Notre Dame Philosophical Reviews*, 2009-10-03. URL = <<http://ndpr.nd.edu/news/experiments-in-ethics/>>
- Doris, John M. and The Moral Psychology Research Group (eds.), 2010, *The Moral Psychology Handbook*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199582143.001.0001
- Doris, John M. and Stephen P. Stich, 2005, “As a Matter of Fact: Empirical Perspectives on Ethics”, in Frank Jackson and Michael Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*, Oxford: Oxford University Press.
- Dunker, Karl, 1939, “Ethical Relativity? (An Enquiry into the Psychology of Ethics)”, *Mind*, 48(189): 39–53. doi:10.1093/mind/XLVIII.189.39
- Ellsworth, Phoebe C., 1994, “Sense, Culture, and Sensibility”, in Shinobu Kitayama and Hazel Rose Markus (eds.), *Emotion and Culture: Empirical Studies of Mutual Influence*, Washington: American Psychological Association.
- Enoch, David, 2009, “How is Moral Disagreement a Problem for Realism?” *The Journal of Ethics*, 13(1): 15–50. doi:10.1007/s10892-008-9041-z.
- , 2011, *Taking Morality Seriously: A Defense of Robust Realism*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199579969.001.0001
- Ericsson, K. Anders, 2014, “Why Expert Performance Is Special and Cannot Be Extrapolated From Studies of Performance in the General Population: A Response to Criticisms”, *Intelligence*, 45: 81–103. doi:10.1016/j.intell.2013.12.001
- Ericsson, K. Anders, Ralf Th. Krampe, and Clemens Tesch-Römer, 1993, “The Role of Deliberate Practice in the Acquisition of Expert Performance”, *Psychological Review*, 100(3): 363–406. doi:10.1037/0033-295X.100.3.363
- Feinberg, Joel, 1965 [1999], “Psychological Egoism”, in *Reason and Responsibility*, Joel Feinberg (ed.), Belmont, CA: Dickenson

- Publishing. Reprinted in various editions including the tenth, co-edited with Russ Shafer-Landau, Belmont, CA: Wadsworth, 1999. Based on materials composed for philosophy students at Brown University, 1958.
- Feltz, Adam and Florian Cova, 2014, “Moral Responsibility and Free Will: A Meta-Analysis”, *Consciousness and Cognition*, 30: 234–246. doi:10.1016/j.concog.2014.08.012
- Feltz, Adam and Melissa Millan, 2013, “An Error Theory for Compatibilist Intuitions”, *Philosophical Psychology*, 28(4): 529–555. doi:10.1080/09515089.2013.865513
- Figdor, Carrie and Mark Phelan, 2015, “Is Free Will Necessary for Moral Responsibility? A Case for Rethinking Their Relationship and the Design of Experimental Studies in Moral Psychology”, *Mind and Language*, 30(5): 603–627. doi:10.1111/mila.12092
- Fischer, John Martin, 1994, *The Metaphysics of Free Will*, Oxford: Blackwell.
- Flanagan, Owen, 1991, *Varieties of Moral Personality: Ethics and Psychological Realism*, Cambridge, MA: Harvard University Press.
- , 2009, “Moral Science? Still Metaphysical After All These Years”, in Darcia Narvaez and Daniel K. Lapsley (eds.), *Personality, Identity, and Character*, Cambridge: Cambridge University Press, pp. 52–78.
- Frankfurt, Harry, 1988, *The Importance of What We Care About*, Cambridge: Cambridge University Press.
- Fraser, Ben and Marc Hauser, 2010, “The Argument from Disagreement and the Role of Cross-Cultural Empirical Data”, *Mind and Language*, 25(5): 541–560. doi:10.1111/j.1468-0017.2010.01400.x
- Fulford, K.W.M., 1989, *Moral Theory and Medical Practice*, Cambridge: Cambridge University Press.
- Gigerenzer, Gerd, 2000, *Adaptive Thinking: Rationality in the Real World*, New York: Oxford University Press. doi:10.1093/acprof:oso/9780195153729.001.0001

- Gigerenzer, Gerd, Peter M. Todd, and the ABC Research Group., 1999, *Simple Heuristics that Make Us Smart*, New York: Oxford University Press.
- Gilovich, Thomas, Dale W. Griffin, and Daniel Kahneman (eds.), 2002, *Heuristics and Biases: The Psychology of Intuitive Judgment*, New York: Cambridge University Press.
- Glickman, Mark E., 1995, “A Comprehensive Guide to Chess Ratings”, *American Chess Journal*, 3: 59–102.
- Goldman, Alvin I., 1970, *A Theory of Human Action*, Englewood-Cliffs, NJ: Prentice-Hall.
- Haidt, Jonathan, Silvia Helena Koller, and Maria G. Dias, 1993, “Affect, Culture, and Morality, Or Is It Wrong to Eat Your Dog?” *Journal of Personality and Social Psychology*, 65(4): 613–28. doi:10.1037/0022-3514.65.4.613
- Haji, Ishtiyaque, 2002, “Compatibilist Views of Freedom and Responsibility”, in Robert Kane (ed.), *The Oxford Handbook of Free Will*, New York: Oxford University Press.
- Hambrick, David Z., Frederick L. Oswald, Erik M. Altmann, Elizabeth J. Meinz, Fernand Gobet, and Guillermo Campitelli, 2014, “Deliberate Practice: Is That All It Takes to Become an Expert?” *Intelligence*, 45: 34–45. doi:10.1016/j.intell.2013.04.001
- Harman, Gilbert, 1999, “Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error”, *Proceedings of the Aristotelian Society*, 99: 315–331.
- , 2000, “The Nonexistence of Character Traits”, *Proceedings of the Aristotelian Society*, 100: 223–226. doi:10.1111/j.0066-7372.2003.00013.x
- , 2009, “Skepticism about Character Traits”, *The Journal of Ethics*, 13(2–3): 235–242. doi:10.1007/s10892-009-9050-6
- Heine, Steven J., 2008, *Cultural Psychology*, New York: W.W. Norton.
- Helzer, Erik G. and David A. Pizarro, 2011, “Dirty Liberals! Reminders of

- Physical Cleanliness Influence Moral and Political Attitudes”, *Psychological Science*, 22(4): 517–522. doi:10.1177/0956797611402514
- Henrich, Joseph, 2015, *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*, Princeton, NJ: Princeton University Press.
- Hobbes, Thomas, 1651 [1981], *Leviathan: Edited with an Introduction by C.B. Macpherson*, London: Penguin Books.
- Horowitz, Tamara, 1998, “Philosophical Intuitions and Psychological Theory”, in Michael R. DePaul and William Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*, Lanham, Maryland: Rowman and Littlefield.
- Hursthouse, Rosalind, 1999, *On Virtue Ethics*, Oxford and New York: Oxford University Press. doi:10.1093/0199247994.001.0001
- Isen, Alice M. and Paula F. Levin, 1972, “Effect of Feeling Good on Helping: Cookies and Kindness”, *Journal of Personality and Social Psychology*, 21(3): 384–388. doi:10.1037/h0032317
- Jackson, Frank, 1998, *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, New York: Oxford University Press. doi:10.1093/0198250614.001.0001
- Jackson, Frank and Philip Pettit, 1995, “Moral Functionalism and Moral Motivation”, *Philosophical Quarterly*, 45(178): 20–40. doi:10.2307/2219846
- Jacobson, Daniel, 2005, “Seeing By Feeling: Virtues, Skills, and Moral Perception”, *Ethical Theory and Moral Practice*, 8(4): 387–409. doi:10.1007/s10677-005-8837-1
- Joyce, Richard, 2006, *The Evolution of Morality*, Cambridge, MA: MIT Press.
- Kahneman, Daniel, 2011, *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, 1982, *Judgment*

- Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511809477
- Kamtekar, Rachana, 2004, "Situationism and Virtue Ethics on the Content of Our Character", *Ethics*, 114(3): 458–91. doi:10.1086/381696
- Kane, Robert, 1996, *The Significance of Free Will*, Oxford: Oxford University Press. doi:10.1093/0195126564.001.0001
- , 1999, "Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism", *Journal of Philosophy*, 96(5): 217–240. doi:10.5840/jphil199996537
- , 2002, "Introduction: The Contours of Contemporary Free Will Debates", in Robert Kane (ed.), *The Oxford Handbook of Free Will*, New York: Oxford University Press.
- Kant, Immanuel, 1785 [1949], *Fundamental Principles of the Metaphysics of Morals*, Translated by Thomas K. Abbott. Englewood Cliffs, NJ: Prentice Hall / Library of Liberal Arts.
- Kitayama, Shinobu and Hazel Rose Markus, 1999, "Yin and Yang of the Japanese Self: The Cultural Psychology of Personality Coherence", in Cervone and Shoda 1999: ch. 8.
- Kitayama, Shinobu and Dov Cohen, 2010, *Handbook of Cultural Psychology*, New York: Guilford Press.
- Kitcher, Philip, 2010, "Varieties of Altruism", *Economics and Philosophy*, 26(2): 121–148. doi:10.1017/S0266267110000167
- , 2011, *The Ethical Project*, Cambridge, MA: Harvard University Press.
- Knobe, Joshua, 2003a, "Intentional Action and Side Effects in Ordinary Language", *Analysis*, 63(279): 190–193. doi:10.1111/1467-8284.00419
- , 2003b, "Intentional Action in Folk Psychology: An Experimental Investigation", *Philosophical Psychology*, 16(2): 309–324. doi:10.1080/09515080307771
- , 2006, "The Concept of Intentional Action: A Case Study in the Uses

- of Folk Psychology", *Philosophical Studies*, 130(2): 203–231. doi:10.1007/s11098-004-4510-0
- , 2010, "Person as Scientist, Person as Moralist", *Behavioral and Brain Sciences*, 33(4): 315–329. doi:10.1017/S0140525X10000907
- , 2014, "Free Will and the Scientific Vision", in Edouard Machery and Elizabeth O'Neill (eds.), *Current Controversies in Experimental Philosophy*, New York and London: Routledge.
- Knobe, Joshua and Brian Leiter, 2007, "The Case for Nietzschean Moral Psychology", in Brian Leiter and Neil Sinhababu (eds.) *Nietzsche and Morality*, Oxford: Oxford University Press. 83–109.
- Kruger, Justin and David Dunning, 1999, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments", *Journal of Personality and Social Psychology*, 77(6): 1121–1134. doi:10.1037/0022-3514.77.6.1121
- Kupperman, Joel J., 2001, "The Indispensability of Character", *Philosophy*, 76(02): 239–50. doi:10.1017/S0031819101000250
- Ladd, John, 1957, *The Structure of a Moral Code: A Philosophical Analysis of Ethical Discourse Applied to the Ethics of the Navaho Indians*, Cambridge, MA: Harvard University Press.
- LaFollette, Hugh (ed.), 2000, *The Blackwell Guide to Ethical Theory*, Oxford: Blackwell Publishing.
- Leikas, Sointu, Jan-Erik Lönnqvist, and Markku Verkasalo, 2012, "Persons, Situations, and Behaviors: Consistency and Variability of Different Behaviors in Four Interpersonal Situations", *Journal of Personality and Social Psychology*, 103(6): 1007–1022. doi:10.1037/a0030385
- Lerner, Jennifer S., Julie H. Goldberg, and Philip E. Tetlock, 1998, "Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility", *Personality and Social Psychology Bulletin*, 24(6): 563–574. doi:10.1177/0146167298246001

- Lewis, David, 1989, “Dispositional Theories of Value”, *Proceedings of the Aristotelian Society*, 63 (supp): 113–37.
- Liao, S. Matthew, Alex Wiegmann, Joshua Alexander, and Gerard Vong, 2012, “Putting the Trolley in Order: Experimental Philosophy and the Loop Case”, *Philosophical Psychology*, 25(5): 661–671. doi:10.1080/09515089.2011.627536
- Loeb, Don., 1998, “Moral Realism and the Argument from Disagreement”, *Philosophical Studies*, 90(3): 281–303. doi:10.1023/A:1004267726440
- Machery, Edouard, 2010, “The Bleak Implications of Moral Psychology”, *Neuroethics*, 3(3): 223–231. doi:10.1007/s12152-010-9063-7
- Machery, Edouard and John M. Doris, forthcoming, “An Open Letter to Our Students: Going Interdisciplinary”, in Voyer and Tarantola forthcoming.
- MacIntyre, Alasdair, 1967, “Egoism and Altruism”, in Paul Edwards (ed.), *The Encyclopedia of Philosophy*, vol. 2, first edition, New York: Macmillan, pp. 462–466.
- Mackie, J.L., 1977, *Ethics: Inventing Right and Wrong*, New York: Penguin Books.
- Macnamara, Brooke N., David Z. Hambrick, and Frederick L. Oswald, 2014, “Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions: A Meta-Analysis”, *Psychological Science*, 25(8): 1608–1618. doi:10.1177/0956797614535810
- Markus, Hazel R. and Shinobu Kitayama, 1991, “Culture and the Self: Implications for Cognition, Emotion, and Motivation”, *Psychological Review*, 98(2): 224–253. doi:10.1037/0033-295X.98.2.224
- May, Joshua, 2011a, “Psychological Egoism”, *Internet Encyclopedia of Philosophy*. URL = <<https://www.iep.utm.edu/psychego/>>
- , 2011b, “Egoism, Empathy, and Self-Other Merging”, *Southern Journal of Philosophy*, 49(s1): 25–39. doi:10.1111/j.2041-6962.2011.00055.x

- , 2011c, “Relational Desires and Empirical Evidence against Psychological Egoism: On Psychological Egoism”, *European Journal of Philosophy*, 19(1): 39–58. doi:10.1111/j.1468-0378.2009.00379.x
- McGrath, Sarah, 2008, “Moral Disagreement and Moral Expertise”, in *Oxford Studies in Metaethics*, volume 3, Russ Shafer-Landau (ed.), New York: Oxford University Press, pp. 87–108.
- , 2011, “Skepticism about Moral Expertise as a Puzzle for Moral Realism”, *Journal of Philosophy*, 108(3): 111–137. doi:10.5840/jphil201110837
- Mehl, Matthias R., Kathryn L. Bollich, John M. Doris, and Simine Vazire, 2015, “Character and Coherence: Testing the Stability of Naturalistically Observed Daily Moral Behavior”, in Miller et al. 2015: 630–51. doi:10.1093/acprof:oso/9780190204600.003.0030
- Mele, Alfred R., 2006, *Free Will and Luck*, New York: Oxford University Press. doi:10.1093/0195305043.001.0001
- , 2013, “Manipulation, Moral Responsibility, and Bullet Biting”, *Journal of Ethics*, 17(3): 167–84. doi:10.1007/s10892-013-9147-9
- Merritt, Maria W., 2000, “Virtue Ethics and Situationist Personality Psychology”, *Ethical Theory and Moral Practice*, 3(4): 365–83. doi:10.1023/A:1009926720584
- , 2009, “Aristotelean Virtue and the Interpersonal Aspect of Ethical Character”, *Journal of Moral Philosophy*, 6(1): 23–49. doi:10.1163/174552409X365919
- Meritt, Maria W., John M. Doris, and Gilbert Harman, 2010, “Character”, in Doris et al. 2010: 355–401.
- Milgram, Stanley, 1974, *Obedience to Authority*, New York: Harper and Row.
- Miller, Christian B., 2003, “Social Psychology and Virtue Ethics”, *The Journal of Ethics*, 7(4): 365–92. doi:10.1023/A:1026136703565
- , 2013, *Moral Character: An Empirical Theory*, Oxford: Oxford

- University Press. doi:10.1093/acprof:oso/9780199674350.001.0001
- , 2014, *Character and Moral Psychology*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199674367.001.0001
- Miller, Christian B., R. Michael Furr, Angela Knobel, and William Fleeson (eds.), 2015, *Character: New Directions from Philosophy, Psychology, and Theology*, New York: Oxford University Press. doi:10.1093/acprof:oso/9780190204600.001.0001
- Mischel, Walter, 1968, *Personality and Assessment*, New York: John J. Wiley and Sons.
- , 1999, “Personality Coherence and Dispositions in a Cognitive-Affective Personality System (CAPS) Approach”, in Cervone and Shoda 1999: ch. 2.
- Montmarquet, James, 2003, “Moral Character and Social Science Research”, *Philosophy*, 78(03): 355–368. doi:10.1017/S0031819103000342
- Moody-Adams, Michele M., 1997, *Fieldwork in Familiar Places: Morality, Culture, and Philosophy*, Cambridge, MA: Harvard University Press.
- Murphy, Dominic, 2006, *Psychiatry in the Scientific Image*, Cambridge, MA: MIT Press.
- Murray, Dylan and Eddy Nahmias, 2014, “Explaining Away Incompatibilist Intuitions”, *Philosophy and Phenomenological Research*, 88(2): 434–467. doi:10.1111/j.1933-1592.2012.00609.x
- Murray, Dylan and Tania Lombrozo, 2016, “Effects of Manipulation on Attributions of Causation, Free Will, and Moral Responsibility”, *Cognitive Science*, 41(2): 447–481. doi: 10.1111/cogs.12338.
- Nado, Jennifer, 2016, “The Intuition Deniers”, *Philosophical Studies*, 173(3): 781–800. doi:10.1007/s11098-015-0519-9.
- Nagel, Thomas, 1970, *The Possibility of Altruism*, Oxford: Oxford University Press.
- , 1986, *The View From Nowhere*, New York and Oxford: Oxford

- University Press.
- Nahmias, Eddy, 2011, “Intuitions about Free Will, Determinism, and Bypassing”, in Robert Kane (ed.), *The Oxford Handbook of Free Will*, second edition, Oxford: Oxford University Press.
- Nahmias, Eddy, Stephen G. Morris, Thomas Nadelhoffer, and Jason Turner, 2009, “Is Incompatibilism Intuitive?” *Philosophy and Phenomenological Research*, 73(1): 28–53. doi:10.1111/j.1933-1592.2006.tb00603.x
- Nichols, Shaun and Joshua Knobe, 2007, “Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions”, *Noûs*, 41(4): 663–685. doi:10.1111/j.1468-0068.2007.00666.x
- Nisbett, Richard E., 1998, “Essence and Accident”, in John M. Darley and Joel Cooper (eds.), *Attribution and Social Interaction: The Legacy of Edward E. Jones*, Washington: American Psychological Association.
- , 2003, *The Geography of Thought: How Asians and Westerners Think Differently ... and Why*, New York: Free Press.
- Nisbett, Richard E. and Eugene Borgida, 1975, “Attribution and the psychology of prediction”, *Journal of Personality and Social Psychology*, 32(5): 932–943. doi:10.1037/0022-3514.32.5.932
- Nisbett, Richard E. and Dov Cohen, 1996, *Culture of Honor: The Psychology of Violence in the South*, Boulder, CO: Westview Press.
- Nisbett, Richard E. and Lee Ross, 1980, *Human Inference: Strategies and Shortcomings of Social Judgment*, Englewood Cliffs, NJ: Prentice-Hall.
- O’Connor, Timothy, 2000, *Persons and Causes: The Metaphysics of Free Will*, New York: Oxford University Press. doi:10.1093/019515374X.001.0001
- Olin, Lauren and John M. Doris, 2014, “Vicious Minds: Virtue Epistemology, Cognition, and Skepticism”, *Philosophical Studies*, 168(3): 665–92. doi:10.1007/s11098-013-0153-3
- Pereboom, Derk, 2001, *Living Without Free Will*, Cambridge: Cambridge

- University Press. doi:10.1017/CBO9780511498824
- , 2014, *Free Will, Agency, and Meaning in Life*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199685516.001.0001
- Petrinovich, Lewis and Patricia O'Neill, 1996, "Influence of Wording and Framing Effects on Moral Intuitions", *Ethology and Sociobiology*, 17(3): 145–171. doi:10.1016/0162-3095(96)00041-6
- Phillips, Jonathan and Alex Shaw, 2014, "Manipulating Morality: Third-Party Intentions Alter Moral Judgments by Changing Causal Reasoning", *Cognitive Science*, 39(6): 1320–47. doi:10.1111/cogs.12194
- Pink, Thomas, 2004, *Free Will: A Very Short Introduction*, New York: Oxford University Press. doi:10.1093/acprof:oso/9780192853585.001.0001
- Prinz, Jesse J., 2009, "The Normativity Challenge: Cultural Psychology Provides the Real Threat to Virtue Ethics", *The Journal of Ethics*, 13(2–3): 117–144. doi:10.1007/s10892-009-9053-3
- Pust, Joel, 2000, *Intuitions as Evidence*, New York: Garland Publishing.
- Rachels, James, 2000, "Naturalism", in LaFollette 2000: 74–91.
- , 2003, *The Elements of Moral Philosophy*, fourth edition, New York: McGraw-Hill.
- Railton, Peter, 1986a, "Facts and Values", *Philosophical Topics*, 14(2): 5–31. doi:10.5840/philtopics19861421
- , 1986b, "Moral Realism", *Philosophical Review*, 95(2): 163–207. doi:10.2307/2185589
- Rawls, John, 1951, "Outline of a Decision Procedure for Ethics", *Philosophical Review*, 60(2): 177–97. doi:10.2307/2181696
- , 1971, *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Rosati, Connie S., 1995, "Persons, Perspectives, and Full Information Accounts of the Good", *Ethics*, 105(2): 296–325. doi:10.1086/293702

- Rose, David and Shaun Nichols, 2013, "The Lesson of Bypassing", *Review of Philosophy and Psychology*, 4(4): 599–619. doi:10.1007/s13164-013-0154-3
- Ross, Lee and Richard E. Nisbett, 1991, *The Person and the Situation: Perspectives of Social Psychology*, Philadelphia: Temple University Press.
- Russell, Daniel C., 2009, *Practical Intelligence and the Virtues*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199565795.001.0001
- , 2015, "From Personality to Character to Virtue", in Mark Alfano (ed.), *Current Controversies in Virtue Theory*, New York: Routledge, pp. 91–106.
- Samuels, Richard and Stephen Stich, 2002, "Rationality", *Encyclopedia of Cognitive Science*, Chichester: Wiley. doi:10.1002/0470018860.s00171
- Samuels, Steven M. and William D. Casebeer, 2005, "A social psychological view of morality: Why knowledge of situational influences on behaviour can improve character development practices", *Journal of Moral Education*, 34(1): 73–87. doi:10.1080/03057240500049349
- Sarkissian, Hagop, 2010, "Minor Tweaks, Major Payoffs: The Problems and Promise of Situationism in Moral Philosophy", *Philosophers' Imprint*, 10(9). URL = <http://hdl.handle.net/2027/spo.3521354.0010.009>
- Sarkissian, Hagop and Jennifer Cole Wright (eds.), 2014, *Advances in Experimental Moral Psychology*, London: Bloomsbury Press.
- Sayre-McCord, Geoffrey, 1988a, "Introduction: The Many Moral Realisms", in Sayre-McCord 1988b: 1–24.
- (ed.), 1988b, *Essays in Moral Realism*, Ithaca and London: Cornell University Press.
- , 2015, "Moral Realism", *The Stanford Encyclopedia of Philosophy*



- (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/moral-realism/>
- Schnall, Simone, Jonathan Haidt, Gerald L. Clore, and Alexander H. Jordan, 2008a, “Disgust as Embodied Moral Judgment”, *Personality and Social Psychology Bulletin*, 34(8): 1069–1109. doi:10.1177/0146167208317771
- Schnall, Simone, Jennifer Benton, and Sophie Harvey, 2008b, “With a Clean Conscience: Cleanliness Reduces the Severity of Moral Judgments”, *Psychological Science*, 19(12): 1219–1222. doi:10.1111/j.1467-9280.2008.02227.x
- Schwitzgebel, Eric and Fiery Cushman, 2011, “Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers”, *Mind and Language*, 27(2): 135–153. doi:10.1111/j.1468-0017.2012.01438.x.
- , 2015, “Philosophers’ Biased Judgments Persist Despite Training, Expertise, and Reflection”, *Cognition*, 141: 127–137. doi:10.1016/j.cognition.2015.04.015
- Shafer-Landau, R., 2003, *Moral Realism: A Defence*, Oxford: Clarendon Press. doi:10.1093/0199259755.001.0001
- Sherman, Ryne A., Christopher S. Nave, and David C. Funder, 2010, “Situational Similarity and Personality Predict Behavioral Consistency”, *Journal of Personality and Social Psychology*, 99(2): 330–343. doi:10.1037/a0019796
- Shweder, Richard A., and Edmund J. Bourne, 1982, “Does the Concept of the Person Vary Cross-Culturally?” in Anthony J. Marsella and Geoffrey M. White (eds.), *Cultural Conceptions of Mental Health and Therapy*, Boston, MA: D. Reidel Publishing.
- Singer, Peter, 1974, “Sidgwick and Reflective Equilibrium”, *Monist*, 58(3): 490–517. doi:10.5840/monist197458330
- Sinnott-Armstrong, Walter P., 2005, “Moral Intuitionism Meets Empirical Psychology”, in Terry Horgan and Mark Timmons (eds.), *Metaethics*

- After Moore*, New York: Oxford University Press. doi:10.1093/acprof:oso/9780199269914.003.0016
- (ed.), 2008a, *Moral Psychology, Vol. 1, The Evolution of Morality: Adaptations and Innateness*, Cambridge, MA: MIT Press.
- (ed.), 2008b, *Moral Psychology, Vol. 2, The Cognitive Science of Morality: Intuition and Diversity*, Cambridge, MA: MIT Press.
- (ed.), 2008c, *Moral Psychology, Vol. 3, The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, Cambridge, MA: MIT Press.
- (ed.), 2014, *Moral Psychology, Vol. 4, Free Will and Moral Responsibility*, Cambridge, MA: MIT Press.
- Slote, Michael Anthony, 2013, “Egoism and Emotion”, *Philosophia*, 41(2): 313–335. doi:10.1007/s11406-013-9434-5
- Smilansky, Saul, 2003, “Compatibilism: the Argument from Shallowness”, *Philosophical Studies*, 115(3): 257–282. doi:10.1023/A:1025146022431
- Smith, Adam, 1759 [1853], *The Theory of Moral Sentiments*, London: Henry G. Bohn. Originally published 1759,
- Smith, Michael, 1994, *The Moral Problem*, Cambridge: Basil Blackwell.
- Snare, F.E., 1980, “The Diversity of Morals” *Mind*, 89(355): 353–369. doi:10.1093/mind/LXXXIX.355.353
- Snow, Nancy E., 2010, *Virtue as Social Intelligence: An Empirically Grounded Theory*, London and New York: Routledge.
- Sober, Elliott and David Sloan Wilson, 1998, *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, MA: Harvard University Press.
- Solomon, Robert C., 2003, “Victims of Circumstances? A Defense of Virtue Ethics in Business”, *Business Ethics Quarterly*, 13(1): 43–62. doi:10.5840/beq20031314
- , 2005, “‘What’s Character Got to Do with It?’”, *Philosophy and Phenomenological Research*, 71(3): 648–655. doi:10.1111/j.1933-

- 1592.2005.tb00478.x
- Sosa, Ernest, 2007, “Intuitions: Their Nature and Epistemic Efficacy”, *Grazer Philosophische Studien*, 74(1): 51–67.  
doi:10.1163/9789401204651\_004
- , 2009, “Situations Against Virtues: The Situationist Attack on Virtue Theory”, in Chrysostomos Mantzavinos (ed.), *Philosophy of the Social Sciences: Philosophical Theory and Scientific Practice*, New York: Cambridge University Press. 274–290.  
doi:10.1017/CBO9780511812880.021
- Sreenivasan, Gopal, 2002, “Errors about errors: Virtue theory and trait attribution”, *Mind*, 111(441): 47–68. doi:10.1093/mind/111.441.47
- Sripada, Chandra Sekhar, 2012, “What Makes a Manipulated Agent Unfree?” *Philosophy and Phenomenological Research*, 85(3): 563–93. doi:10.1111/j.1933-1592.2011.00527.x
- Stich, Stephen, 1990, *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*, Cambridge, MA: The MIT Press.
- Stich, Stephen, John M. Doris, and Erica Roedder, 2010, “Altruism”, in Doris et al. 2010: 147–205.
- Stich, Stephen and Kevin P. Tobia, 2016, “Experimental Philosophy and the Philosophical Tradition”, in Sytsma and Buckwalter 2016: 3–21.  
doi:10.1002/9781118661666.ch1
- , 2018, “Intuition and Its Critics”, in Stuart, Fehige, and Brown 2018: ch. 21.
- Stichter, Matt, 2007, “Ethical Expertise: The Skill Model of Virtue”, *Ethical Theory and Moral Practice*, 10(2): 183–194.  
doi:10.1007/s10677-006-9054-2
- , 2011, “Virtues, Skills, and Right Action”, *Ethical Theory and Moral Practice*, 14(1): 73–86. doi:10.1007/s10677-010-9226-y
- Strawson, P.F., 1982, “Freedom and Resentment”, in Gary Watson (ed.), *Free Will*, New York: Oxford University Press. Originally published,





- 1962,
- Strawson, Galen, 1986, *Freedom and Belief*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199247493.001.0001
- Strohming, Nina, Richard L. Lewis and David E. Meyer, 2011, “Divergent Effects of Different Positive Emotions on Moral Judgment”, *Cognition*, 119(2): 295–300.  
doi:10.1016/j.cognition.2010.12.012
- Stuart, Michael T., Yiftach Fehige, and James Robert Brown (eds.), 2018, *The Routledge Companion to Thought Experiments*, New York: Routledge.
- Sturgeon, Nicholas L., 1988, “Moral Explanations”, in Sayre-McCord 1988b: 229–255.
- Sumner, William Graham, 1908 [1934], *Folkways*, Boston: Ginn and Company.
- Sunstein, Cass R., 2005, “Moral Heuristics”, *Behavioral and Brain Sciences*, 28(4): 531–42. doi:10.1017/S0140525X05000099
- Swanton, Christine, 2003, *Virtue Ethics: A Pluralistic View*, Oxford: Oxford University Press. doi:10.1093/0199253889.001.0001
- Sytsma, Justin and Wesley Buckwalter (eds.), 2016, *A Companion to Experimental Philosophy*, Oxford: Blackwell.
- Tetlock, Philip E., 1999, “Review of *Culture of Honor: The Psychology of Violence in the South* by Robert Nisbett and Dov Cohen”, *Political Psychology*, 20(1): 211–13. doi:10.1111/0162-895X.t01-1-00142
- Tiberius, Valerie, 2015, *Moral Psychology: A Contemporary Introduction*, New York: Routledge.
- Tobia, Kevin Patrick, Gretchen B. Chapman, and Stephen Stich, 2013, “Cleanliness is Next to Morality, Even for Philosophers”, *Journal of Consciousness Studies*, 20(11 and 12): 195–204.
- Tversky, Amos and Daniel Kahneman, 1973, “Availability: A heuristic for judging frequency and probability”, *Cognitive Psychology*, 5(2): 207–232. doi:10.1016/0010-0285(73)90033-9

- , 1981, “The Framing of Decisions and the Psychology of Choice”, *Science*, 211(4481): 453–463. doi:10.1126/science.7455683
- Upton, Candace L., 2009, *Situational Traits of Character: Dispositional Foundations and Implications for Moral Psychology and Friendship*, Lanham, Maryland: Lexington Books.
- Vargas, Manuel, 2005a, “The Revisionist’s Guide to Responsibility”, *Philosophical Studies*, 125(3): 399–429. doi:10.1007/s11098-005-7783-z
- , 2005b, “Responsibility and the Aims of Theory: Strawson and Revisionism”, *Pacific Philosophical Quarterly*, 85(2): 218–241. doi:10.1111/j.0279-0750.2004.00195.x
- Valdesolo, Piercarlo and David DeSteno, 2006, “Manipulations of Emotional Context Shape Moral Judgment”, *Psychological Science*, 17(6): 476–477. doi:10.1111/j.1467-9280.2006.01731.x
- Velleman, J. David, 1992, “What Happens When Someone Acts?” *Mind*, 101(403): 461–81. doi:10.1093/mind/101.403.461
- Voyer, Benjamin G. and Tor Tarantola (eds.), forthcoming, *Moral Psychology: A Multidisciplinary Guide*, Springer.
- Vranas, Peter B.M., 2005, “The Indeterminacy Paradox: Character Evaluations and Human Psychology”, *Noûs*, 39(1): 1–42.
- Watson, Gary, 1996, “Two Faces of Responsibility”, *Philosophical Topics* 24(2): 227–48. doi: 10.5840/philtopics199624222
- Webber, Jonathan, 2006a, “Character, Consistency, and Classification”, *Mind*, 115(459): 651–658. doi:10.1093/mind/fzl651
- , 2006b, “Virtue, Character and Situation”, *Journal of Moral Philosophy*, 3(2): 193–213. doi:10.1177/1740468106065492
- , 2007a, “Character, Common-Sense, and Expertise”, *Ethical Theory and Moral Practice*, 10(1): 89–104. doi:10.1007/s10677-006-9041-7
- , 2007b, “Character, Global and Local”, *Utilitas*, 19(04): 430–434. doi:10.1017/S0953820807002725
- Westermarck, Edvard, 1906, *Origin and Development of the Moral Ideas*,

2 volumes, New York: MacMillan.

- Wiegmann, Alex, Yasmina Okan, and Jonas Nagel, 2012, “Order Effects in Moral Judgment”, *Philosophical Psychology*, 25(6): 813–836. doi:10.1080/09515089.2011.631995
- Williams, Bernard, 1973, “A Critique of Utilitarianism”, in *Utilitarianism: For and Against*, by J.J.C. Smart and Bernard Williams, Cambridge: Cambridge University Press.
- , 1985, *Ethics and the Limits of Philosophy*, Cambridge, MA: Harvard University Press.
- Woolfolk, Robert L., John M. Doris and John M. Darley, 2006, “Identification, Situational Constraint, and Social Cognition: Studies in the Attribution of Moral Responsibility”, *Cognition*, 100(2), 283–301. doi:10.1016/j.cognition.2005.05.002
- Zhong, Chen-Bo, Brendan Strejcek, and Niro Sivanathan, 2010, “A Clean Self Can Render Harsh Moral Judgment”, *Journal of Experimental Social Psychology*, 46(5): 859–862. doi:10.1016/j.jesp.2010.04.003
- Zimbardo, Philip G., 2007, *The Lucifer Effect: Understanding How Good People Turn Evil*, Oxford: Blackwell Publishing Ltd.

## 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

- Moral Psychology Research Group
- Philosophy Experiments
- Your Morals Blog

## Related Entries

altruism: empirical approaches | egoism | ethics: virtue | experimental moral philosophy | moral realism | practical reason | reasoning: moral

## Notes to Moral Psychology: Empirical Approaches

1. For overviews of this literature see Tiberius (2015) and Alfano (2016). For useful collections of papers see Sinnott-Armstrong (2008a,b,c, 2014), Doris et al. (2010), Decety & Wheatley (2015), and Voyer & Tarantola (forthcoming).

2. For an overview of this work see the SEP entry on experimental moral philosophy. For some recent papers see Sarkissian & Wright 2014.

3. This scientific analogy is not unique to our exposition. Singer (1974: 517; cf. 493) understands Rawls' (1971) method of reflective equilibrium as "leading us to think of our particular moral judgments as data against which moral theories are to be tested ..." As Singer (1974: 493n3) notes, Rawls (1951) in an early paper made the analogy with scientific theory choice explicit. Though the claim that intuitions are used as evidence in philosophy is widely accepted (Pust 2000; Sosa 2007), it has recently been challenged, most notably by Cappelen (2012) and Deutsch (2015). For a useful overview of this debate, see Nado (2016).

4. For more on the appeal to the intuitions of philosophical "experts" see Deutsch (2015: 134–135), Stich & Tobia (2016: 14–17; 2018).

5. Authors such as Baron (1994) and Sunstein (2005) have argued that the distorting influences of "heuristics and biases" uncovered in the recent psychological literature on reasoning, judgment and decision making are widespread in everyday ethical reflection. For overviews of the relevant psychological literature see Nisbett & Ross (1980), Kahneman et al. (1982), Baron (2001), Gilovich et al. (2002), and Kahneman (2011).

6. Rosati (1995) raises analogous concerns with regard to highly idealized "full information accounts" of the good: the ideal, fully informed agent, is likely to be so psychologically different from the actual, very imperfectly informed, agent that there is little reason to think that the actual agent would be motivated by what motivates the ideal agent.

7. There is a wealth of terminological and doctrinal variation here. For example, Haji uses "authenticity" in the neighborhood where Kane uses "ultimacy". Very often, these issues are couched in terms of "freedom" instead of (or together with) "responsibility"; indeed, as Kane (2002: 4–5) observes, the issues are frequently not sharply distinguished. Here, we focus entirely on responsibility attribution, and will be silent on the relation (or lack of a relation) between these issues and issues discussed under the heading of "freedom" or "free will".

8. "Identification" is a troublesome notion, afflicted with more philosophical complexity than can feasibly be operationalized in empirical work. Woolfolk et al. attempted to depict behaviors that could plausibly be construed as manifesting identification, or its lack, without resolving the conceptual issues. For some discussion of the unresolved philosophical difficulties surrounding identification, see Velleman (1992), Bratman (1996), and Watson (1996).

9. The effect appears to be easily replicable. For example, the Woolfolk group (unpublished data) obtained consistent results in variations concerning the positively valenced behavior of kidney donation, and the negatively valenced behavior of an atrocity committed by a soldier under orders.

10. For surveys of the empirical literature, see Doris (2002), Miller (2013, 2014), and Ross and Nisbett (1991).

11. For character skepticism in philosophy, see Doris (1998, 2002, 2005, 2006, 2010, forthcoming) and Doris & Stich (2005). For related positions, see Alfano (2013), Harman (1999, 2000), Machery (2010), Merritt (2000, 2009; Merritt, Doris, & Harman 2010), and Vranas (2005). In psychology, see Mischel (1968) and Ross & Nisbett (1991); a variety of later perspectives are collected in Donnellan et al. (2009).

12. For responses to character skepticism, see Adams (2006), Annas (2005), Appiah (2008: 33–72), Arpaly (2005), Badhwar (2009), Besser-Jones (2008), Flanagan (2009: 54–6), Kamtekar (2004), Knobe & Leiter (2007), Kupperman (2001), Miller (2003), Montmarquet (2003), Prinz (2009); Russell (2009: 227–331), Samuels & Casebeer (2005), Sarkissian (2010), Solomon (2003, 2005), Snow (2010), Sreenivasan (2002), Swanton (2003: 30–1), Upton (2009), and Webber (2006a,b; 2007a,b).

13. For an extensive bibliography, see: <http://philpapers.org/browse/skepticism-about-character>.

14. The terms “egoism” and “altruism” are used here for descriptive views about the nature of human motivation. Some prefer to call these views “psychological egoism” and “psychological altruism” to distinguish them from “ethical egoism” and “ethical altruism”, which are *prescriptive* views about how people should behave.

15. Interpretation of historical texts is, of course, often less than straightforward. While there are passages in the works of each of these philosophers that can be interpreted as advocating psychological egoism, scholars might debate whether these passages reflect the author’s considered option.

16. For more on the history and philosophical implications of the debate, see Broad (1930), MacIntyre (1967), Nagel (1970), Batson (1991: Chs. 1–3), Sober & Wilson (1998: Ch. 9), and Dixon (2008).

17. For surveys of Batson’s work, see Batson (1991) and Batson (2011).

18. Versions of the standard account have been offered by many authors including Broad (1950), Feinberg (1965 [1999]), Sober and Wilson (1998: Chs. 6 & 7), Rachels (2003: Ch. 6), Joyce (2006: Ch. 1), Kitcher (2010, 2011: Ch. 1), May (2011a), and many others. For discussion of other accounts of altruism, see the entry on empirical approaches to altruism, sections 2, 3 & 4.

19. For some substantive discussion of the question see Cialdini et al. (1997), Sober and Wilson (1998: ch. 7), Stich et al. (2010), May (2011a, 2011b & 2011c) and Slote (2013).

20. For a classic statement of this account of practical reasoning, see Goldman (1970).

21. As with other philosophical isms, the tidy moniker “moral realism”, covers a substantial diversity of opinion. Nevertheless, many moral realists would accept something in the neighborhood of Smith’s characterization of moral objectivity (see Railton 1986a,b; Sayre-McCord 1988a, 2015; Boyd 1988; Sturgeon 1988; Brink 1989; Enoch 2011; Cuneo 2014: Ch. 1).

22. Recently, some moral realists have denied that moral realism is committed to convergence. For example, Shafer-Landau (2003: 228; cf. Moody-Adams 1997: 109; Bloomfield 2001) argues “disagreement poses no threat to realism of any stripe, and so, *a fortiori*, poses no threat to moral realism in particular”. See Doris & Plakias (2008) for argument to the effect that realists cannot afford to be sanguine about disagreement.

23. See Williams (1985: 136): “In a scientific inquiry there should ideally be convergence on an answer, where the best explanation of the convergence involves the idea that the answer represents how things are; in the area of the ethical, at least at high level of generality, there is no such coherent hope”.

24. It is plausible to suppose that convergence does not require total unanimity. However, this plausible qualification raises hard questions: How much dissent can obtain in ideal conditions before “substantial disagreement” is a more apt characterization than “less-than-unanimous-convergence”? As is usual in philosophy, one can’t be very precise about the percentages, but one may suspect that the relevant notion of convergence—always remembering that this is under *ideal* conditions—should be thought to allow only minimal dissent.

25. Brandt is not the only philosopher working in the Anglo-American “analytic” tradition to produce ethnography. Ladd (1957) reports field work with the Navaho; his conclusions (e.g., 1957: 328) about the difficulties posed by disagreement seem somewhat more sanguine than, though perhaps not radically at odds with, Brandt’s.

26. For remarks on situational meaning with affinities to what follows, see Snare (1980: 356–62). Thanks to Walter Sinnott-Armstrong for pointing us to Snare’s valuable discussion.

27. For a more detailed discussion of Moody Adams’ view, see Doris & Plakias (2008).

28. The last clause is important, since realists (e.g., Brink 1989: 200) sometimes argue that apparent moral disagreement may result from cultures applying similar moral values to different economic conditions (e.g., differences in attitudes towards the sick and elderly between poor and rich cultures). But this explanation seems of dubious relevance to the described differences between *contemporary* northerners and southerners, who are plausibly interpreted as applying *different values* to *similar economic conditions*.

29. The legal scholarship that Nisbett & Cohen (1996: 57–78) review makes it clear that southern legislatures are often willing to enact laws reflecting the culture of honor view regarding the circumstances under which violence is justified, which suggests there is at least some support among southerners for the idea that honor values should be universalizable.

30. In addition to the expedients considered here, realists may plausibly appeal to, *inter alia*, requirements for internal coherence and the different “levels” of moral thought (theoretical v. popular, abstract v. concrete, general v. particular) at which moral disagreement may or may not be manifested. Brink (1989: 197–210) and Loeb (1998) offer valuable discussions with considerably more detail than offered here, Brink manifesting realist sympathies and Loeb tending toward anti-realism.

31. Nisbett and Cohen will likely fare pretty well under such scrutiny. See Tetlock’s (1999) favorable review.

Copyright © 2020 by the authors

John Doris, Stephen Stich, Jonathan Phillips, and Lachlan Walmsley