University of Groningen

Intuitions, Rationalizations, and Justification

Hindriks, Frank

Published in:
Journal of value inquiry

DOI:
[10.1007/s10790-014-9419-z](https://doi.org/10.1007/s10790-014-9419-z)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2014

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hindriks, F. (2014). Intuitions, Rationalizations, and Justification: A Defense of Sentimental Rationalism. *Journal of value inquiry*, 48(2), 195-216. <https://doi.org/10.1007/s10790-014-9419-z>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Intuitions, Rationalizations, and Justification: A Defense of Sentimental Rationalism

Frank Hindriks

Published online: 9 April 2014
© Springer Science+Business Media Dordrecht 2014

People sometimes make moral judgments on the basis of brief emotional episodes. I follow the widely established practice of referring to such affective responses as intuitions (Haidt 2001, 2012; Bedke 2012, Copp 2012). Recently, a number of moral psychologists have argued that moral judgments are never more than emotion- or intuition-based pronouncements on what is right or wrong (Haidt 2001, Nichols 2004, Prinz 2007). A wide variety of empirical findings seem to support this claim. For example, some argue that arbitrary emotional responses or intuitions induced under hypnosis elicit moral judgments (Wheatly and Haidt 2005). Furthermore, intuitions function as the point of last resort in attempts to justify moral judgments (Haidt, Björklund, and Murphy 2000). On the basis of such evidence, psychologists such as Jonathan Haidt (2001, 2012) and philosophers such as Shaun Nichols (2004) and Jesse Prinz (2007) defend what I call ‘Subjective Sentimentalism’, which consists of three claims about moral intuitions and the moral judgments they give rise to.¹ The first claim is that moral intuitions are affective responses. The second is that moral intuitions cause but do not justify moral judgments. And the third is that the attempts people sometimes make to justify these judgments typically fail, because moral reasoning tends to be biased or confabulated.²

¹ Nichols (2004) and Prinz (2007) formulate their theories in terms of emotions and sentiments rather than in terms of intuitions. The way in which these feature in their theories, however, matches the way in which moral psychologists use the term ‘intuition’ and how they conceive of their relation to emotions. Discussing their theories in terms of intuitions facilitates an insightful comparison to other positions.

² In section 1.1 I consider whether reasoning plays other more positive roles in some versions of Subjective Sentimentalists.

Around the turn of the nineteenth century, Henry Sidgwick (1874) and G.E. Moore (1903) defended a position known as Moral Intuitionism.³ The notion of an intuition lies at the heart of this theory about moral judgment, just as in the case of Subjective Sentimentalism. In contrast to Subjective Sentimentalists, however, Moral Intuitionists do not regard emotions as a proper source of moral judgment. Emotions are too fickle to fulfill this purpose. According to Moral Intuitionism, genuine moral intuitions are not based on emotions. Instead, they are self-evident and require rational capacities to be appreciated as such. Whereas Subjective Sentimentalists deny that moral judgments can be justified, Moral Intuitionists affirm this. And they hold that moral knowledge is possible, because reason provides justification for moral beliefs.

These two positions provide contrasting perspectives on intuition, emotion, and reason, as well as on the roles they play in moral judgment. Although I reject both positions, the alternative I defend draws on each of them. I defend two claims. First, even though this kind of reasoning is often of poor quality, the available evidence supports the claim that at least some of the time moral reasoning provides justification for moral judgments. In other words, a substantial proportion of the population engages in unbiased moral reasoning and does not confabulate (see section 2 for a discussion of the data). Subjective Sentimentalists over-generalize and fail to appreciate the significance of the responses of the minorities in the empirical evidence they present. The second claim that I defend here concerns the way in which moral reasoning works: affect and cognition interact in the formation of moral judgments. This claim derives support from experiments concerning cognitive dissonance. As it turns out, an apparent cognitive conflict can trigger an unpleasant affective response, which in turn motivates reasoning aimed at resolving the cognitive conflict. Although such reasoning is often biased, there is no reason to believe that this is always the case. The reasoning triggered by cognitive dissonance can in fact provide warrant for the moral judgment that results from it, or so I argue below. If it does, affect plays a constructive role in moral judgment. This implies that, *pace* Moral Intuitionism, emotions can contribute to the justification of moral beliefs.

As it combines elements from sentimentalism and rationalism, I call the position that I defend in this paper ‘Sentimental Rationalism’.⁴ The central idea of Sentimental Rationalism is that affect and cognition can contribute to the justification of moral beliefs together. Others have criticized Subjective Sentimentalism before, and Sentimental Rationalism is not the only hybrid proposal (Fine 2006, Allman and Woodward 2008, Craigie 2012, and Sauer 2012a, 2012b).

³ See Ross (1930), Huemer (2005) and Audi (2004, 2013) for more recent defenses of this position. Roeser (2011) argues that more than a century earlier Reid (1785) defended a version of moral intuitionism.

⁴ Sentimental Rationalism is to be distinguished from Rational Sentimentalism. Rational Sentimentalism is the view that values are constituted by sentiments that are subject to rational assessment (D’Arms and Jacobson 2000). Sentimental Rationalism, in contrast, is the view that affect and cognition can contribute to the justification of moral beliefs together. Note that Rational Sentimentalism is not restricted to moral values, but extends to (judgments about) non-moral values. Sentimental Rationalism is not restricted to moral value judgments but also encompasses other moral judgments including those about right and wrong.

However, the evidence concerning cognitive dissonance and moral reasoning presented in section 2.2 – Albert Bandura’s moral disengagement studies – has not yet been discussed in this context. And I use this evidence to construct an account of how moral judgments could be justified that is consistent with these studies as well as other recent empirical findings. More specifically, I propose a normative theory of moral judgment in which the mutual interaction between affect and cognition plays a central and distinctive role. This theory is based on a descriptive account of moral reasoning to which I refer as ‘the Cognitive Dissonance Model’ (section 3). Before turning to the evidence concerning cognitive dissonance, I critically discuss how Subjective Sentimentalists invoke evidence concerning moral reasoning in support of their position (section 2.1). I begin by introducing Subjective Sentimentalism and Moral Intuitionism (section 1).

1 Intuitions in Ethics

1.1 Subjective Sentimentalism

Perhaps the most striking evidence in favor of Subjective Sentimentalism are the hypnosis experiments mentioned in the introduction (Wheatly and Haidt 2005). In these experiments, some connection is forged between morally irrelevant words such as ‘take’ and ‘often’ on the one hand, and morally charged emotions such as disgust on the other hand. Subsequently, participants read morally neutral scenarios that contain the relevant words. As it turns out, they are then inclined to say that the agent in the scenario did something wrong even though they cannot explain why. Apparently, an intuitive response – in this case a flash of disgust – is enough to trigger a moral verdict.

I refer to these experiments as ‘the dumbfounding experiments’, as they involve participants who are dumbfounded in the sense that they are unable to back up their moral judgments with arguments when asked to do so (Haidt, Koller, and Dias 1993). These experiments do not involve scenarios concerning harm or rights violations but scenarios that are morally charged in some other way. Examples are using a flag to clean a toilet, masturbating with a chicken before eating it, and eating your pet dog that just died on the street in front of your house. People have the intuition that these actions are morally wrong. When pressed to defend their verdicts, they provide justifications that are explicitly ruled out by the scenario, or they invent victims. And once they realize that they have run out of arguments, they express their emotions, for instance by saying that the relevant action is simply disgusting. Jonathan Haidt (2001, 2012) argues that the arguments people provide tend to be biased or without any basis: they commonly constitute rationalizations or confabulations.⁵ Haidt points out that ultimately we are left with nothing but our

⁵ Ever since Davidson (1963), philosophers typically take rationalizations to consist of explanations that successfully explicate the reasons on the basis of which the agent acted. In contrast to this, Haidt uses the term in a derogatory way to designate bad or biased arguments. Confabulations are invented arguments that have no basis in reality at all (see also Nisbett and Wilson 1977, Wilson 2002, and Johansson et al. 2005).

emotions as a basis for the judgment made. Reason functions as a lawyer of those emotions, and not as an unbiased and objective judge.⁶

As I see it, the core of Subjective Sentimentalism (SS) consists of the three claims concerning moral intuitions and the moral judgments they give rise to mentioned in the introduction (here I present more precise versions of them). The first claim is: (1) Moral intuitions are affective responses (or dispositions to display such responses). As Subjective Sentimentalists hold that moral intuitions form the immediate basis of moral judgments, this implies that emotions are the ultimate and only source of moral judgments. This means that rational thought does not provide for an alternative or a complementary source. The second claim concerns the question whether moral intuitions are justified or provide justification for moral judgments: (2) Moral intuitions are a-rational in that they cause but cannot justify moral judgments. The third statement addresses the way in which people try to justify their moral judgments: (3) Apparent justifications of moral judgments typically fail in that moral reasoning tends to be biased or without any basis.

SS can be reformulated in the language of Dual System Theory along the following lines.⁷ Intuitions are generated by System I, which means that they are unmediated and a-rational. System II, the conscious planner or reasoner, attempts to provide a justification for these responses, but is unable to offer good arguments.⁸ The conclusion of those arguments is fixed beforehand, and consists of whatever judgment happens to be supported by the intuitive response at hand. Hence, rationalizations are always post hoc. Furthermore, the arguments people provide in favor of their moral judgments are either biased or unfounded, they constitute rationalizations or confabulations. Given that emotions are a-rational on SS, this could not be any other way.

In the introduction, I mentioned Haidt, Nichols and Prinz as proponents of SS.⁹ All of them postulate an intimate connection between intuitions and moral judgments. Prinz contrasts his own view to those of Nichols and Haidt by pointing out that, on his view, this relation is not causal but constitutive (2007: 98–99). Both views, however, are consistent with SS. One might worry that statement (3), the claim that apparent justifications typically fail, is not underwritten by all three of them. Haidt has provided a lot of evidence for it and endorses it explicitly. Nichols (2004) and Prinz (2007), however, hold that reasoning has an important and constructive role to play, which could mean they do not subscribe to claim (3). It

⁶ Haidt's (2012) Moral Foundations Theory encapsulates the claim that moral judgments are based on emotions. According to his Social Intuitionist Model, moral reasoning is mostly post hoc rationalization or confabulation.

⁷ See Evans (2008) for a review of Dual System Theories.

⁸ I follow Kahneman (2011) in talking about the two systems as distinct entities without committing myself to the claim that they are more than analytically distinct.

⁹ Nichols (2004) presents his Sentimental Rules Theory according to which an affect-based normative theory provides the basis for moral judgment. Prinz's (2007) Constructive Sentimentalism is a subject relative form of sentimentalism that focuses instead on individual dispositions to experience moral emotions. Another difference worth noting here concerns the object of intuition. Whereas Nichols holds that moral intuitions pertain to norms or principles, Prinz and Haidt take moral intuitions to concern particular kinds of actions.

will be useful to look at their views in some detail and check whether the claim that they subscribe to SS needs to be qualified in some respect.

On Nichols' view (2004: 100–101), reasoning pertaining to moral judgments primarily consists of reasoning about the content of the normative theory. In addition to norm violations, moral reasoning can concern intentions and consequences (*ibid.*: 103–104). Such reasoning can be pretty advanced, as it often requires people to evaluate counterfactuals. The key question that reasoning serves to answer on this picture is whether a particular action genuinely violates a norm. As Nichols approaches it, this seems to be a factive question. A reasoned answer to this question can provide warrant to the relevant moral judgments. This reasoning does not, however, extend to discussions about the aptness of the norms or the appropriateness of affective responses, which are genuinely normative matters. At this more fundamental level, then, warrant plays no role. Statement (3) concerns this level. In light of this I see no reason to doubt that Nichols subscribes to all three theses of SS after all.

Prinz (2007) argues that values are constituted by sentiments, which he takes to be dispositions to display affective responses. Non-moral reasoning can play an important role in his theory insofar as it serves to determine whether a given value applies to a given case (*ibid.*: 124–125). Perhaps his theory even allows for intrapersonal moral reasoning, as individuals might have to weigh different values in a given case. On Prinz's view, however, moral judgment is subjective. He goes as far as to claim that, when two individuals form different judgments, they do not disagree. Each moral judgment is after all relative to the particular agent who makes it. There is no genuine inconsistency between my view that it is wrong to eat your pet dog, and someone else's view that it is permissible to do so. Prinz's subjectivism implies definite limits to the extent to which moral reasoning can be warranted. All this means that, even if Prinz might see reason to qualify thesis (3), he does not fundamentally disagree with it.¹⁰

Haidt (2012) might instead have problems with the second thesis of Subjective Sentimentalism (2). On the one hand, he describes our affective receptors as taste buds, which suggests emotions are a-rational and cannot be justified.¹¹ On the other hand, however, he claims that emotions typically involve cognitive appraisals of events that prepare people to respond appropriately (*ibid.*: 44–45). This suggests that they can be rational. I do not know how to resolve this tension. To the extent that this second feature of his position is more true to Haidt's account, his allegiance to Subjective Sentimentalism has to be qualified just as some of my criticisms.

¹⁰ There is in fact little room to maneuver here for Prinz. He rejects an analysis of sentiments in terms of normal conditions for normal subjects (Prinz 2006, 2007). This means he cannot appeal to the conditions under which a moral judgment is made in order to explain why it is wrong. Hence, the space for moral error in Prinz's theory is rather small.

¹¹ Haidt clarifies the relation between emotions and intuitions as follows: 'Moral emotions are one type of moral intuition, but most moral intuitions are more subtle; they don't rise to the level of emotions.' (2012: 45) Later he adds that 'most of our affective reactions are too fleeting to be called emotions' (*ibid.*: 55). In addition to appraisals, affective intuitions encompass alterations of attention and vigilance (*ibid.*: 329n40). This is why I use the term 'affective response'.

Before turning to the empirical findings adduced in support of Subjective Sentimentalism, I introduce Moral Intuitionism.

1.2 Moral Intuitionism

Moral Intuitionists such as Henry Sidgwick, G.E. Moore, and W.D. Ross have a very different view of what intuitions are. Intuitions are not unconsidered responses, but require reflection. An intuition has rational standing by definition. Intuitions concern general moral principles such as the duty to keep promises.¹² Moral Intuitionism (MI) can be characterized in terms of three claims that run parallel to those introduced above for SS. Again, the first statement concerns the nature of moral intuitions. Moral intuitionists take moral intuitions to be akin to perceptions in part because of their immediacy. Moral intuitions appear in someone's consciousness, but there is no trace of any antecedents there might be. This motivates (some) Moral Intuitionists to postulate a moral sense. (1) Moral intuitions are due to a special moral sense: moral intuition.¹³

The evidentiary role of moral intuitions is made explicit in the second statement: (2) Moral intuitions are typically warranted. They derive their warrant at least in part from the fact that they are the intuitions of rational agents. This second statement is diametrically opposed to the claim of SS that moral judgments cannot be warranted, and are based on emotions rather than reasoning. The third claim sheds more light on exactly how moral intuitions provide warrant for moral judgments. (3) Moral intuitions justify moral judgments non-inferentially.¹⁴

Moral Intuitionists regard moral intuitions as self-evident in a way that is similar to the self-evidence of mathematical axioms. They provide the basis for moral knowledge. The step from moral intuition to moral judgment is not a matter of induction, deduction, or abduction on their view. Whereas a lot of true moral statements bear logical relations to each other, some are special in that they provide the foundation for other statements while they are not themselves based on other statements. Moral intuitions are (typically regarded as) such 'unmoved movers'. They form the foundations of moral knowledge, they are not inferred from anything else, and provide independent warrant for moral judgments.¹⁵

Just as Subjective Sentimentalists, Moral Intuitionists are well aware of the fact that emotions often form the causal basis of moral judgments. However, they have a very different view of their status. They are hardly concerned with what happens to cause someone's judgments, if at all. Instead they ask what *should* be the basis of moral judgments. Emotions form a distraction insofar as this project is concerned, or so they maintain. Rational reflection provides for a way of avoiding such distorting

¹² Audi (2013) hold that moral intuitions can also concern particular cases.

¹³ Not all intuitionists accept this thesis. Ross (1930) and Huemer (2005) explicitly deny it.

¹⁴ Cowan combines statements (1)–(3) when he defines MI as follows: 'Normal ethical agents can and do have epistemically independent non-inferential justification for first-order ethical beliefs that is the result of substantive ethical thinking.' (forthcoming)

¹⁵ At least contemporary Moral Intuitionists regard moral intuitions as *defeasible* evidence for moral judgments (Audi 2004, Huemer 2005).

factors. The true normative theory is that which survives rational reflection. Hence, Moral Intuitionists are rationalists rather than sentimentalists.¹⁶

Even though SS and MI are very different views, it is useful to discuss them in combination because we can learn from both. As discussed, MI holds that intuitions are rational insights. SS maintains instead that intuitions are a-rational affective responses. In section 2.1 I criticize SS for rejecting moral justification and argue that it mistakenly dismisses reasoning as a source of moral justification.¹⁷ In section 2.2 I argue that MI mistakenly rejects affect as an appropriate source of moral judgment. This argument invokes research concerning cognitive dissonance and moral judgment that suggests that cognition and affect can contribute to the justification of moral judgments together. I present my alternative, Sentimental Rationalism, in section 3. Just as SS, it regards affect as a proper basis of moral judgment. Just as MI, it regards intuitions as sources of justification.

2 The Evidence

The data collected in the past two decades or so have established beyond doubt that emotions influence moral judgments in significant ways. Whether they do so for better or worse, however, remains to be seen. The data as such do not reveal whether emotions distort moral judgments, or whether they can serve to justify them. The thing to do in order to make progress with respect to this issue, I argue in this section, is to evaluate the role of emotion and of reasoning together. This is in effect what Haidt does in the dumbfounding experiments (section 2.1) and Bandura in what I call ‘the disengagement experiments’ (section 2.2). These experiments reveal that SS is too pessimistic about the role of cognition and that MI is too negative about the role of affect. In section 3, Sentimental Rationalism emerges as the synthesis that avoids the pitfalls and combines the virtues of both.

2.1 Moral Dumbfounding

Haidt and his colleagues present those who participate in the moral dumbfounding experiments with scenarios that concern harmless taboo violations (Haidt, Björklund, and Murphy 2000). The underlying idea is that people will be inclined to regard the depicted actions as wrong because they are taboo violations. At the same time, it will be difficult for them to explain why those actions are wrong because they do not involve rights violations or harm. These scenarios have been carefully constructed not to include rights violations or harm.

In addition to the flag, dog, and chicken scenarios mentioned in section 1.1, Haidt presents a scenario concerning sibling sex. The brother and sister in the story decide to commit incest. However, they expect to enjoy having sex with each other. They

¹⁶ Roesser (2011) forms an exception in this respect. According to her Affectual Intuitionism, the non-inferential moral beliefs that form the foundation of moral knowledge are emotions. Audi (2013: 123–124) denies that affective responses can be intuitions. He does, however, regard them as a source of intuitions and he allows for the possibility that intuitions partly constitute emotions (ibid.: 172 and 133).

¹⁷ See also Craigie (2012), Fine (2006), and Sauer (2012a, 2012b).

decide to do it only once in order to be on the safe side and avoid any negative consequences having sex with a sibling might have. As expected, they do not experience any psychological distress afterwards. Furthermore, they use contraception in order to rule out the possibility of pregnancy. Hence, there is no possibility of a baby being born with birth defects due to the genetic risks involved in having sex with relatives. After reading scenarios such as this one, participants are asked whether the act depicted in them is wrong.

Most people say yes. They are then invited to provide arguments for their verdict. As it turns out, most answers are clearly beside the point. People raise worries that have been ruled out in the scenario, such as that incest increases the probability of birth defects. When it is pointed out to them that the problems they raise are not there, people try to come up with other complaints. They go as far as inventing victims, as when they suggest that the family would get sick from eating their pet's meat (Haidt 2012: 24). In the end, however, most people are left with nothing else to say than that it is wrong but that they do not know why, or that it is simply disgusting.¹⁸ Thus, they discover that they are unable to give reasons for a judgment they thought they could justify.

In his attempt to explain moral dumbfounding, Haidt focuses on the role that emotions might play in the formation of moral judgments. He suggests that affective responses such as disgust trigger people to form the belief that the act at issue is wrong. Subsequently, they put a lot of effort in making that verdict look justified. Their reasoning functions, as Haidt puts it, as a lawyer of the emotion-based verdict rather than as a disinterested judge.

The first thing I would like to point out in response is that if SS is true, it is hard to understand why people engage in any moral reasoning at all. According to SS, (basic) moral judgments are based on emotions and cannot be justified. This leaves the puzzle of explaining why people would bother to adduce any support at all. These people would be making a category mistake, if SS were true. As moral judgments express sentiments rather than beliefs, they do not require any evidence. It will not do to say that people simply feel the need to justify themselves. People feel little pressure to provide arguments in defense of their personal preferences (Haidt 2012: 44). In light of this, Subjective Sentimentalists face the challenge to explain the following: if moral judgments are not susceptible to warrant, just like personal preferences, why do people feel the urge to justify the former but not the latter?¹⁹

The second thing to note is that there is more to the evidence than what was mentioned thus far (Haidt, Björklund, and Murphy 2000, Haidt 2012: 36–40). Some of the participants changed their view when they discovered that their arguments were defective. Rather than affirming the consequent and stating that it is simply wrong, they reconsidered their view and ended up admitting that the act was in fact

¹⁸ Haidt, Koller and Dias (1993) distinguish four kinds of justifications: autonomy, community, divinity, or norm statements. They regard appeals to disgust as divinity statements, and claims to the effect that an action is wrong as norm statements.

¹⁹ In a similar vein, Ditto, Pizarro and Tannenbaum (2009: 323) observe that people feel more pressure to provide justification for moral claims as compared to aesthetic claims. Their explanation is that, whereas aesthetics is usually seen as a matter of taste, most people are naive moral realists most of the time.

unproblematic. These participants treat their inability to come up with good reasons against the action as an indication that in the case at issue there is no reason to forbid it. This observation reveals that Haidt over-generalizes when he qualifies moral reasoning as post hoc rationalization and confabulation. This is an important conclusion to draw as it opens up the possibility that in each of these experiments a minority of the participants might argue in a way that approximates that of a judge and is pretty far removed from that of a biased lawyer.

In order to see how rash Haidt's overgeneralization is, we need to consider the minority that was not morally dumbfounded. Those who change their mind tend to be people with a high socio-economic status (SES). Apparently, moral dumbfounding is influenced by individual differences concerning, for instance, education and income. Another factor that matters is time. More people change their mind when they are given the time to reflect on the issue, or when they are presented with better arguments (Haidt 2012: 69, Paxton, Ungar and Greene 2012). It is far from implausible that these factors are systematically related to quality of judgment formation.

Daniel Batson and his colleagues (1997, 1999) have conducted research concerning moral hypocrisy that also sheds light on which factors determine the quality of moral judgments. Batson argues that people care more about appearing to be moral than about being moral.²⁰ This means that the quality of their reasoning depends on how easy it is to get away with flaws, and that depends on circumstantial factors. People flip a coin, for example, to appear fair, but ignore the outcome of the coin flip when they can get away with it in the sense that the violation of the standard is neither too obvious to the participants themselves nor apparent to others. Facing a mirror during the experiment, however, turns out to make a difference. Looking in a mirror increases your self-awareness. And in contrast to the original setup, a substantial number of people who face a mirror do abide by the outcome of the coin flip.

It may well be that circumstantial factors such as mirrors sometimes serve to increase the quality of someone's moral reasoning. The idea would be that mirrors heighten self-awareness, which makes it more difficult to avoid noticing the discrepancy between the envisaged action and the moral standard of fairness at issue. As a consequence, the participant might no longer be able to deceive himself into believing that the conflict is only apparent. The discrepancy can plausibly be taken to trigger cognitive dissonance. It is relatively easy for lots of people either not to notice the conflict, or to refrain from incurring any self-sanctions. Such moral disengagement becomes more difficult, however, when the agent's self-awareness is high. The upshot is that there is a range of factors that affects the quality of moral reasoning and moral judgment. This provides reason to doubt the second thesis of SS (2), the claim that moral intuitions are a-rational. After all, to the extent that these factors do indeed bear on the quality of moral judgments, there must be a measure of their quality.

²⁰ Haidt makes a similar claim when he maintains that 'we care more about looking good than about truly being good' (2012: 190).

The second thing to see is that, as soon as it is granted that moral judgments can be justified or appropriate, it does not always matter much whether people care only about appearing to be moral or whether they genuinely want to be moral. The important point is that in some circumstances they will not get away with appearing to be moral without being moral, or approximating the moral ideal. Factors such as SES, heightened self-awareness, and time to reflect, as well as the presence of others make it more difficult to deceive others as well as themselves about the morality of their actions. Wanting to appear moral will all by itself already have favorable consequences for the quality of moral reasoning in some contexts.²¹ The upshot is that, *pace* SS, the evidence concerning moral dumbfounding and moral hypocrisy supports the idea that the moral judgments people form are or come close to being justified or appropriate at least some of the time.²²

2.2 Moral Disengagement

Moral disengagement facilitates people to do things that conflict with their own moral standards. It typically concerns an action that an agent envisages herself as performing that appears to be in conflict with her moral standards. The agent is tempted to violate her own norms, but experiences anticipatory guilt feelings about doing so. This affective response triggers a process of reasoning aimed at double-checking whether the apparent conflict is genuine. Mechanisms of moral disengagement are types of rationalization strategies people can use in order to deceive themselves into believing that the apparent conflict is nothing more than that, an *apparent* conflict that dissolves on closer inspection. Bandura et al. (1996) refer to this process as moral disengagement because the agent disengages moral self-sanction from possibly immoral conduct.

In one of the disengagement studies, the participants are prison personnel from maximum-security penitentiaries (Osofsky, Bandura, and Zimbardo 2005). These participants rate statements on a 5-point Likert scale ranging from strongly agree (2) to uncertain (0) to strongly disagree (−2). One of the statements is: ‘Capital punishment is just a legal penalty for murder.’ Michael Osofsky, Albert Bandura, and Philip Zimbardo take this to be an example of euphemistic labeling, one of the eight disengagement mechanisms that Bandura distinguishes.²³ They classify the statements ‘An execution is merciful compared to a murder’ and ‘Those who carry out state executions should not be criticized for following society’s wishes’ as

²¹ Social psychologists distinguish accuracy, impression, and defense motivation. Rather than about accuracy, agents sometimes care primarily about the image or impression others have about them or about their self-image and whether it remains defensible (Chaiken et al. 1996). In these terms, the argument presented in the main text comes down to the claim that impression and defense motivation sometimes have effects that are rather similar to those that accuracy motivation would have.

²² This is confirmed by research concerning dishonesty. When people are in the position to be dishonest, they behave dishonestly but only to a relatively small extent. The line they do not cross is that beyond which they can no longer maintain their self-image as honest people (Haidt 2012: 83; Mazar et al. 2008).

²³ See Bandura et al. (1996), and Bandura (1999). Four mechanisms of moral disengagement pertain to conduct and its consequences: moral justification, advantageous comparison, euphemistic labeling, disregarding or distorting the consequences. Both displacement and diffusion of responsibility pertain to the agent. Dehumanization, and attribution of blame pertain to the victim.

advantageous comparison and displacement of responsibility respectively (*ibid.*: 379). As it turns out, executioners employ all of the eight kinds of rationalizations. Among the other participants in the study were members of the support teams that provide solace and emotional support to the families of the victims and the condemned inmate. They are unlikely to engage in dehumanization and moral justification (*ibid.*: 387).²⁴

The executioners are expected to kill as part of their job. Experiential reports reveal that they experience resistance to killing and that they manage their thought processes in order to enable themselves to go through with it. Note that, even if it were sometimes morally justified to execute prisoners, the extent of moral disengagement as well as its content strongly suggests that many of the facilitating thoughts are of mediocre quality at best. Even if they endorse these claims, they should be able to see on reflection that they do not support the relevant action. The thing that matters here, however, is that they help the agents believe that, in spite of initial appearances, there is in the end no substantial conflict between the agent's norms and the envisaged action. The idea is that, once the consistency has been established, the remaining obstacle for performing the action has vanished and the agent goes ahead and carries out his plan.

The disengagement experiments provide little reason for optimism about moral conduct or about moral reasoning. The extent of moral disengagement, however, is not uniform across the population. Factors that bear on the extent to which people disengage include age, gender, education and race. Older people, men, the lesser educated, and people from a Caucasian background are more prone to morally disengage than others.²⁵ These findings concerning individual differences imply that some people are significantly more likely to behave in accordance with their moral standards than others.²⁶

Bandura's moral disengagement theory is rather explicit about the underlying mechanism. Although he rarely if ever uses the term, the mechanism can plausibly be said to concern cognitive dissonance (Moore 2008). What is fascinating about cognitive dissonance is that cognition and affect both play a role in it in rather revealing ways (Festinger 1957). The mechanism concerns agents who have internalized certain standards of rationality such as consistency. Such agents use these standards to evaluate the actions they consider performing. They register conflicts between such actions and their moral standards. Those cognitive conflicts trigger affective responses. People are susceptible to feel guilt prior to the act that clashes with their moral standards, a feeling that is known as 'anticipatory guilt

²⁴ Agreeing or disagreeing with a statement is indicative of moral disengagement only when the statement conflicts with the moral commitments of an agent. Osofsky, Bandura, and Zimbardo (2005) do not measure such discrepancies directly. They do, however, interview the participants and discuss their emotional reactions in preparation, during, and after an execution and the ways in which they tried to manage their stress (*ibid.*: 381, 389–390). To the extent that their stress is indicative of cognitive dissonance, the interviews provide indirect evidence in support of moral disengagement.

²⁵ As it happens, age is the only one of these factors that makes a difference in the study just discussed (Osofsky, Bandura, and Zimbardo 2005: 381 and 387).

²⁶ According to Aquino and Reed (2002), people can have a moral identity in the sense that their moral standards can play a role in their self-conception. They provide evidence that people who have a strong and accessible moral identity violate their moral standards less often than others.

feeling'. The term 'cognitive dissonance' refers to anticipatory guilt feelings that are caused by cognitive conflicts. As a consequence of this affective response, the agent initiates a process of reasoning aimed at resolving the cognitive conflict in one way or another. When it is resolved, the affect dissolves.

The conflict can, of course, be resolved by refraining from the envisaged action. However, it can also be resolved by arriving at the conclusion that the conflict was only apparent. And perhaps it was. The possibility to which the theory of moral disengagement highlights, however, is that in which some mechanism of disengagement provides the agent with a new perspective on the action such that it no longer appears to be in conflict with his moral standards even though it in fact still is. The agent deludes himself into thinking that it is not. The disengagement experiments suggest that people are rather good in deceiving themselves, in mistaking bad reasons for good ones, when it comes to moral matters. The driving force underlying disengagement is the desire to maintain self-consistency. The preceding suggests that *apparently* maintaining self-consistency is not so difficult for most of us.

How can a positive view on the formation of moral judgments be defended in the face of such a gloomy conclusion? As in section 2.1, the first step is to turn our attention to the minority. In contrast to Haidt, Bandura does not overgeneralize, but is sensitive to the fact that different people disengage to different extents. In Celia Moore's (2008) terms they differ in their 'propensity to morally disengage'. In light of the findings presented earlier, it might be that highly educated young women from a non-Caucasian background tend to have a rather low propensity to morally disengage.²⁷ The second step is to consider what role cognitive dissonance might play in this. The account of cognitive dissonance just presented is detailed enough to identify what I call 'the cognitive dissonance mechanism' that proceeds from a cognitive conflict, to anticipatory guilt feelings, to reasoning, and finally to moral disengagement. When people with a low propensity to morally disengage notice a cognitive conflict, they are likely to experience an affective response. However, they are less likely to fall into the trap of any one of the disengagement mechanisms. And the arguments that convince them, if any, are likely to be relatively unbiased. They will not easily be satisfied with the thought that the conflict is merely apparent. The upshot is that their judgments as to whether the envisaged action is in conflict with their moral standards tend to be better than those of the rest of us.

Given this way of describing the differences between people with a low and a high propensity to morally disengage, it is natural to conclude that moral reasoning can serve to reduce or eliminate certain errors. This makes sense only if moral judgments can in principle be justified. If this were impossible – as SS supposes it is – moral reasoning would be pointless. Note that justification is a matter of degree. A moral judgment can be more or less unbiased or justified. And the discussion of cognitive dissonance suggests that the affective responses people experience in reaction to a cognitive conflict can be rather informative. They indicate that the

²⁷ All we know is that the individual features correlate with propensity to morally disengage. This can be true even if some combinations of these features do not.

agent's cognitions might be flawed in some respect. This in turn triggers a process of reasoning aimed at resolving the cognitive conflict.

An agent's sensitivity to cognitive conflicts can be likened to a car alarm in that it alerts the agent to possible flaws. Over time System II can fine-tune or calibrate System I, just as a car alarm can be set such that it is appropriately sensitive to genuine threats and not too sensitive in order to avoid too many false alarms (see also section 3.1). The important point to appreciate here is that cognitive dissonance involves interaction between cognition and affect that serves to eliminate possible defeaters and to register evidence. Given how they interrelate, cognition and affect are conducive to justified moral judgments. This implies that a combination of informed affect and sound reasoning can give rise to adequate moral judgments.²⁸

Before I turn to the question how these insights can be used to construct an alternative to SS and MI, let me comment on the kind of argument I have presented in this section. In section 2.1, I presented what I will call 'the argument from overgeneralization' in order to motivate an investigation of minority responses in the dumbfounding experiments. A number of features were identified that decrease the probability that an agent's reasoning and judgment is biased. In this section, I have discussed the cognitive dissonance mechanism and I have explicated how it functions under ideal circumstances. The methodological point that I have made is that differences in degree sometimes provide information about processes that averages cannot reveal. More specifically, minority responses can turn out to be a model for all in the sense that they exemplify an ideal that others try to emulate. The thing to appreciate is that, even though in quantitative terms the minority might be insignificant, it may be that it should play a dominant role when it comes to conceptualizing the underlying mechanism. I will refer to this second line of argument as 'the dominant minority argument'.

A dominant minority is a group that is only a small fraction of the overall population but holds a disproportionate amount of power. Just as a dominant minority has a lot of social power, the minorities involved in the experiments discussed in this paper are rather revealing about the determinants of people's moral psychology than the majorities. Rather than social power, they have a lot of explanatory power. This holds even if no member of the minority actually reaches the moral ideal – for instance the ideal of never morally disengaging.

One may want to object to the dominant minority argument and point out that the fact that people are climbing a mountain does not imply that they reach the summit. Similarly, the thought would be, the fact that some of the time people reason in an unbiased way does not entail that they acquire moral knowledge or even that there is such a thing as moral knowledge. In response I would like to point out that the fact that people are climbing does imply that there is a mountain and that some people may reach a higher altitude than others. This holds even if nobody reaches the summit. Suppose that the summit represents moral certainty, and the slope of the mountain (the degree of) justification. Disoriented climbers make little or no

²⁸ This line of reasoning suggests that it is too simple to think, as is often thought, that System I is fast but unreliable, and System II slow but normatively superior. Note also that on my view affect and cognition can each play a role in both systems.

headway, if they do not loose altitude. It may well be, however, that some climbers make progress. The research discussed provides some reason to believe that there is altitude to be gained and thereby reason to think that Subjective Sentimentalism is false. If it is, we need an alternative.

3 Sentimental Rationalism

3.1 Emotion and Reasoning in Moral Judgment

Affect as well as cognition play a constructive role in the formation of moral judgments some of the time. More specifically, both affect and cognition can contribute to the justification of moral beliefs, and they often do so together. This is the core thesis of Sentimental Rationalism (SR), the alternative view that I defend in this section.²⁹ Just as SS and MI were explicated in terms of three claims above, I characterize SR in terms of three claims here. The first statement concerns the nature of moral intuitions: (1) Moral intuitions are often affective responses (or dispositions to display such responses). When someone has a moral intuition, she will spontaneously regard some moral judgment as compelling. It will often require effort to resist forming the relevant belief. This may be due to an affective response. SR is more ecumenical than SS in that it does not insist that all moral intuitions are affective. For all I have argued, some might be intellectual rather than affective appearances (Huemer 2005), or even self-evident beliefs, i.e. beliefs that are basic in that their content is not inferred from anything else (Audi 2004). In contrast to (some versions of) MI, SR allows for intuitions that are affective. Furthermore, it does not require recourse to a moral sense.

The second claim concerns the rationality of moral intuitions: (2) Moral intuitions themselves can be more or less warranted. This is in part due to the fact that they are the intuitions of rational agents. The notion of rationality that features in this claim is relatively undemanding. It requires people to be disturbed by inconsistencies in their beliefs some of the time, and entails that people can to some extent distinguish between appropriate and inappropriate emotional responses. Note that this second statement is a restricted version of the claim that MI makes about the warrant of moral judgments. It is restricted because SR allows for a-rational and unwarranted intuitions of the kind SS is concerned with. SR, however, does more justice to the moral sensibilities people can have that provide them with input from the outside world. These sensibilities serve, for instance, to register harm, or to indicate how appropriate help would be.

The third thesis concerns the question how moral intuitions can confer warrant on moral judgments: (3) Moral intuitions can justify moral judgments non-inferentially. The relation between moral appearances and moral judgments is non-inferential by definition. What SR has in common with MI is that appearances can be warranted, and that such warrant can be transmitted to the beliefs they give rise to. SR differs

²⁹ See the introduction for other proposals that combine elements from sentimentalism and rationalism. Note 4 comments on how SR relates to Rational Sentimentalism.

from MI in how it conceives of the justificatory power of moral intuitions. In contrast to SS, SR does not regard affective responses as a-rational.

In section 1.2 I noted that proponents of MI typically regard moral intuitions as unmoved movers. SR rejects this view. Typically, the justification of a moral judgment that is caused by a moral intuition does not only depend directly on the warrant of that intuition, but also indirectly on past appearances and reasoning processes and on how they shaped the agent's moral sensibilities. As a consequence, it will depend in part on other intuitions and beliefs. These other justifying factors need not be present in order for the intuition to confer warrant on the relevant judgment. And the agent need not be able to recount them. The fact that intuitions have been shaped by normatively relevant factors themselves, however, reveals that they are not normatively foundational in the way they are often taken to be.³⁰

Another difference between SR and MI is that some intuitions are, as I will say, procedural rather than substantive. Substantive moral intuitions have cognitive content, or they are representational appearances that can give rise to cognitions. These are the intuitions MI is concerned with. SR acknowledges them and maintains that at least some of them are moral emotions. Moral emotions have affective and cognitive aspects 'that cannot be pulled apart' (Zagzebski 2003: 109). Some affective responses, however, do nothing else than indicate that the agent should be alert and put more effort in forming a particular judgment. The point of this is to draw attention to a possible distorting factor the influence of which should be mitigated if not eliminated. Such procedural intuitions have little or no cognitive content. They can, however, signal that the agent should exercise her moral sensibilities more carefully, or that she should consider more arguments for and against the belief she is inclined to form.³¹

At this point, the analogy between the cognitive dissonance mechanism and a car alarm mentioned towards the end of section 2.2 becomes relevant. Too many car alarms are too sensitive and go off in response to a loud noise or a gust of wind. This discourages people from taking them seriously. It is, however, possible to make a car alarm less sensitive. When the settings of an alarm are changed accordingly, false alarms become rare. Now it makes sense to pay attention to the alarm. In a somewhat similar way, one might say, someone's cognitive dissonance mechanism can be fine-tuned. This changes the agent's sensitivity to cognitive conflicts. Anticipatory guilt feelings alert the agent to possible flaws, but rarely when this is out of place and almost always when there is a genuine cognitive conflict. The better the mechanism is fine-tuned, the more justified the affective responses or moral intuitions are.

This idea can be developed in terms of Dual System theory (or more generally in terms of more or less conscious responses).³² Over time System II can fine-tune or calibrate System I. Consider Shy. Shy is very accommodating, and is used to accept

³⁰ Audi (2004) and Huemer (2005) maintain that intuitions sometimes have to be discarded due to other intuitions with which they are compared in an attempt to achieve reflective equilibrium. This resolves at least some of the worries I express in the main text.

³¹ Audi (2013: 155) acknowledges that emotions can contribute to the formation of reliable intuitions. However, as mentioned in note 16, he denies that affective responses can be intuitions.

³² See Craigie (2012) for another Dual Systems account of moral reasoning and moral judgment.

excuses from other people even if they are not particularly good. At some point she notices that this way of responding is not very fruitful. It seems that people just take advantage of her kind and considerate attitude. Meanwhile Shy talks with friends who confirm her diagnosis. Over time she becomes less timid and less compliant. She asserts herself more and begins to radiate more confidence. People start taking her more seriously and are less inclined to take advantage of her. This change of heart is facilitated by a change in sentiment. As a consequence, she no longer turns inward, but becomes annoyed when she notices others do not care about her feelings. In sum, Shy acquires different sensibilities.

Presumably Shy's initial shy responses are automatic and unreflective, which means that they are generated by System I. They may, of course, have been shaped by social norms and other people's expectations. However, at this point in time, Shy has never spent much time thinking them over. Over time she becomes more conscious of her strategy of responding and of how inadequate it is. Shy becomes particularly conscious of this when she discusses it with friends. In this way, System II gets involved. She starts asserting herself more, which initially requires conscious effort. For some time, System II regulates her behavior in the relevant circumstances. Over time, however, her new way of responding becomes habitual. In the end, no conscious thought is needed anymore and her newly acquired attitude influences her behavior automatically. Along the way, System I takes over responds without a need for deliberation.³³

During this process, one might say, system II calibrates system I. System I becomes more sensitive to inappropriate responses from other people, to responses that indicate that they are about to take advantage of Shy. She comes to recognize that this is not how it should be. She notices a discrepancy between how she is treated on the one hand, and how she should be treated according to her own norms on the other. This triggers an affective response, which in turn prompts her to consider alternative ways of reacting. She settles on a reaction that differs from what she used to do in settings like this one. Over time this new way of reacting becomes habitual. At some level, Shy is aware of her improved skills and learns to rely on them. Why think of this process as one of recalibration? System I becomes more sensitive to particular circumstances. Furthermore, System I starts associating the input she collects in those circumstances to different behavior. In other words, not only the input, but also the output that is linked to it changes.³⁴

Cognitive dissonance plays a central role in the process of change. When Shy becomes aware of a discrepancy between how she expects to be and how she believes that she should be treated, she becomes conscious of a cognitive conflict. This cognitive conflict triggers an affective response. As she is not the cause of the conflict, this response will not be one of guilt, which is the response involved in moral disengagement. Instead she is annoyed by if not angry with the other person.

³³ Fine makes a similar point when she argues that currently automatic responses might be influenced by prior conscious reflection or reasoning (2006: 85 and 93).

³⁴ This line of argument presupposes that each of the two systems has its good and its bad side, a view that is defended, for instance, by Kahneman (2011). The idea is that the heuristics of System I can be conducive to adequate responses in some circumstances, and lead to biased responses in other. They can be quick and dirty, but they can become more sophisticated over time.

This affective response in turn prompts a process of reasoning. The alternative courses of action that Shy considers differ in adequacy. She will be looking for actions that might prevent the other from taking advantage of her or that will discourage the other from doing so at future occasions. Shy might also take into account how her reaction affects her reputation. She stands to gain from an image of a more assertive person. Settling on a particular action can resolve the cognitive conflict just mentioned, and pave the way to performing that action. More often than not it will trigger the other to treat her properly. If, however, she is not treated properly, she will find an adequate response to it. Also in that case, the affect will dissolve. Note that, in order for System I to be recalibrated in this way, an agent typically has to undergo a number of such bouts of cognitive dissonance.

In order to connect Shy's story directly to the earlier discussion on SS, MI, and SR, it has to be retold in terms of intuitions. Although, in contrast to SS, SR does not assume that intuitions are always generated by System I, it maintains that this is often the case insofar as intuitions are concerned that are affective responses. Shy's intuitions will initially be fairly primitive affective responses. Over time, however, her responses and reactions become more attuned to her environment. For some time her new way of responding to her surroundings requires conscious effort and deliberation, which means that it involves System II. Due to practice it becomes ingrained. Once it has become automatic, Shy's new sensibilities have become part of System I. Along the way her intuitions acquire more and more justification, as her responses to the environment become more adequate. In contrast to SS, then, SR regards moral intuitions as potentially justified. Shy's judgments inherit any justification her intuitions have, as the process by which she forms her moral beliefs involves the cognitive dissonance mechanism that has been fine-tuned over time. Rather than being unmoved movers, affective dispositions can be recalibrated so as to increase their evidential significance.

3.2 Intuitions and Cognitive Dissonance

Shy's story illustrates how significant the role of cognitive dissonance can be in shaping people's intuitions. It exemplifies what I call 'the Cognitive Dissonance Model of moral reasoning'. Formulated in terms of Dual System Theory, this model has it that both Systems I and II play a role in generating reliable and justified intuitions. System II plays a role as a standby ready to step in when conscious reasoning is required. It is activated when System I registers an irregularity that requires special attention.³⁵ It might be that no such irregularities are detected because no complex situations are encountered or because System I is so well developed that it can adequately handle the complicated situations it encounters on its own. At the same time, the probability that it fails to alert system II when it is about to respond inadequately is very small. This is the ideal case of

³⁵ Craigie (2012: 67–68) defends a similar claim in terms of virtual control, a notion she borrows from Philip Pettit. See also Clarke (2008: 809).

the virtuous person who can exert his practical wisdom without much if any deliberation.³⁶

In more humdrum cases, System I will every so now and then notice certain features of the situation that deserve more attention at which point System II is engaged. In such a situation, conscious effort is required to come up with a more adequate response. Although it does not always do so, explicit deliberation can increase the degree to which the response and reaction is warranted. Note, however, that the fact that, when necessary, System II will be engaged reflects positively on System I responses. More precisely, the extent to which they are justified depends on the quality of its input/output settings. Suppose that System II functions properly as a perceptive standby that takes charge when it receives signals that indicate that System I functions in a way that is suboptimal, which could mean it is about to make a mistake. If this were indeed the case, it would confer further warrant on the agent's moral judgments.³⁷

How can an intuition provide justification? Intuitions are generated without the agent being conscious of any features that might justify the response. When they give rise to a judgment, no occurrent belief provides warrant for it. This does not mean it is not justified. Instead, it means that the justification it has is indirect. The intuition points in a certain direction and thereby functions as a signpost. The extent to which it is a good idea to travel in a particular direction depends on the quality of the signpost. As an intuition is only as good as the system that generates it, the quality of an intuition depends on the calibration of System I. More specifically, it depends on whether System I is appropriately sensitive to irregularities, and whether it provides the requisite signals to System II so it can step in when necessary.

Detecting an irregularity can initiate a process of cognitive dissonance. The affective response indicates that the envisaged reaction might be inadequate. This triggers a line of reasoning that may resolve the conflict. The dissolution of the emotion subsequently functions as a sign of coherence. Such processes of cognitive dissonance can improve the quality of the agent's moral judgments. In part by means of explicit reasoning, the agent's moral intuitions can become more sophisticated over time, and the warrant that moral intuitions confer on moral judgments can increase.

Intuitions, then, often function as indicators or proxies of reasons. Although initially they were of little use, at some point Shy's intuitions become reliable and she can treat them as defeasible sources of justification. They tell us in which direction we should think, or even which belief to form. And they can be relied upon in this respect because they transmit the warrant of the reasoning they are based upon. They will often involve emotions. However, this need not always be the case. A response can be so habituated that it is simply retrieved from memory without affective mediation (see also Prinz 2007). How do intuitive responses compare to conscious processes of reasoning? To some extent, intuitions or affective responses

³⁶ Sentimental Rationalism has affinities not only with Moral Intuitionism, but also with Aristotelian virtue ethics.

³⁷ See also Fine (2006: 90–92), who argues that controlled processes can interfere with automatic processes, and that prejudices can in principle be overcome in this way.

can be seen as a heuristic for conscious thought. The distress I experience when I hurt someone – let's suppose my elbow accidentally ends up in someone's stomach – and the empathic understanding I arrive at almost automatically point me in the same direction as the thought that I have harmed someone. I need to help the other or redress the pain I have caused. At this point, it might appear that Systems I and II are perfect substitutes. They do not need each other, and each can perform its task just as well as the other. The only difference is temporal. System I is much quicker than System II and functions as a shortcut for explicit reasoning. This conclusion, however, does not follow. The two systems specialize in different aspects of judgment formation and they need each other.

Earlier I argued that System I works well only if System II functions as a standby. The thing to see now is that System II needs input and can acquire it from System I. The quality of someone's affective responses depends in part on the agent's reasoning capacities, both when they are engaged in a particular case and when they play a role in the background. Intuitions sometimes function as a substitute for reasoning, which is useful in particular when a relatively unconscious affective response provides for quick input for action. I take moral judgments to express cognitive attitudes that can be based on moral reasoning. And even when they are based on intuitions instead, they can be insightfully reconstructed as the conclusions of processes of reasoning. In light of this, I refer to the position I defend as a kind of *rationalism*.

Sentiments do, however, play an indispensable role at least some of the time. Moral arguments can be complex and moral judgments often require balancing of reasons. Such balancing requires affective responses at least in some cases. And affect might be required to even begin to realize that something is wrong. Consider the example in which my elbow ends up hurting someone. Perhaps I am in a hurry because I am going to see a movie together with a friend. Presumably I have some duty towards my friend to be in time for the movie. How this should be weighed in the situation or how it is to be balanced against the harm the stranger incurs is not necessarily easy to determine. Perhaps mere cognition will not do the trick. Emotional distress and empathic understanding both with respect to my friend and to my victim may be required in order to determine the appropriate weight of certain considerations. The upshot is that affective capacities are indispensable. This is why the position I defend is called '*Sentimental Rationalism*'.³⁸

All in all, intuitions turn out to provide quick responses that can function as indicators for moral judgment. They give the agent a sense of which judgment is correct, but leave her unable to provide reasons for the judgment she makes. When she has developed her moral sensibilities, someone like Shy need not be worried by this, as she can rely on her intuitions. Conscious cognitions, however, are slow but propositional, and provide explicit justification.

SR portrays moral reasoning in a way that differs drastically from that of SS. In particular Haidt paints a bleak picture of private moral reasoning as biased and

³⁸ Sentimental Rationalism differs from the kind of rationalism defended by Greene et al. (2001; see also 2008) in that it does not regard reasoning as a rival of emotion and rejects the idea that reasoning is bound to be superior to emotion.

confabulated reasoning. In a sense, he argues that what is sometimes called ‘the game of giving and asking for reasons’ is mere play. On the basis of the dominant minority argument in combination with new evidence concerning moral disengagement, I have argued that this is often a serious game, not only when you play it with others, but also when you play it with yourself, and not only when you play it well, but also when your performance is mediocre.

4 Conclusion

Subjective Sentimentalism (SS) makes a mistake when it leaves no constructive role for reasoning to play. Moral Intuitionism (MI) is misguided when it depicts affect as no more than a source of distortion. The synthesis of the most plausible features of these two positions is Sentimental Rationalism (SR). SR ascribes a positive role to both affect and cognition. Both emotions and arguments can play a causal as well as a justificatory role with respect to moral judgment. Most of the arguments I presented against Subjective Sentimentalism (SS) or in favor of SR were based on empirical findings. I have presented *the argument from overgeneralization* on which the data do not support the sweeping claims that supporters of SS have defended on their basis. The second argument that I have presented is *the dominant minority argument*. This argument serves to appreciate that exceptions to generalizations can serve to point towards a mechanism on which moral judgments can be justified. On this alternative picture, both emotional responses and arguments can contribute to moral judgments together.

SR is a synthesis of SS and MI in that it transcends the oppositions that they harbor. Rather than either or, it says: both reason and emotion. Instead of always or never, it says: justified to a certain degree and only some of the time. Due to the fact that it has so much in common with its rivals, relatively small empirical peculiarities that can in some other cases safely be ignored turn out to make a big difference concerning which position is supported. Just as SS, SR recognizes that affective responses typically form the immediate basis of moral judgments. It does not, however, go so far as to claim that they always do. Similarly, just as MI, SR acknowledges that the rationality of moral agents contributes in important ways to the warrant they have for their moral judgments. It does, however, allow for other factors – in particular their affective moral sensibilities – to contribute as well.

Acknowledgments I thank Jochen Bojanowski, Joel Rickard, Sabine Roeser, Hanno Sauer, Markus Schlosser, and Peter Timmerman for helpful comments.

References

- J. Allman and J. Woodward, “What Are Moral Intuitions and Why Should We Care About Them? A Neurobiological Perspective,” *Philosophical Issues* 18 (2008): 164–85.
- K. Aquino and A. Reed, “The Self-Importance of Moral Identity,” *Journal of Personality and Social Psychology* 83 (2002): 1423–40.
- R. Audi, *The Good in the Right: A Theory of Intrinsic Value* (Princeton, Princeton University Press, 2004).

- R. Audi, *Moral Perception* (Princeton, Princeton University Press, 2013).
- A. Bandura, C. Barbaranelli, G.V. Caprara, and C. Pastorelli, "Mechanisms of Moral Disengagement in the Exercise of Moral Agency," *Journal of Personality and Social Psychology* 71 (1996): 364–74.
- A. Bandura, "Moral Disengagement in the Perpetration of Humanities," *Personality and Social Psychology Review* 3 (1999): 193–209.
- C.D. Batson, D. Kobrynowicz, J.L. Dinnerstein, H.C. Kampf, and A.D. Wilson, "In a Very Different Voice: Unmasking Moral Hypocrisy," *Journal of Personality and Social Psychology* 72 (1997): 1335–48.
- C.D. Batson, E.R. Thompson, G. Seufferling, H. Whitney, and J.A. Strongman, "Moral Hypocrisy: Appearing Moral to Oneself Without Being So," *Journal of Personality and Social Psychology* 77 (1999): 525–37.
- M.S. Bedke, "Intuitionist Epistemology in Ethics," *Philosophy Compass* 5 (2012): 1069–83.
- S. Chaiken, R. Giner-Sorolla, and S. Chen, "Beyond Accuracy: Defense and Impression Motives in Heuristic and Systematic Information Processing," in P.M. Gollwitzer and J.A. Bargh (eds.), *The Psychology of Action: Linking Cognition and Motivation to Behavior* (New York, Guilford Press, 1996), 553–78.
- S. Clarke, "SIM and the City: Rationalism and Psychology and Haidt's Account of Moral Judgment," *Philosophical Psychology* 21 (2008): 799–820.
- D. Copp, "Experiments, Intuitions, and Methodology in Moral and Political Theory," in R. Shafer Landau (ed.), *Oxford Studies in Metaethics: Volume 7* (Oxford, Oxford University Press, 2012), 1–36.
- R. Cowan (forthcoming) "Clarifying Ethical Intuitionism," *European Journal of Philosophy*, doi: [10.1111/ejop.12031](https://doi.org/10.1111/ejop.12031).
- J. Craigie, "Thinking and Feeling: Moral Deliberation in a Dual-Process Framework," *Philosophical Psychology* 24 (2012): 53–71.
- J. D'Arms and D. Jacobson, "Sentiment and Value," *Ethics* 118 (2000): 722–48.
- D. Davidson, "Actions, Reasons and Causes," *Journal of Philosophy* 60 (1963): 685–700.
- P.H. Ditto, D.A. Pizarro, and D. Tannenbaum, "Motivated Moral Reasoning," in D.M. Bartels, C.W. Baliman, L.J. Skitka, and D.L. Medin (eds.), *Psychology of Learning and Motivation*, vol. 50 (2009): 307–38.
- J.S.B.T. Evans, "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition," *Annual Review of Psychology* 59 (2008): 255–78.
- L. Festinger, *A Theory of Cognitive Dissonance* (Stanford, Stanford University Press, 1957).
- C. Fine, "Is the Emotional Dog Wagging Its Rational Tail or Chasing It?," *Philosophical Explorations* 9 (2006): 83–98.
- J.D. Greene, "The Secret Joke of Kant's Soul," in W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development* (Cambridge (MA), MIT Press, 2008), 35–79.
- J.D. Greene, R.B. Sommerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen, "An fMRI Investigation of Emotional Engagement in Moral Judgment," *Science* 293 (2001): 2105–08.
- J. Haidt, F. Björklund, and S. Murphy (2000) "Moral Dumbfounding: When Intuition Finds No Reason," *Unpublished Manuscript*.
- J. Haidt, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108 (2001): 814–34.
- J. Haidt, *The Righteous Mind: Why Good People Are Divided by Politics And Religion* (Pantheon Books, 2012).
- J. Haidt, S.H. Koller, and M.G. Dias, "Affect, Culture, and Morality, or Is it Wrong to Eat Your Dog?," *Attitudes and Social Cognition* 56 (1993): 613–28.
- M. Huemer *Ethical Intuitionism* (New York, Palgrave MacMillan, 2005).
- P. Johansson, L. Hall, S. Sikström, and A. Olsson, "Failure to Detect Mismatches Between Intention and Outcome in a Simple Decision Task," *Science* 310 (2005): 116–19.
- D. Kahneman, *Thinking Fast and Thinking Slow* (London, Allen Lane, 2011).
- N. Mazar, O. Amir, and D. Ariely, "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance," *Journal of Marketing Research* 45 (2008): 633–44.
- D. Moore, "Moral Disengagement in Processes of Organizational Corruption," *Journal of Business Ethics* 80 (2008): 129–39.
- G.E. Moore, *Principia Ethica* (Cambridge, Cambridge University Press, 1903).

- S. Nichols, *Sentimental Rules: On the Natural Foundations of Moral Judgment* (Oxford: Oxford University Press, 2004).
- R. Nisbett and T.D. Wilson, "Telling More than We can Know: Verbal Reports on Mental Processes," *Psychological Review* 84 (1977): 231–59.
- M.J. Osofsky, A. Bandura, and P.G. Zimbardo, "The Role of Moral Disengagement in the Execution Process," *Law Human Behavior* 29 (2005): 371–93.
- J. Prinz, "The Emotional Basis of Moral Judgments," *Philosophical Explorations* 9 (2006): 29–43.
- J. Prinz, *The Emotional Construction of Morals* (Oxford, Oxford University Press, 2007).
- J.M. Paxton, L. Ungar, and J.D. Greene, "Reflection and Reasoning in Moral Judgment," *Cognitive Science* 36 (2012): 163–77.
- T. Reid, *Essays on the Intellectual Powers of Man* (Cambridge (MA), MIT Press, 1969 [1785]).
- S. Roeser, *Moral Emotions and Intuitions* (Basingstoke, Palgrave Macmillan, 2011).
- W.D. Ross, *The Right and the Good* (Oxford, Oxford University Press, 1930).
- H. Sauer, "Educated Intuitions. Automaticity and Rationality in Moral Judgment," *Philosophical Explorations* 15 (2012a): 255–75.
- H. Sauer, "Psychopaths and Filthy Desks: Are Emotions Necessary and Sufficient for Moral Judgment?," *Ethical Theory and Moral Practice* 15 (2012b): 95–115.
- H. Sidgwick, *The Methods of Ethics* (London, MacMillan, 1874 [1907]).
- T. Wheatly and J. Haidt, "Hyponotic Disgust Makes Moral Judgments More Severe," *Psychological Science* 16 (2005): 780–84.
- T.D. Wilson, *Strangers to Ourselves: Discovering the Adaptive Unconscious* (Cambridge (MA), Belknap Press of Harvard University Press, 2002).
- L. Zagzebski, "Emotion and Moral Judgment," *Philosophy and Phenomenological Research* 66 (2003): 104–124.