

Moral Realism, Moral Disagreement, and Moral Psychology

Simon Fitzpatrick

Abstract: This paper considers John Doris, Stephen Stich, Alexandra Plakias, and colleagues' recent attempts to utilize empirical studies of cross-cultural variation in moral judgment to support a version of the argument from disagreement against moral realism. Crucially, Doris et al. claim that the moral disagreements highlighted by these studies are not susceptible to the standard 'diffusing' explanations realists have developed in response to earlier versions of the argument. I argue that plausible hypotheses about the cognitive processes underlying ordinary moral judgment and the acquisition of moral norms, when combined with a popular philosophical account of moral inquiry—the method of reflective equilibrium—undercut the anti-realist force of the moral disagreements that Doris et al. describe. I also show that Stich's recent attempt to provide further theoretical support for Doris et al.'s case is unsuccessful.

1. Introduction

To a first approximation, moral realism is the view that there are objective answers to moral questions, that the objectivity of morality is underwritten by a set of stance-independent moral facts,¹ and that these moral facts are discoverable through careful moral inquiry.

Thus understood, moral realism has been thought to face significant difficulties when it comes to explaining the extent and persistence of *moral disagreement* in the world. This paper will consider John Doris, Stephen Stich, Alexandra Plakias, and colleagues' recent attempts to use empirical studies of cross-

¹ Moral facts are stance-independent if they obtain independently of the evaluative attitudes of any actual or hypothetical individual or group (Shafer-Landau, 2003). Moral realism is thus to be contrasted with various constructivist meta-ethical views, which hold that the moral facts obtain relative to a particular evaluative stance, and with moral error theory, which holds that moral statements are systematically false because there no moral facts to make them true, and non-cognitivism, which denies that moral statements are capable of being true or false.

cultural variation in moral judgment to support such an anti-realist argument from disagreement (Doris and Stich, 2005; Machery, et al., 2005; Doris and Plakias, 2008a; Stich, 2009). Crucially, Doris et al. claim that the moral disagreements highlighted by these studies are not susceptible to the standard 'diffusing' explanations realists have developed in response to earlier versions of the argument from disagreement.

I will argue that current views in moral psychology about the cognitive processes that underlie ordinary moral judgment and the acquisition of moral norms, when combined with a popular philosophical account of moral inquiry—the method of reflective equilibrium—suggest a plausible diffusing explanation for the moral disagreements that Doris et al. highlight. While this does not show that the cross-cultural data poses no challenge to moral realism, it does show that Doris et al. have more work to do to make their case. Finally, I will show that Stephen Stich's (2009) recent attempt to provide further theoretical support for Doris et al.'s case is unsuccessful.

2. Moral Realism and Moral Disagreement

What has become known as the 'argument from disagreement' against moral realism has been formulated in many different ways (see Enoch, 2009), but here is a standard formulation:² moral realists hold not only that there are objective moral facts, but also that these moral facts are epistemically accessible to us—i.e. that moral inquiry provides a reliable means of discovering them. However, this would seem to imply a gradual convergence in moral views over time, as individuals and communities engage in such inquiry. Yet, what we actually see are persistent *failures* of convergence. Disagreement over moral issues is not only widespread, but also seems remarkably resistant to resolution. Importantly, so it is argued, this makes morality look unlike areas of inquiry that plausibly admit a realist interpretation, such as natural science. While disagreements over scientific claims are certainly common, empirical scientific methods generally seem able to filter out such

² This formulation derives from Mackie (1977) and subsequent discussion (e.g. Brink, 1989; Loeb, 1998).

disagreements over time. The methods of moral inquiry, on the other hand, appear unable to do this; moral disagreements often appear to be permanent. This, anti-realists claim, suggests either that there are no stance-independent moral facts for us to discover, or, if such facts do exist, the methods of moral inquiry fail to provide reliable means to discover them. Either way, persistent moral disagreement is bad news for realism.

In response to this argument, realists have suggested a number of what Doris et al. call 'diffusing' explanations for the persistence of moral disagreement (e.g. Boyd, 1988; Brink, 1989; Smith, 1994). To begin with, it can be argued that some of the apparent moral diversity in the world—cross-cultural and historical differences in attitudes towards infanticide, for instance—is not really genuine moral disagreement, but the result of the same moral principles being applied to different circumstances—for instance, applying the principle, 'killing is only permitted in extreme circumstances' under different material conditions. Also, as in other areas of inquiry, the truth may be epistemically accessible to us, and yet convergence of opinion fail to obtain, if people do not always form and revise their moral beliefs in rational ways. Some may be irrationally wedded to their beliefs, or fail to take all the relevant issues in account. In addition, people may maintain divergent moral views as a result of insufficient or faulty information about the moral situation at hand (whether X intended to do A, was aware of the consequences of doing A, etc.), or unresolved disagreements over relevant non-moral issues (what cosmological and religious views are correct, whether a foetus is a 'person', etc.). Since more than truth is normally at stake in moral debates, partiality and vested interest may also play a role in explaining why people persistently fail to reach agreement on moral issues. Moreover, realists need not hold that moral inquiry is *guaranteed* to get us to the truth—there is, after all, no such guarantee for scientific inquiry—only that it provides a *generally* reliable means for acquiring and improving knowledge of the moral truths. They can also accept that there are some moral questions that are too difficult for us to answer (presumably some scientific questions may also escape our epistemic reach), or do not admit determinate answers (for instance, because some of the relevant properties, such as the property of being a 'person', have

indeterminate borders). Hence, realism need not entail complete convergence of opinion, even amongst fully informed, rational, and impartial inquirers.

As Doris et al. see it, such responses to the argument from disagreement commit the realist to the claim that *most* moral disagreements are *superficial* rather than *fundamental*. The suggestion seems to be that it is only widespread, truly fundamental moral disagreement—genuine disagreement that would remain unsettled once all epistemic disadvantages on the part of disputants (irrationality, partiality, lack of relevant non-moral information, etc.) have been corrected—that poses a problem for realism. Hence, the realist is committed to the conjecture that genuine moral disagreement will not be widespread under *ideal* epistemic conditions: when people are fully rational, reflective, impartial, and informed of all relevant non-moral facts.³ This appears to shift the focus of the debate from *actual* moral disagreements, which likely do not occur under anything like ideal epistemic conditions, to the question of whether there could exist certain *hypothetical* disagreements, leading some to despair that the meta-ethical debate over moral disagreement has reached its own intractable impasse, reducing to loosely grounded speculation about what would or would not happen under remote hypothetical conditions.⁴

In contrast to this pessimistic view of the debate, Doris et al. emphasize the relevance of empirical data on the nature of actual moral disagreements. In particular, they claim that if we can find a significant number of cases of actual moral disagreement that cannot easily be explained away in terms of the above sorts of diffusing explanations (irrationality, partiality, non-moral disagreements,

³ As Doris and Plakias (2008a, p306-310) note, not all realists accept this commitment (e.g. Shafer-Landau, 2003). Such 'divergentist' realists hold that disagreement in science and other realism-apt domains is more widespread and persistent than moral anti-realists acknowledge, and that this poses no challenge to realism. Thus, such disagreement shouldn't impugn moral realism either. Since Doris et al. direct their empirical case against realists happy to accept commitments to convergence, I will not discuss divergentist positions here, but it is worth considering the extent to which realism does actually require such commitments.

⁴ Realists clearly face difficulties in characterizing these ideal conditions. It would be counter-productive to make them so demanding as to escape real human beings (e.g. requiring some form of omniscience), yet the weaker the conditions the more obvious it will be that widespread disagreement could persist when they obtain. Following Doris et al., I will assume that realists can characterize them without undermining their position, though I will return to the issue of whether the epistemic standards set by realists are too onerous for human beings in Section 7.

and so forth), then that would cast significant doubt on the realist conjecture that most moral disagreements will be settled under ideal epistemic conditions. Hence, even though the debate may turn on hypothetical rather than actual disagreements, both realists and their anti-realist opponents have a clear empirical imperative: to investigate the extent of actual moral diversity in the world and why it is so persistent. Doris et al.'s project is to show that the convergence conjecture is rendered implausible by the empirical evidence.

3. Empirical Evidence Against The Convergence Conjecture?

One of Doris et al.'s key examples is the work of Richard Nisbett and Dov Cohen (1996) on differences in attitudes between White non-Hispanic Northerners and Southerners in the USA towards the use of violence. According to Nisbett and Cohen, many areas of the American South exhibit features of a *culture of honor*, typified by emphasis on male reputation for strength and an imperative for males to respond with violence to slights upon their honor.⁵ In support of this claim, Nisbett and Cohen conducted surveys of Southern and Northern attitudes towards violence. They found that White Southerners were significantly more likely than White Northerners to regard violent retaliation by males to insults and other affronts as 'extremely justified', and more likely to express disapproval of males who failed to so respond with violence. A field survey also found that Southern employers were more likely to take a sympathetic attitude towards a (fictitious) male job applicant who had been convicted of accidentally killing a man in a brawl, after the man had boasted about sleeping with the applicant's fiancé, than a male applicant convicted of stealing a car when his family was in desperate need of money. In contrast, Northern employers showed no difference in attitude towards the manslaughter and theft applicants.

⁵ Nisbett and Cohen argue that such cultures tend to emerge in areas where law enforcement agencies are absent or ineffective at deterring theft of valuable resources. Individuals therefore have to develop their own deterrents, such as acquiring a reputation for extreme retaliation for the slightest affront. Once established, honor cultures may persist long after socio-economic conditions have altered, as in areas of the American South settled by herders who had to protect valuable livestock.

Doris et al. regard this apparent disagreement between White Southerners and White Northerners about the permissibility of violence as important evidence for the claim that a significant amount of moral disagreement would likely persist under ideal epistemic conditions. They claim that it is implausible that either group is simply irrational, or that their judgments are driven by partiality—it seems doubtful, for instance, that ‘Southerners’ economic interests are served by being quick on the draw...’ (Doris and Plakias, 2008a, p320). They also claim that the differences cannot plausibly be explained away by disagreement over non-moral issues. White Northerners and White Southerners don’t appear to disagree about what counts as an insult or provocation, nor do there seem to be relevant differences in background religious or cosmological views. Doris et al. conclude: ‘Nisbett and colleagues’ work represents one potent counterexample to the convergence conjecture...’ (Doris and Plakias, 2008a, p321).

In addition to the Nisbett and Cohen data, Doris et al. also report data of their own (Peng et al., unpublished data) highlighting cross-cultural differences in responses to the famous *magistrate and the mob* thought experiment. Chinese undergraduates and American undergraduates of European descent were presented with a vignette in which a murder has taken place in a town with a history of ethnic violence. Participants were told that if the town’s authorities do not immediately punish someone for the crime the townspeople will riot and many people will be killed. The authorities do not know the identity of the murderer. However, they can frame an innocent man for the crime, thus preventing the rioting. Participants were then told that the authorities decided to frame the innocent man, and asked whether this was the right thing to do. The American undergraduates tended to say that the authorities did not do the right thing. However, the Chinese undergraduates tended to respond that framing the man was the right thing to do, and that it was actually the townspeople who were responsible for his being punished.

This disagreement between American and Chinese undergraduates is claimed as another case that casts doubt on the convergence conjecture. Importantly, participants were also asked various factual questions about the case, including whether the innocent man would suffer as result of being punished, and

whether the riots would have caused pain and suffering. No significant differences were found in the answers given by the two groups. According to Doris and Plakias (2008a, p324), 'the differences are not readily to be attributed to differences in conceptions of the *non-moral facts*.' Nor, they claim, is it plausible to regard either group of participants as irrational, or motivated by partiality; the hypothetical nature of the scenario suggests that both groups can be regarded as impartial observers.

Doris et al. do not claim to have ruled out *all* possible 'diffusing' explanations for these apparent moral disagreements. However, their contention is that these studies, along with a number of other studies of cross-cultural evaluative diversity, 'motivate confident speculation to the effect that the [convergence] conjecture is unlikely to be satisfied' (Doris and Plakias, 2008a, p327).⁶

Doris et al.'s claims have prompted various critical responses (e.g. Bloomfield, 2008; Leiter, 2008; Sneddon, 2009; Fraser and Hauser, 2010). In particular, Leiter (2008) and Fraser and Hauser (2010) question whether Nisbett and Cohen's work provides clear evidence of disagreement between White Northerners and White Southerners about the *permissibility* of violence, rather than just a difference in their willingness to *forgive* violent acts, or the *degree* to which they regard violence as permissible in a given situation.⁷ Moreover, while Doris et al. emphasize the need for philosophers to obtain more accurate understanding of the extent of moral diversity, they only discuss research on disagreement. Research seemingly pointing in the other direction also has to be considered—for instance, Mikhail's (2011) work documenting striking cross-cultural similarities in judgments and norms concerning battery. Indeed, to some extent, the Peng et al. data demonstrates core agreement between Chinese and Americans that individual

⁶ Increasing globalization might reduce such diversity via the hegemony of one culture over others, but this is not the kind of 'convergence' Doris et al. have in mind, and clearly couldn't be taken to indicate convergence to the *truth*.

⁷ Leiter and Fraser and Hauser are nonetheless supporters of Doris et al.'s cause. Leiter argues that the history of moral disagreement in philosophy provides the clearest challenge to realism, since professional philosophers come closest to approximating ideal epistemic conditions. Fraser and Hauser regard some of their own cross-cultural work as providing firmer evidence for Doris et al.'s conclusions. My response to Doris et al. applies equally to Fraser and Hauser. Doris and Plakias (2008b) respond to Leiter on the relevance of disagreement in philosophy.

liberty and community harmony are moral goods, even if they disagree about how they ought to be traded off—re-call that the Chinese participants tended to blame the townspeople for punishing the innocent man, suggesting moral concern about the infringement of his liberty (see also Wong, 2006 on Chinese versus Western moral traditions). Hence, we need a broader evaluation of the true extent of moral diversity, being careful to distinguish genuine disagreement on core moral principles and values from different views on how the same principles/values ought to be applied in particular cases—persistent disagreement on the latter arguably being less troubling for realism.

Rather than pursue these sorts of issues, I will argue that, in any case, plausible hypotheses about the psychological processes underlying moral judgment and the acquisition of moral norms, when combined with an account of moral inquiry endorsed by many contemporary realists, suggest that the disagreements cited by Doris et al. need not challenge the convergence conjecture. In so doing, I will highlight some important considerations that need to be taken into account when evaluating how actual agreements or disagreements bear on the plausibility of moral realism.

4. A Dual-Process Model of Moral Judgment and Norm Acquisition

The model of human moral psychology that I want to appeal to is, ironically, one that has been developed by Stich, along with Chandra Sripada (Sripada and Stich, 2006).⁸ The Sripada and Stich (S&S) model synthesizes a large body of empirical research that has been conducted over the last few decades into the nature of moral judgment and normative cognition more generally. In particular, it is consistent with the increasingly popular view that most of our moral judgments are not the product of, or much influenced by, deliberative and consciously guided moral reflection. Rather, they are, in the psychological sense, moral *intuitions*: fast, automatic, and

⁸ Stich (2009) argues that this model can be used to provide further theoretical support for Doris et al.'s case. I will discuss these claims in Section 6. In Section 5, I will argue that it actually highlights a plausible diffusing explanation for moral disagreements.

(often) emotion-laden responses that are the product of entirely unconscious cognitive processes.

One important piece of evidence for this view comes from work on the phenomenon of *moral dumbfounding* (Haidt, 2001), which shows that people often react with immediate, visceral moral disapproval to certain actions (e.g. consensual sibling incest), while being completely unable to articulate moral principles or beliefs that can explain why they judge the action to be wrong. In addition, brain-imaging studies have shown activity in emotional and affective areas of the brain to be more commonly implicated in moral judgment than activity in areas associated with controlled reasoning (Paxton and Greene, 2010). Affect manipulations, such as engendering heightened disgust, have also been shown to influence moral judgment (Wheatley and Haidt, 2005).

Such results provide support for the view that most moral judgments are automatic, intuitive responses to situations, closely linked with emotional and affective responses, rather than the product of controlled moral reasoning and reflection on the part of the agent.⁹ They do not, however, show that controlled reasoning *never* plays a role in moral cognition. The history of ethical thought suggests that people are able to explicitly formulate and reason from a set of consciously accessible moral beliefs. This is confirmed by brain-imaging studies, showing activation in brain areas associated with controlled reasoning in some instances of moral judgment, and behavioral studies showing that people can consciously alter their initial moral judgments to bring them in line with explicitly held moral principles, and that reflective thought can be increased by experimental manipulations (Pizzaro et al., 2003; Paxton and Greene, 2010). Nonetheless, the current empirical literature does suggest that such consciously guided, reflective thinking is a comparatively rare component of our moral lives.

⁹ This psychological sense of 'intuition' does not involve the claim, associated with the intuitionist tradition in moral philosophy, that moral intuitions are direct apprehensions of moral truths, or that they have privileged epistemic status. Nor does it imply that no process of reasoning whatsoever is involved in their generation. Complex computational processes may still be operating below the level of consciousness (Mikhail, 2011).

This has motivated numerous researchers to postulate *dual-process* models of moral cognition (e.g. Saunders, 2009; Cushman et al., 2010; Kahane, 2012). Dual-process models of cognition (e.g. Evans and Over, 1996) distinguish between two types of cognitive process that may play role in a given cognitive domain: System 1 processes, which are automatic, fast, and not consciously accessible, and System 2 processes, which are effortful, slow, and consciously accessible. The suggestion is that a suite of System 1 processes, closely linked with emotion and affect, drive much of our moral cognition, and these are responsible for our moral intuitions. We also have a System 2 capacity for controlled and consciously accessible moral reasoning, which underlies our ability to explicitly formulate, and make moral evaluations based upon, a set of moral convictions. This may allow us to override our System 1 intuitions in some cases, but plays a minimal role in everyday moral judgment, compared to System 1 processes.¹⁰

4.1 *The S&S Model*

Sripada and Stich's 'A framework for the psychology of norms' (2006) is an attempt to explicate what such a dual-process model of moral cognition might look like. The model is represented diagrammatically in Figure 1. One important question raised by the dual-process picture concerns where our System 1 moral judgments come from, if not from some process of controlled reasoning and reflection. Merely highlighting the link with emotional and affective responses does not explain how such responses are triggered by often quite complex moral stimuli (Mikhail, 2011). Sripada and Stich ask a similar question about normative judgments more generally. They describe empirical work showing that though there are significant differences in their content, every human social group has norms governing aspects of life such as appropriate behavior, dress, and diet. Violations of these norms typically elicit strong emotional reactions, such as anger and disgust, and these provide individuals

¹⁰ Greene and colleagues (e.g. Paxton and Greene, 2010) argue that the difference between intuitive and reflective processes is associated with a difference in the *content* of the resultant moral judgments: the former reflecting implicit deontological commitments, the latter reflecting utilitarian reasoning. Kahane (2012) critiques these claims, arguing for a content-independent dual-process model similar to that described here.

with strong intrinsic (rather than instrumental) desires for norm compliance and for the punishment of norm violators, even at significant cost to themselves. Yet, many of these norms are not explicitly articulated or taught. Rather, they are implicit in the behavior and normative judgments of other group members. How is it, then, that children come to internalize the norms of their group?

[Insert Figure 1 here]

To address this question, Sripada and Stich posit two key psychological mechanisms, both of which are culturally universal, innate, automatic, and entirely unconscious—hence, they can both be regarded as System 1 processes. These are the *acquisition mechanism* and the *execution mechanism*. The acquisition mechanism is a dedicated learning mechanism that enables individuals to unconsciously identify and internalize the prevailing moral and non-moral norms of their group from proximal cues in the environment. As Sripada and Stich note, there is a range of possible hypotheses about how exactly the norm acquisition mechanism does this. According to what they call the ‘Pac-Man’ thesis, although the mechanism is specifically tuned to identifying norm-implicating behavior in the environment, it internalizes all and only those norms present in the environment. Another possibility is that the mechanism is guided by an innate normative ‘grammar’ that constrains the types of norms it can acquire.

In any case, once the content of a prevailing norm is identified by the acquisition mechanism, a mental representation of the norm gets stored in the *norm database*. This is a central part of the execution mechanism, which underlies everyday capacity for normative evaluation. This mechanism applies the internalized norms in the database to particular situations to determine what compliance with these norms would involve, and whether a particular action on the part of the individual, or a third party, would conform to, or violate, an internalized norm. The model links the norm database with motivational and desire generating processes that lead to the individual being intrinsically motivated to comply with the norms in the norm database, desire that others comply with them, and being

motivated to behave punitively towards those who violate these norms. The model also has the execution mechanism output evaluative judgments about particular situations into the *judgment* box via the *emotion system*, giving normative judgments their characteristic emotional valence.

However, when an individual observes, or thinks about, a norm violation, all she has conscious access to is the normative judgment and the emotional valence attached to it. Though the execution mechanism may have to perform some complex processing in order to apply the relevant norm to particular situations, the individual is not consciously aware of going through any process of reasoning to reach the judgment. Rather, such judgments are fast, automatic responses to the relevant stimuli, the content of which are determined by the norms stored in the norm database, which are in turn the product of the acquisition mechanism internalizing the norms present in the individual's cultural environment. The individual has no conscious control over which norms are internalized, nor does she have conscious access to these norms. Hence, she will not be able to articulate the norms that give rise to her judgments; rather, she will only be able to provide a post hoc reconstruction of them by studying her intuitions. This accounts for the phenomenon of moral dumbfounding.

In contrast to the acquisition and execution mechanisms, the *belief*, *judgment*, and *explicit reasoning* boxes can be regarded as System 2 processes. It is via the operation of these latter processes that reflective moral reasoning takes place. An individual's explicitly held moral principles and convictions may be acquired via explicit reasoning, testimony, or explicit learning. Once acquired, they are stored, not in the norm database, but in the belief box. Application of such principles and convictions to concrete situations will take place not via the execution mechanism, but via the operation of explicit reasoning.

4.2 *The S&S Model and Reflective Equilibrium*

Many important architectural questions are left open by the S&S model. However, this dual-process framework for understanding our normative psychology is well supported by the current empirical evidence.

One interesting implication of the S&S model that has been explored by Leland Saunders (2009) is that potential interactions between the System 1 and System 2 processes may provide a psychological basis for a popular philosophical account of moral inquiry and justification: the method of *reflective equilibrium*, suggested by John Rawls in *A Theory of Justice*, but perhaps most influentially articulated by Norman Daniels (1979). On this account, moral inquiry should be conceived as a dynamic, ongoing process of mutual adjustment between one's considered moral judgments, moral theories, and non-moral beliefs, the ideal endpoint being a fully coherent network of belief—a reflective equilibrium.

The method is supposed to work like this: as a moral theorist, one has no choice but to start with one's initial moral beliefs—for example, the intuitive judgments that one has about the rightness or wrongness of particular actions. The first stage of the method is to sort among these to find those *considered* judgments that one has some initial confidence in—for instance, those that are stable rather than fleeting, are not the product of distress, anger, or vested interest. The second stage is to formulate moral principles that can systematize and explain these judgments. Such systematization is regarded as mutually reinforcing, providing justification for both the considered judgments and the systematizing principles. Achieving a coherent system of moral beliefs will not be easy, however. Any candidate set of moral principles will typically come into conflict with some of one's considered judgments. One will then have to revise either or both the discrepant judgments and the moral principles, depending on which one has least commitment to, and on what kind of revision would bring about the most coherent system. Neither is to be held immune from potential revision in the search for equilibrium. However, one should not just settle for a *narrow* reflective equilibrium—an internally coherent moral system. One should also strive towards a *wide* reflective equilibrium, which involves bringing one's moral beliefs into a state of systematic coherence with one's other beliefs, including, for instance, background scientific and philosophical beliefs. This should also involve comparing candidate moral systems with potential alternatives, considering possible arguments for and against each of them, so as to assess which would fit most coherently into one's overall network of

beliefs. Though one may never actually reach the ideal endpoint of a full wide equilibrium, the degree of overall coherence attained constitutes a measure of the rational justification for one's beliefs.¹¹

According to Saunders' psychologized account of reflective equilibrium, System 1 processes can supply the individual with a set of starting intuitive moral judgments. System 2 reasoning processes may then use these System 1 judgments to get the process of moral inquiry off the ground, attempting to develop systematizing principles, which would get stored in the belief box. Saunders then envisions discrepant System 1 judgments potentially being overridden by System 2 judgments derived from the individual's consciously held moral and non-moral beliefs. More speculatively, Saunders suggests that an individual's reflectively held moral beliefs could become embedded in the norm database. For instance, the norm acquisition mechanism might take an individual's verbal expression of her moral beliefs as inputs, causing them to be internalized into the norm database. Hence, explicit moral reasoning might lead to an individual acquiring new moral intuitions that conform to her explicitly held moral convictions.

These possibilities are important for moral realism, since many realists have endorsed a reflective equilibrium account of how it is that we are to acquire knowledge of the moral facts (e.g. Boyd, 1988; Brink, 1989; Smith, 1994).¹² The method is attractive to realists since it allows them to avoid having to posit a mysterious faculty of rational intuition that provides us with direct non-inferential access to the moral facts. They can concede that moral truths cannot be directly inferred from empirical observation without having to embrace the idea that they

¹¹ Thus understood, reflective equilibrium is normally taken to assume a coherentist account of epistemic justification, according to which a belief is justified in proportion to the coherence of the network of beliefs of which it is a part, and thus implies a rejection of foundationalism, according to which justification must ultimately trace back to set of non-inferentially justified foundational beliefs. However, the search for coherence as *method* of inquiry can, in principle, be separated from the idea that justification is reducible to coherence, and not all advocates of reflective equilibrium reject foundationalism.

¹² Rawls famously shied away from a realist interpretation of reflective equilibrium, preferring to see it as a kind of psychology, helping to reveal our fundamental 'moral sensibility'. In this respect, there are important differences between Rawls and Daniels' treatments of reflective equilibrium (see Mikhail, 2010). The description of the method given here follows Daniels' treatment, which has been most influential in realist circles.

are somehow self-evident, or that our moral intuitions have some privileged epistemic status.¹³ Similarly, it suggests that our pre-reflective moral beliefs can still play a role in moral inquiry, even if it is recognized that cultural and biological forces have shaped their content. These merely provide the starting point for moral inquiry. The pursuit of wide reflective equilibrium may lead to significant revisions of these beliefs, or to their systematization. Crucially, that our moral beliefs may have non-rational origins need not undermine their potential justification on a coherence account of moral justification. Reflective equilibrium therefore provides a response to the notion that cultural or biological explanations for the content of our moral beliefs are necessarily debunking. Indeed, according to Richard Boyd, reflective equilibrium undermines the notion that moral and scientific inquiry are all that distinct, since 'the dialectical interplay between observations, theory and methodology which, according to the realist, constitutes the *discovery* procedure for scientific inquiry *just is* the method of reflective equilibrium...' (Boyd, 1988, p199-200).

With this conception of moral inquiry in mind, I want to argue that the S&S model highlights a plausible diffusing explanation for the moral disagreements that Doris et al. cite.

5. A Diffusing Explanation

To set the stage for this response, it should be emphasized that realists are only committed to claims about convergence given the assumption that some appropriate method of moral inquiry provides a generally reliable procedure for acquiring and improving knowledge of the moral truths. As Doris and Plakias put it, the standard way of framing the problem of disagreement is 'that application of the same method may, for different individuals and cultures, yield divergent moral judgments that are equally acceptable by the lights of the method, even in reflective

¹³ This is not to say that all realists have abandoned such notions (e.g. Shafer-Landau, 2003; Audi, 2004). Though some modern intuitionists *may* be able to utilize the kind of diffusing explanation I will soon articulate, intuitionism is, I think, threatened both by the sort of cross-cultural data that Doris et al. cite and the psychological account of our moral intuitions just described (on the latter, see Sinnott-Armstrong, 2006). Thus, I will restrict my focus to non-intuitionist realisms.

conditions that the method countenances as ideal' (2008a, p326). Crucially, this means that in assessing whether actual moral disagreements ground 'confident speculation' that the convergence conjecture is unlikely to be satisfied, it matters how exactly realists think this method is supposed to work, what the relevant 'ideal' reflective conditions are, and what sorts of cognitive processes are actually involved in sustaining the disagreements. Actual moral disagreements would only seem to put genuine pressure on the convergence conjecture if they are sustained by, and thus persist in spite of, individuals (or communities of individuals) diligently forming and revising their moral beliefs according to whatever method of inquiry realists think provides us with epistemic access to the moral facts, since it is the ability of this method, properly deployed, to settle moral disagreements that is at issue. If a disagreement is not sustained by the relevant set of cognitive processes that constitute the implementation of this method, or is somehow causally insulated from them, this will leave open the possibility that the disagreement may be resolved were the method to be deployed. Thus, such disagreements need not present the kind of threat to the convergence conjecture that Doris et al. are trying to pose. Indeed, in principle, this conjecture is consistent with any amount of persistent diversity in moral belief where those beliefs are insulated from the relevant method of inquiry.

Let us assume, then, that the S&S model is broadly correct. Let us also assume that something like the method of reflective equilibrium is the preferred realist conception of how we are to discover the moral facts (which it does seem to be). As Saunders has shown, the S&S model seems to allow for the psychological possibility of pursuing the method of reflective equilibrium.¹⁴ However, the empirical work underlying the model also suggests that most ordinary people do not engage in this process, or do so only very rarely. Saunders argues that this is partly because it is time-consuming and cognitively costly. Most ordinary people probably have little motivation to reflect upon, systematize, and potentially revise

¹⁴ There is also fascinating research by Epley and colleagues (Epley et al., 2004; Epley and Gilovich, 2005) on how people adjust their initial judgments in various cognitive domains, which shows some of the hallmarks of reflective equilibrium.

the moral judgments that they have internalized from their community. Indeed, as Sripada and Stich note, in many societies there are strong social pressures against deviation from prevailing normative standards.

Consequently, the S&S model suggests that most moral judgments made by ordinary people, including those made by participants in the studies cited by Doris et al., are brute System 1 intuitions: the product of the acquisition mechanism automatically and unconsciously internalizing the socially-generated norms present in the environment. While such System 1 judgments may count as considered judgments, in so far as they are stable rather than fleeting, not simply the product of stress, partiality, and so forth, such judgments are only the *starting point* for reflective equilibrium—the input, rather than the output.

My suggestion, then, is that realists can argue that, in so far as it is plausible to hold that the moral disagreements at issue between White Northerners and Southerners and between Chinese and American college students exist at the level of System 1 intuitions that have not been subject to careful reflection and potential revision via the operation of System 2 processes, they should not, without further argument, be regarded as having direct empirical bearing on the convergence conjecture. This is because such disagreements are not sustained by individuals reaching different moral conclusions via the method supposed by realists to have to the power to eliminate moral disagreements under ideal conditions. Rather, they are sustained by individuals unconsciously internalizing different norms from different cultural environments.

Importantly, the nature of System 1 moral cognition allows us to offer a diffusing explanation for why these moral disagreements may persist even when the disputants don't appear to disagree in any directly relevant way on the non-moral facts, and there is no clear evidence of outright irrationality or partiality. The norms that give rise to the moral judgments of White Southerners and Northerners and Chinese and American college students are culturally entrenched as a result of the fact that we, as humans, automatically and unconsciously internalize the prevailing norms of our social group, rarely, if ever, reflect on the judgments these internalized norms give rise to, and thus rarely bring them into contact with the rest of our

beliefs. This leaves open the possibility that such disagreements may be resolved were the parties to the disagreement to consciously reflect on their brute moral intuitions, bring them into contact with their non-moral beliefs, and engage in the pursuit of wide reflective equilibrium.

To be clear, I am *not* claiming that we have good reason to think that the moral disagreements at issue *will* be eliminated under ideal conditions, or that they are *irrelevant* to the meta-ethical debate. Rather, I am claiming that the cross-cultural data that Doris et al. cite do not, as presented, and without further argument, support an empirical case against the realist convergence conjecture. This is because the moral judgments at issue are plausibly not ones that have withstood application of the kind of method of moral inquiry that realists claim will lead to the elimination of (most) moral disagreement under ideal conditions. Hence, they plausibly do not put the convergence conjecture to the kind of empirical test that Doris et al. claim.

It might be argued that this makes things too easy for realists. They cannot arbitrarily stipulate a particular method of moral inquiry, and then claim that only moral disagreements sustained by that method provide a challenge to realism, since they could then guarantee the absence of problematic disagreement by playing around with the notion of moral inquiry. For example, realists could stipulate as constitutive of moral inquiry a version of reflective equilibrium that includes a set of inviolable moral beliefs, in order to rule out any potential disagreement over these beliefs as being a threat to realism. That would indeed be an ad hoc move. However, realists shouldn't be accused of making such a move in this context. Reflective equilibrium is not being plucked out of the air just to explain away moral disagreement. It is a central commitment of many prominent realists. Indeed, it is arguably the orthodox view of inquiry in contemporary philosophy.

In any case, the version of the argument from disagreement that Doris et al. advocate only gets off the ground given the assumption that the disagreements in question persist in spite of diligent application of a particular method of forming and revising moral beliefs that is supposed by the realist to provide us with epistemic access to the moral facts. As noted earlier, Doris and Plakias seem to

acknowledge this when they state that their way of setting up the meta-ethical relevance of the cross-cultural data assumes that 'moral disagreements proceed in surroundings typified by substantial *methodological agreement*. On this construction, the problem of disagreement is that application of the same method may, for different individuals and cultures, yield divergent moral judgments that are equally acceptable by the lights of the method, even in reflective conditions that the method countenances as ideal' (2008a, p326). However, they go on to claim that not only is there significant diversity of moral belief, there is also likely to be significant disagreement about the correct method for forming and revising these beliefs, and claim that this makes the problem of disagreement worse for realism: 'the methodological restriction is in want of an argument—an argument that does not beg substantive methodological questions' (ibid.). But that is to completely misunderstand why moral disagreement is a problem for realism. No realist should be committed to the claim that there will be convergence of moral belief amongst individuals that use *any* method whatever for forming and revising their beliefs. Commitments to convergence only follow from claims about epistemic access, and no realist should claim that any method whatever provides epistemic access to the moral facts.¹⁵

Doris et al. may be right that there is cultural diversity in conceptions of correct moral methodology, but that just serves to emphasize that the adoption of particular methods rather than others requires justification. There are deeply difficult issues to address here for any philosophical theory that endorses particular methods of inquiry. Standard strategies for justifying methods such as reflective equilibrium may not work in the face of this kind of cultural diversity (Stich, 1990). For example, one might not be able to offer it as an analysis of *the* concept of moral inquiry or moral justification. But that doesn't change the fact that the epistemic problem of moral disagreement cannot even be posed unless particular norms of

¹⁵ Consider the analogous situation for scientific realism: in so far as realists are committed to there being a general convergence of opinion on scientific truths, it is only for agents who form and revise their beliefs according to what are regarded as reliable scientific methods. Persistent failures of convergence between those who utilize such methods and those who do not (e.g. astrologers and creationists) presents no problem, since realists are not committed to the epistemic reliability of other possible methods.

moral inquiry endorsed by realists are accepted for the sake of argument.¹⁶ Hence, given that it is the ability of diligent moral inquiry, as realists understand it, to produce moral convergence that is at issue in the argument from disagreement, the point remains that, if the S&S model is on the right track, and reflective equilibrium is the conception of moral inquiry adopted by realists, Doris et al.'s data sheds little empirical light on this. At the very least, Doris et al. have more work to do to challenge the convergence conjecture.

6. Reflective Equilibrium and The Prospects for Convergence

One might be skeptical that even impeccable application of the method of reflective equilibrium would likely resolve the sorts of disagreements that seem to exist between White Northerners and Southerners, and between Chinese and American college students. If such skepticism were legitimate, these disagreements could still threaten the convergence conjecture. Though he does not consider the sort of response to Doris et al.'s argument described in the previous section, Stich (2009) has effectively sought to highlight grounds for such skepticism in an effort to provide additional theoretical support for Doris et al.'s project.

Stich agrees with Saunders that the S&S model motivates a psychologized reflective equilibrium account of moral inquiry. He then argues that this gives us good reason to *expect* there to be widespread fundamental moral disagreement, simply because inquirers born into different cultures will tend to internalize different socially generated norms, and start out with different considered judgments. Stich relies on a view of reflective equilibrium, common amongst critics of the method, according to which the set of moral beliefs that an individual is likely to end up with, even in wide reflective equilibrium, will largely be determined by those she started with. While some revision of these pre-reflective beliefs may occur in the initial filtering process, in the process of ironing out inconsistencies that she may come to find in her beliefs, and upon consideration of alternative moral views,

¹⁶ Interestingly, Shafer-Landau (2003, p224-226) responds to the argument from disagreement by accepting methodological pluralism: if there is no single correct method for determining moral truths, no argument from disagreement can be posed.

the end state is largely going to be a function of the starting state.¹⁷ Hence, all the method is likely to do is to polish up the individual's cultural inheritance—the norms embedded in her norm-database. If this were so, then it would seem obvious that there could remain significant moral disagreement even amongst ideal inquirers, who agree on all the non-moral facts, but started from different sets of judgments.¹⁸

Realists may pursue various responses to Stich. First, he assumes that it will be easy for people in different cultures to construct equally coherent but radically incompatible networks of moral and non-moral belief. However, much turns on how we are to understand the relevant notion of 'coherence'. Critics of coherence methods in ethics typically take it to involve merely removing logical inconsistency in one's beliefs. Yet, proponents of such methods normally have a much more demanding notion in mind, which also involves maximizing things like *evidential* consistency (that none of one's beliefs constitute evidence against holding any of one's other beliefs), *connectedness* (there being mutually reinforcing inferential and explanatory connections between beliefs), *comprehensiveness*, and so on (Sayre-McCord, 1996). Realists may argue that, on such a demanding understanding of coherence, it will be no easy feat to turn a folk morality into a coherent moral system and blend it with a coherent network of non-moral belief, without subjecting it to extensive revision—revision that will filter out most of the starting disagreements that may or may not exist between individuals from different cultures.

Second, Stich also assumes that our non-moral beliefs place weak constraints on the kinds of moral views that can be in wide reflective equilibrium. Here, realists

¹⁷ This motivates the criticism that reflective equilibrium is merely intuitionism rebranded: any justification that an individual can have for the beliefs she holds in reflective equilibrium will have to derive from her original intuitions.

¹⁸ Since realists needn't hold that moral inquiry is guaranteed to get us to the truth (even in the long run), a realist conception of reflective equilibrium is consistent with individuals reaching different moral views and being equally justified in holding their respective networks of beliefs in virtue of their overall coherence. It is also consistent with the possibility of multiple full wide reflective equilibria, even for ideal agents who agree on all the non-moral facts. What would seem to be problematic for realism is if there could be *widespread* moral disagreement between such ideal agents in full wide reflective equilibrium. Stich seems to assume that this is highly likely, given cultural differences in starting points.

may want to follow Daniels (1979): striving for a genuinely *wide* reflective equilibrium is likely to reveal complex and surprising interconnections between our moral and non-moral beliefs, and thus highlight unobvious ways to make starting moral disagreements more tractable. For example, Daniels (following Parfit) suggests that part of the dispute between utilitarian and Rawlsian theories of justice may come down to their fitting better with different theories of the person: 'The problem between the utilitarian and the contractarian thus becomes the (possibly) more manageable problem of determining the acceptability of competing theories of the person, and only one of many constraints on that task is the connection of the theory of the person to the resulting moral principles' (1979, p263). For many naturalistic realists, who hold that moral properties supervene on, or are reducible to, ordinary natural facts about what promotes basic human and societal needs (e.g. Boyd, 1988), the assumption is also that the expansion of background scientific knowledge in the human sciences will substantially constrain the kinds of moral beliefs that can be in genuine wide reflective equilibrium. Hence, realists may be inclined to bet that the non-moral constraints on our moral beliefs will leave little scope for substantial moral disagreement between fully rational, impartial, and reflective inquirers, informed of all the relevant non-moral facts.

Finally, some advocates argue that reflective equilibrium should not be seen as a detached and purely intellectual process. DePaul (1988) argues that it should involve a requirement to seek out moral *experiences* apt to cause one to change one's mind. DePaul is concerned with 'moral naïveté', where people maintain moral views that they would otherwise reject as a result of limited life experiences. Thus, he emphasizes the importance of being exposed to a range of challenging moral experiences—including, for instance, acquaintance with vivid representations of morally significant events in art and literature, as well as personal interaction with people who hold different moral perspectives. If such a role for moral experience is built into the method, and it is recognized that such experiences can cause radical shifts in one's moral thinking, there may be further hope for convergence.

Given these potential responses to Stich, which is the more plausible speculation: optimism or pessimism about the prospects for convergence under

ideal conditions? Unfortunately, I don't think that we are on solid enough ground to arrive at a principled verdict. This is partly due to the remoteness of the hypothetical situations we are trying to imagine, but it is also due to the residual vagueness of the method of reflective equilibrium itself. Though there have been attempts by proponents of the method to articulate norms for filtering initial judgments, and for making revisions to one's beliefs in order to bring them into coherence, these norms tend to leave important questions about the method unanswered. For instance, can one reject a non-moral belief if doing so would bring one's moral beliefs into coherence with the rest of one's belief network, or do non-moral beliefs always trump moral ones? How is one to choose between the likely numerous different possible ways of bringing any particular set of moral and non-moral beliefs into coherence? The remarks made in the previous section about how an argument from disagreement needs to proceed suggest that, without a much more detailed account of how reflective equilibrium is actually meant to work, we cannot reasonably assess whether there could exist widespread moral disagreement between ideal epistemic agents in full wide reflective equilibrium, given particular culturally variable starting points. This places the onus on realists to be much more specific about the norms that they think should govern the pursuit of reflective equilibrium. But it also shows that, while realists should not rest easy, Stich has failed to add any substantive weight to Doris et al.'s case. Indeed, we seem to be back to the kind of impasse that Doris et al. originally wanted to break.

7. Concluding Remarks

It should be noted that the picture of human moral psychology I have been relying on is a double-edged sword for the realist. First, it suggests that apparent instances of moral agreement may not be the output of a process of moral inquiry either, but rather the product of joint cultural inheritance, or perhaps the constraints imposed by an innate normative grammar. Thus, contrary to realists (e.g. Smith, 1994) that have sought to use arguments from agreement to establish the reliability of moral inquiry, convergence on the truth may not provide the best explanation for the cases of moral agreement that we do find. As with arguments from disagreement, what

matters in assessing such arguments are the cognitive processes involved in producing the agreement.¹⁹

Second, while the apparent rarity of reflective moral inquiry can be seen as a diffusing explanation for the persistence of moral disagreement, it seems that when we do engage in such reflection we are not especially good at it. Haidt's (2001) work on moral dumbfounding suggests that when people are forced to engage in controlled moral reasoning, all they tend to do is search for supporting evidence for their pre-existing judgments. When they reach the point of being dumbfounded, they simply stop reasoning altogether, but do not revise their judgments. Other work has shown that moral judgments are susceptible to a variety of framing effects and biases (see Sinnott-Armstrong, 2006), and, more generally, that people tend to be poor at noticing and correcting inconsistency and incoherence in their own beliefs and arguments, and recognizing the merits of views with which they disagree (see Mercier and Sperber, 2011). There is little reason to think that professional philosophers or other 'moral experts' are much less susceptible to these kinds of biases. Though there is evidence that people are able to overcome, or at least mitigate the effects of these biases, this should give realists pause in so far as it raises concerns about the practical feasibility of reflective equilibrium as a method for discovering the moral facts. It could be argued that we suffer from too many cognitive shortcomings to ever be in a legitimate position to adopt a realist attitude towards any particular set of moral beliefs, if the epistemic standards embodied in reflective equilibrium are to be our guide to whether we are entitled to hold that moral truth has been achieved.

These are important issues for realists to face up to. Indeed, I suspect that the latter concern has the makings of a much more potent epistemic challenge to realism than the argument from disagreement that Doris et al. advocate—in particular, it does not rely on speculations about what would happen under hypothetical ideal conditions. Nonetheless, what I have shown in this paper is that if something like the S&S model of our moral psychology is correct, and reflective

¹⁹ See Nichols (2004) for an interesting non-realist account of the convergent evolution of common harm norms.

equilibrium is the preferred realist account of moral inquiry, the actual moral disagreements cited by Doris et al. need not provide a significant challenge to the realist conjecture that most moral disagreement will be eliminated under ideal epistemic conditions. I have also argued that Stich's recent attempt to use the S&S model to provide further theoretical support for the argument from disagreement is unsuccessful. Moral realists cannot afford to be sanguine about moral disagreement, but in this instance at least, empirical critics of realism have more work to do to make their case.²⁰

John Carroll University
sfitzpatrick@jcu.edu

References

- Audi, R. 2004: *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton: Princeton University Press.
- Boyd, R. N. 1988: How to be a moral realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Ithaca, NY: Cornell University Press.
- Bloomfield, P. 2008: Disagreement about disagreement. In W. Sinnott-Armstrong (ed.), *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press.
- Brink, D. O. 1989: *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Cushman, F. A., Young, L., and Greene, J. 2010: Our multi-system moral psychology. In J. Doris et al. (eds.), *The Oxford Handbook of Moral Psychology*. New York: Oxford University Press.
- Daniels, N. 1979: Wide reflective equilibrium and theory acceptance in ethics. *Journal of Philosophy*, 76, 256-282.
- DePaul, M. 1988: Naïveté and corruption in moral inquiry. *Philosophy and*

²⁰ My thinking on these issues evolved from discussions with Steve Stich during workshops of the *Culture and the Mind Project*, funded by the UK Arts and Humanities Research Council. Many thanks to two referees, Harry Gensler, Steve Stich, and an audience at John Carroll University for helpful criticisms and suggestions.

Phenomenological Research, 48, 619-635.

Doris, J. and Stich, S. 2005: As a matter of fact: empirical perspectives on ethics. In F. Jackson and M. Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*. Oxford: Oxford University Press.

Doris, J. and Plakias, A. 2008a: How to argue about disagreement: evaluative diversity and moral realism. In W. Sinnott-Armstrong (ed.), *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press.

Doris, J. and Plakias, A. 2008b: How to find a disagreement: philosophical diversity and moral realism. In W. Sinnott-Armstrong (ed.), *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press.

Enoch, D. 2009: How is moral disagreement a problem for realism? *Journal of Ethics*, 13, 15-50.

Epley, N., Van Boven, L., Kesar, B., and Gilovich, T. 2004: Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 87, 327-339.

Epley, N. and Gilovich, T. 2005: When effortful thinking influences judgmental anchoring. *Journal of Behavioral Decision Making*, 18, 199-212.

Evans, J. and Over, D. 1996: *Rationality and Reasoning*. Hove: Psychology Press.

Fraser, B. and Hauser, M. 2010: The argument from disagreement and the role of cross-cultural empirical data. *Mind & Language*, 25, 541-560.

Haidt, J. 2001: The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychology Review*, 108, 814-834.

Kahane, G. 2012: On the wrong track: process and content in moral psychology. *Mind and Language*, 27, 519-545.

Leiter, B. 2008: Against convergent realism: the respective roles of philosophical argument and empirical evidence. In W. Sinnott-Armstrong (ed.), *Moral Psychology Volume 2, The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press.

- Loeb, D. 1998: Moral realism and the argument from disagreement. *Philosophical Studies*, 90, 281-303.
- Machery, E., Kelly, D., and Stich, S. 2005: Moral realism and cross-cultural normative diversity. *Behavioral and Brain Sciences*, 28(6), 830.
- Mackie, J. L. 1977: *Ethics: Inventing Right and Wrong*. New York: Penguin.
- Mercier, H. and Sperber, D. 2011: Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34, 57-111.
- Mikhail, J. M. 2010: Rawls' concept of reflective equilibrium and its original function in *A Theory of Justice*. *Washington University Jurisprudence Review*, 3(1), 1-30.
- Mikhail, J. M. 2011: *Elements of Moral Cognition*. New York: Cambridge University Press.
- Nichols, S. 2004. *Sentimental Rules*. New York: Oxford University Press.
- Nisbett, R. E. and Cohen, D. 1996: *Culture of Honor: The Psychology of Violence in the South*. Boulder, CO: Westview Press.
- Paxton, J. M. and Greene, J. D. 2010. Moral reasoning: hints and allegations. *Topics in Cognitive Science*, 2, 511-527.
- Pizarro, D. A., Uhlmann, E., and Bloom, P. 2003: Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology*, 39, 653-660.
- Rawls, J. 1971: *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Saunders, L. 2009: Reason and intuition in the moral life: a dual-process account of moral justification. In J. Evans and K. Frankish (eds.), *In Two Minds: Dual Processes and Beyond*. New York: Oxford University Press.
- Sayre-McCord, G. 1996: Coherentist epistemology and moral theory. In W. Sinnott-Armstrong and M. Timmons (eds.), *Moral Knowledge?* New York: Oxford University Press.
- Shafer-Landau, R. 2003: *Moral Realism: A Defense*. Oxford: Clarendon Press.
- Sinnott-Armstrong, W. 2006: Moral intuitionism meets empirical psychology. In M. Timmons and T. Horgan (eds.), *Metaethics after Moore*. New York: Oxford University Press.
- Smith, M. 1994: *The Moral Problem*. Oxford: Blackwell.
- Sneddon, A. 2009: Normative ethics and the prospects of an empirical contribution

to assessment of moral disagreement and moral realism. *Journal of Value Inquiry*, 43, 447–455.

Sripada, C. S. and Stich, S. 2006: A framework for the psychology of norms. In P. Carruthers, S. Laurence, and S. Stich (eds.), *The Innate Mind (Volume 2): Culture and Cognition*. New York: Oxford University Press.

Stich, S. 1990: *The Fragmentation of Reason*. Cambridge, MA: MIT Press.

Stich, S. 2009: The persistence of moral disagreement. Lecture 2 of the *Leverhulme Lectures on Moral Psychology*, University of Sheffield, UK, May 11th 2009. Video available online at: http://www.philosophy.dept.shef.ac.uk/hangseng/videos/02_persistence_moral_disagreement.php.

Wheatley, T. and Haidt, J. 2005: Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16, 780-784.

Wong, D. 2006: *Natural Moralities: A Defense of Pluralistic Relativism*. New York: Oxford University Press.

Note: Figure 1 omitted. Please see the published version.