

# The Limits of Emotion in Moral Judgment

Joshua May

Published in *The Many Moral Rationalisms*,  
Karen Jones & François Schroeter (eds.), Oxford University Press (2018), pp. 286-306.

**Abstract:** I argue that our best science supports the rationalist idea that, independent of reasoning, emotions aren't integral to moral judgment. There's ample evidence that ordinary moral cognition often involves conscious and unconscious reasoning about an action's outcomes and the agent's role in bringing them about. Emotions can aid in moral reasoning by, for example, drawing one's attention to such information. However, there is no compelling evidence for the decidedly sentimentalist claim that mere feelings are causally necessary or sufficient for making a moral judgment or for treating norms as distinctively moral. I conclude that, even if moral cognition is largely driven by automatic intuitions, these shouldn't be mistaken for emotions or their non-cognitive components. Non-cognitive elements in our psychology may be required for normal moral development and motivation but not necessarily for mature moral judgment.

**Key words:** sentimentalism, disgust, psychopathy, moral/conventional distinction, reason/emotion dichotomy

## 1. Introduction

As a psychological theory, moral *sentimentalism* is the view that “gives the emotions a constitutive role in evaluative judgment” (D’Arms & Jacobson 2014: 254) or at least that ultimately “moral judgment is grounded in affective response” (Nichols 2004: 83). The contrasting tradition of moral *rationalism* maintains that moral judgment fundamentally “derives from our rational capacities” (Kennett 2006: 70) or is ultimately “the culmination of a process of reasoning” (Maibom 2010: 999). According to rationalists, emotions are either merely the natural consequences of reasoning or provide just one way of instigating or facilitating reasoning.

To a great extent the debate between rationalists and sentimentalists is hostage to empirical research. Psychologists used to emphasize the role of reasoning and reflection in mature moral judgment, particularly in the tradition of Lawrence Kohlberg (1973) inspired by Kant. More recently, there has been something of an “affect revolution” in moral psychology, as Jonathan Haidt puts it (2003: 852). There is apparently converging scientific evidence that emotions play a foundational role in moral judgment. Jesse Prinz, for example, proclaims: “Current evidence favors the conclusion that ordinary moral judgments are emotional in nature” (2006: 30). Similarly, following Hume’s famous derogation of reason, Haidt concludes that “the emotions are in fact in charge of the temple of morality and that moral reasoning is really just a servant masquerading as the high priest” (2003: 852).

Against the sentimentalist orthodoxy, we’ll see that rationalism is well supported by our best science of morality. We’ll start by building a brief presumptive case in favor of rationalism by appealing to evidence that much of moral judgment involves reasoning that is sensitive primarily to the outcomes of an action and how involved an agent was in

bringing them about. We'll then reconsider the empirical support for sentimentalism. It will emerge that there is no compelling evidence that the affective component of emotions is causally necessary or sufficient for making a moral judgment or for treating norms as distinctively moral. Ultimately, while moral judgment is largely driven by automatic intuitions, these should not be mistaken for emotions, or at least not for their non-cognitive components.

## 2. Reason vs. Emotion

We need at least a working conception of reasoning versus emotional processing in order to determine their role in moral judgment. We'll see that there is no bright line dividing reason from emotion (Damasio 1994; Huebner 2015). So the division between rationalism and sentimentalism is likewise fuzzy at the edges (Maibom 2010). Both reasoning and emotion plausibly play some role in moral thinking (May and Kumar 2018). However, important questions remain about which is more important or fundamental and whether the cognitive, rather than the affective, elements of emotions are doing the explanatory work.

*Reasoning* is, roughly, a kind of inference in which new beliefs, or belief-like states, are formed on the basis of pre-existing ones that serve as reasons or evidence. Many psychologists define reasoning more narrowly as *conscious* inference (e.g. Haidt 2001: 818; Mercier & Sperber 2011: 57; Greene 2013: 136), but that's overly restrictive. Ordinary usage and dictionary definitions don't make conscious awareness essential to reasoning, and neither should we. The archetype of reasoning is indeed deliberate, relatively slow, and drawn out in a step-wise fashion. For example, you calculate your portion of the bill, weigh the pros and cons of divorce, or deliberate about where to eat for lunch.

However, reasoning can be a rather automatic, unconscious, implicit, and unreflective process. For example, you just find yourself concluding that your son is on drugs; you automatically infer from your boss's subtly unusual demeanor that she's about to fire you; or you suddenly realize in the shower the solution to a long-standing problem. These beliefs seem to pop into one's head but they aren't born of mere feelings or non-inferential associations. There is plausibly inference on the basis of representations that function as providing reasons for a new belief. Reasoning occurs; it's just largely outside of awareness and more rapid than conscious deliberation.

No doubt some of one's beliefs are formed without engaging in even unconscious reasoning. Basic perceptual beliefs are perhaps a good example. You believe that the door opening in front of you retains a rectangular shape, but arguably you don't form this judgment on the basis of beliefs about angles in your field of vision. Rather, your visual system generates such perceptual constancies by carrying out computational work among "subpersonal" or "subdoxastic" states (Stich 1978)—states that are relatively inaccessible to introspection and isolated from other patterns of belief-formation. As the visual experience of a rectangular door is generated, you believe that the door is rectangular by simply taking your visual experience at face value. In such cases it isn't appropriate to posit unconscious reasoning (about angles and the like) at least because the relevant transitions aren't among genuine beliefs.

Nevertheless, some inferential transitions between genuine beliefs are unconscious. So, within the category of unconscious mental processes, some generate beliefs on the basis of prior beliefs (e.g. inferring that your son is on drugs) but others do not (e.g. perceiving an opening door as constantly rectangular).

There is unfortunately even less consensus about the boundaries of emotions. Famously, there are staunch cognitivist theories on which emotions have cognitive content, much like (or even exactly like) beliefs. Martha Nussbaum, for example, argues that our emotions contain “judgments about important things” which involve “appraising an external object as salient for our own well-being” (2001: 19). Non-cognitivist theories maintain that emotions lack cognitive content. Prinz, for example, maintains that emotions are “somatic signals... not cognitive states” although they do “represent concerns” (2007: 68). We’ll do better to adopt a more neutral account that points to common features of emotional responses at least one of which is *feeling* or *affect*. Heidi Maibom provides a useful characterization of *emotions* as “mental states associated with feelings, bodily changes, action potentials, and evaluations of the environment” (2010: 1000; cf. Haidt 2003: 853). Such states can certainly facilitate reasoning, even good reasoning, although it’s often automatic and unconscious.

So at times it is artificial or fruitless to divide reason from emotion, but sometimes important contrasts can be made. Consider the brute fear one feels toward a large, hairy, but harmless spider. The affective response may involve information processing by providing representations of danger, but the feeling is not a state fit for inferential processes characteristic of reasoning. At any rate, if sentimentalism is to be distinguished from rationalism, moral judgment must be ultimately grounded in mere feelings or affect that do not constitute beliefs or merely facilitate reasoning (Prinz 2006: 31; Maibom 2010: 1000).

### **3. Dispassionate Moral Cognition**

#### *3.1 Unconscious Moral Reasoning*

It’s certainly a familiar idea that morality is codified in terms of norms, particularly those that transcend mere convention (see §4.1). The relevant rules, however, are not always consciously considered or accessible. Much like other forms of cognition (Reber 1989), moral judgment involves in part the unconscious application of abstract concepts or rules that identify relevant considerations. These can then be applied automatically with little conscious access to their content (Cushman et al 2006; Hauser et al 2007). Often what seem like emotionally-driven moral judgments are simply quick and automatic intuitions with complex computation occurring unconsciously. Indeed, there is experimental evidence that we easily track rule violations in moral dilemmas, even when the scenarios aren’t emotionally charged (Nichols & Mallon 2006). Just as Hume famously urged that “calm passions” can be mistaken for reason, tacit reasoning can be mistaken for passion.

One can perhaps draw an analogy with linguistic judgment, which involves automatic intuitions about the meaning and grammaticality of sentences (Dwyer 2009; Mikhail 2011). But we needn’t commit ourselves to the idea that moral cognition is largely universal or innate or that it arises from a module which is insensitive to conscious reasoning. There is, after all, evidence that moral norms vary substantially

across cultures (Prinz 2007) and that automatic moral intuitions can be shaped by experience through implicit learning mechanisms (Railton 2014) or overridden by deliberate reasoning (Paxton et al 2011).

Our concern is with mature moral judgment and to what extent it involves reasoning. Ample experimental evidence suggests that tacit rules in ordinary moral reasoning are sensitive to a range of factors that commonly feature in theoretical accounts of ethics, including: *good versus bad consequences*, *actions versus omissions*, and *intentional versus unintentional acts* (see May 2014b). The most prevalent method of detecting such unconscious considerations is to ask ordinary people to evaluate actions described in short vignettes that are experimentally varied in terms of only one factor. Differences in patterns of responses, if statistically significant, can help reveal what factors are driving moral intuitions.

Let's first consider studies on the role of consequences in ordinary moral cognition. It's quite obvious that good outcomes matter, and several experiments confirm this platitude. Famously, people tend to think it's morally acceptable to save five innocent people from certain death even if one will die as a mere side-effect (e.g. Mikhail 2011; Greene 2013). There is even some experimental evidence that this judgment is widespread across cultures and classes (Hauser et al 2007). Studies also reveal that consequences matter to us in more than just cases of death by trolley, such as moral dilemmas involving bodily injury, financial loss, emotional harm, and property damage (e.g. Gold et al 2013; Nichols & Mallon 2006). Even when moral rules seem to prohibit certain acts, such as violently killing an innocent person or engaging in incest, moral intuitions are apparently sensitive to the further norm that overall consequences also matter (Nichols & Mallon 2006; Rozman et al 2011). We're especially inclined to accept harming one person for the greater good when the alternative leaves no one better off, in line with the rule of "Pareto optimality" (Huebner et al 2011).

In ordinary moral thinking, though, good overall outcomes aren't only what matters. We certainly don't operate on the one utilitarian rule that one ought to do whatever maximizes happiness for everyone. This is at least evident from the highly robust result that ordinary people don't tend to regard bringing about a greater good as morally acceptable if it requires sacrificing someone in a violent, intentional, and personal way (May 2014b; Greene 2013). In the famous Footbridge case, for example, people from an array of cultures don't believe it's right to push one innocent person into the path of a train in order to save five others (see e.g. Hauser et al 2007; Mikhail 2011).

The non-consequentialist character of ordinary thinking is not merely sensitive to what's up-close and personal. Numerous experiments demonstrate, for example, the importance attached to outcomes arising from mere *omissions*, regardless of whether these are personal or impersonal. When presented with a wide range of scenarios, participants systematically rate them as more morally problematic when a negative outcome is brought about actively rather than through omission, such as killing an enemy versus letting him die (see e.g. Cushman et al 2006). In other words, ordinary moral thinking seems tacitly committed to something like the famous Doctrine of Doing and Allowing (May 2014b).

More importantly perhaps is whether a bad outcome was brought about *intentionally*. Even if someone dies as a result of her co-worker putting poison in her coffee, the act is much less likely to be condemned if it resulted from accidentally

mistaking the poison for sugar (see Young & Tsoi 2013). There's also some evidence that we ordinarily treat harming as a *means* as worse than harming as a side-effect (e.g. Hauser et al 2007; Mikhail 2011). A meta-analysis suggests the effect is weak and difficult to square with a traditional form of the famous Doctrine of Double Effect (Feltz & May 2017), but something close to it seems to shape ordinary moral judgment.

In sum, a growing body of research seems to reveal tacit reasoning in the formation of many ordinary moral judgments. We often rapidly infer the morally relevant properties of agents and actions (e.g. intentionality, outcomes) and then categorize the action as either moral or immoral on the basis of these patterns of implicit inference. Many of us, for example, seem to tacitly reason from something like the Doctrine of Doing and Allowing. It's unclear whether we have tacit beliefs in such basic moral principles (cf. Mikhail 2011: ch. 6) or are merely disposed to reason in accordance with them. Thankfully, we needn't decide between these two options since, like the rules of logic (e.g. Modus Ponens), one can count as reasoning in accordance with a moral principle even if it isn't represented in the content of a belief (Horgan & Timmons 2007: 280; Zimmerman 2013: 6). Ultimately, we have evidence of moral reasoning, regardless of whether the relevant principles are universal, innate, or mere dispositions to generate certain patterns of moral judgment.

### 3.2 Moral Cognition, Fast and Slow

The science is certainly revealing that a surprisingly large amount of mature moral judgment is driven by automatic and unconscious processes. But conscious moral reasoning is also possible. Indeed, there is ample evidence that moral judgment exhibits a *dual process* character like other areas of human cognition. Moral thinking can be quick, automatic, and largely unconscious, as we've seen when people evaluate actions that involve sacrificing one person for the greater good in a prototypically violent way. But moral reasoning can also be slow, deliberate, and more consciously accessible, as when one reflects over the course of a few days about whether to turn a good friend into the authorities for his criminal activity. We can see these two processes diverging in the laboratory in various ways. For example, experimenters can disrupt some kinds of moral judgment by increasing the strain on one's ability to consciously deliberate (e.g. Greene et al 2008).

Of course, some researchers contend that automatic moral intuitions are largely driven by emotional heuristics. Some neuroimaging studies suggest that these intuitions are correlated with areas of the brain independently associated with emotional processing (Greene 2013), and some changes in moral judgment seem to result after damage to these areas of the brain (more on this in §5). However, other brain imaging studies suggest that automatic moral intuitions are correlated instead with processing intuitive, as opposed to counter-intuitive, moral dilemmas (Kahane et al 2012). Moreover, brain imaging and lesion studies are rather coarse-grained tools for identifying emotions, and more fine-grained measures suggest there isn't a strong correlation between automatic moral judgments and the experience of emotions (Horne and Powell 2016). Regardless, though, rationalists needn't reject the idea that some moral intuitions are correlated with emotion.

What about evidence that conscious reasoning doesn't cause moral judgment but is rather mere post hoc rationalization of automatic moral intuitions? People do

sometimes evaluate moral dilemmas immediately and stick to their judgment even if they can't identify a satisfactory reason. This phenomenon of "moral dumbfounding" is evident when people evaluate apparently harmless taboo violations, such as eating one's pet dog that has died or passionately kissing one's sibling (Haidt et al 1993). Participants often stick to the condemnation of such acts without being able to identify a compelling reason, even to their own satisfaction.

However, at best such studies reveal only that moral judgments are sometimes arrived at via *unconscious* reasoning while conscious reasoning is recruited to uncover the rationale (Mallon & Nichols 2010). Indeed, it's unclear whether the taboo violations in the moral dumbfounding studies are truly perceived as harmless (Jacobson 2012; Royzman et al 2015). Either way, the inability to properly articulate unconscious reasoning is familiar from non-moral cognition too, as when one makes judgments about the appropriateness of sentences in one's native language (Dwyer 2009). Even if moral and linguistic cognition differ in important ways, conscious reasoning is not precluded from playing a causal role in some moral judgments (see May 2018).

Ultimately, the scientific literature is converging on the idea that automatic and controlled processes *interact* to generate mature moral judgment. Intuitive reactions can flexibly, even if implicitly, shape deliberative reasoning (Railton 2014), and conscious reasoning can automate later reactions, making them intuitive (Kennett & Fine 2009). Moreover, reflective deliberation can monitor and guide intuitive reactions as they occur (Craigie 2011). This is perhaps most evident in ordinary experience when we realize that some intuitive moral judgments are inconsistent, such as the belief that dog fighting is immoral versus the belief that factory farming is morally acceptable (Campbell & Kumar 2012). So, whether moral cognition is fast or slow, it often appears to involve a complex interplay between implicit and explicit reasoning.

We now have a presumptive empirical case in favor of rationalism. One might now ask: Must moral cognition involve reasoning all the way down? Won't reasoning bottom out in some basic moral beliefs, such as "It's wrong to fail to maximize happiness" or "Treat others as you'd like to be treated"? Sentimentalists, after all, can admit that reasoning plays a role in moral cognition, just not at its foundations. Sentimentalists would no doubt claim that our basic moral beliefs must be grounded in emotions or dispositions to feel them (Prinz 2007: 24-6) or that moral reasoning ultimately assesses the fittingness of emotional responses, which remain constitutive of moral judgment (D'Arms & Jacobson 2014). Rationalists could simply embrace a conception of moral cognition without foundations. (Perhaps we simply continue to reason to the most internally coherent system of moral beliefs or we continue to justify our moral beliefs without end.) Rationalists who prefer a more foundationalist framework can simply deny that basic moral beliefs must be backed by affect. Either way, the sentimentalist alternative is compelling only if emotions are essential to distinctively moral judgment, independently of reasoning processes. We'll see that the empirical case is unconvincing.

#### **4. Moralizing with Emotions**

Let's first consider evidence that emotions alone can substantially influence moral judgment. There are certainly studies showing that moral judgments are *correlated* with emotions (Moll et al 2005; Haidt et al 1993), but that is no evidence in favor of sentimentalism. Since we care deeply about moral issues, rationalists can happily accommodate emotions being a consequence of moral judgment (Prinz 2006: 31; Huebner et al 2009). Ideally sentimentalists would be able to show that simply feeling powerful emotions can make us *moralize* an action—take us from one moral judgment to another.

To establish this, we must be able to disentangle emotions from any morally relevant thoughts or beliefs that come along for the ride. Many emotions, such as anger and compassion, are intimately bound up with beliefs about the morally relevant facts. Being angry that someone stole my wallet is too closely bound up with my thoughts about the lack of justification for the perpetrator's action. Was it the raw feeling that made me judge it wrong or the belief that it was harmful and unwarranted? It's difficult to disentangle these experimentally. Such emotions aren't "pre-normative" (to borrow a term from Royzman et al 2009). Sentimentalists have thus been drawn to the explosion of research on disgust in which this feeling is *incidental* or unrelated to the action being evaluated. Being grossed out by someone using the bathroom, for example, can be divorced from morally relevant thoughts about embezzlement. I am thus in agreement with psychologists Justin Landy and Geoff Goodwin that "studies of incidental disgust arguably provide the best current evidence for the neo-sentimentalist idea that moral judgment is largely driven by emotion" (2015: 522).

#### 4.1 Moralizing Conventions with Disgust

One prominent sentimentalist strategy is to establish that emotions are essential to moralizing or to making distinctively *moral* judgments. For this to work, we need a characterization of the concept of morality or some core aspect of it.

One mark of moral norms is that they appear to be distinct from mere conventions. The norms of etiquette require that I not spit in my napkin at the dinner table, and school rules dictate that children wear a certain uniform. These are typically regarded as conventions that can be contrasted with moral norms, such as those that proscribe physically harming innocent people or invading someone's privacy. Wearing pajamas to school seems less serious and more acceptable if an authority deems it permissible (e.g. the school rules have changed). Moreover, explanations for why one shouldn't violate a convention are less likely to point to paradigm moral considerations, like harm, fairness, or rights.

A large body of empirical evidence confirms this. In general, compared to moral transgressions, we treat violations of conventions as *less serious, more permissible, contingent on authority, valid more locally* than universally, and involving *distinct justifications* that don't primarily appeal to another's welfare or rights (Turiel 1983). This distinction between types of norms appears to develop quite early—around age 4—and appears to be universal across many cultures, religions, and social classes (Nucci 2001).

Drawing heavily on this research, Shaun Nichols (2004) has argued that what makes us moralize a norm is that it's backed by strong feelings or affect. While rules or norms are essential to moral judgment, they aren't sufficient, for they may be

conventional, not moral. What makes a judgment *moral* has to do with our emotions toward the norm that has been violated (or upheld, presumably).

The key test of this “sentimental rules account” comes from a clever study in which Nichols (2002) sought to demonstrate that people would moralize the violation of a convention if they were especially disgusted by it (e.g. a person snorting and spitting into his napkin at dinner). In the first experiment, Nichols found evidence that participants would treat transgressions that are disgusting as more like moral transgressions compared to violations of emotionally neutral conventions. The people in his small sample were inclined to rate the disgusting transgressions as slightly more serious, less permissible, and less authority contingent (while justifications varied). In the second experiment, Nichols divided participants up into those that are highly disgust-sensitive, based on their score on a disgust scale, previously validated by other researchers. Participants especially sensitive to disgust tended to treat disgusting transgressions as less conventional, compared to the other group. However, while disgust-sensitive participants rated the disgusting transgressions as more serious and less authority contingent, there was no difference between the groups’ *permissibility* ratings (2002: 231).

Does this provide strong evidence that emotions alone can moralize? There are several reasons for doubt. First, disgust was not manipulated in either experiment, and in the second disgust was merely identified as likely to be more intense in a certain group. We can’t be sure that the different responses these groups provided were merely due to differing levels of disgust experienced, rather than another factor. Second, permissibility ratings are arguably the core of moral judgment, yet there was no difference among those participants who were especially disgust-sensitive. While these participants did rate disgusting transgressions as more serious and less authority contingent, this is a far cry from moralizing. It is interesting that elevated disgust seems to correspond to treating a transgression as less authority contingent. However, third, Nichols did not directly measure whether more disgusting violations strike people as involving more psychological harm, which fails to pry the emotion apart from a morally relevant belief. Follow-up studies by Royzman and his collaborators (2009) suggest that perception of harm accounts for some of the moralization of disgusting transgressions. Moreover, with a much larger sample size Royzman and colleagues were not able to replicate Nichols’s original result when the disgust scale was administered two weeks prior to soliciting moral reactions to the hypothetical transgressions. With this improved design, participants were less likely to be aware of the hypothesis being tested or to have their assessments of the transgressions influence their responses on the disgust scale.

A broader problem is that it’s unclear whether characteristic features of the moral/conventional distinction are an appropriate measure of moralizing. Daniel Kelly and his collaborators (2007) had participants evaluate a broader range of harmful actions than the usual “school yard” transgressions found in work on the moral/conventional distinction. The results provide some evidence that not all violations of moral rules yield the signature pattern of responses. For example, most of their participants thought that it’s very bad to train people in the military using physical abuse but only if government policy prohibits it. So the norm is apparently regarded as a moral one even though its status is authority-dependent.



While there may be concerns about some aspects of the study conducted by Kelly and colleagues (Kumar 2015), there are good theoretical reasons for expecting such data. As Heidi Maibom (2005: 249) points out, many norms that would be dubbed mere “conventions” often seem moral. For example, if I speak without the talking stick in hand, then I’ve violated a rule that’s not very serious, not exactly highly impermissible, and dependent on an authority who set the rule. If the councilor says anyone can talk, with or without the stick, then there’s no transgression. Nevertheless, when the rule is in place, speaking without the stick is rude, and many would deem it morally wrong.

In sum, it doesn’t seem sentimentalists can find strong support in research on disgust and the moral/conventional distinction. The distinction is a valuable heuristic for distinguishing many moral rules from non-moral ones, perhaps even as a way of characterizing the essence of a norm’s being moral (Kumar 2015). But it’s unclear in this context whether one group of people count as moralizing a norm just because they treat a transgression as slightly less conventional than another group does. More importantly, even if treating a rule as slightly less conventional suffices for moralization, we lack solid evidence that this is driven by raw emotions, like feelings of disgust, rather than tacit thoughts about increased psychological harm.

#### *4.2 Amplifying with Incidental Emotions*

A better route to sentimentalism appeals to research that manipulates emotions and directly measures moral judgment. Dozens of studies have now manipulated incidental disgust and measured its effects on moral opinions about hypothetical scenarios. Many philosophers and scientists champion these experiments as vindicating the role of emotions in either all of moral judgment (e.g. Haidt 2001; Prinz 2007; Chapman & Anderson 2013; Sinhababu 2017) or large swaths of ordinary moral thinking (e.g. Nado, Kelly, & Stich 2009; Horberg, Oveis, & Keltner 2011; Kelly 2011; Plakias 2013; Greene 2013).

Unfortunately, however, the evidence again underwhelms. Rather than support sentimentalism, the studies suggest that incidental emotions hardly influence moral judgment and are instead often a mere consequence. But let’s first consider some of the key evidence.

The disgust experiments all involve having participants provide their moral opinions about hypothetical scenarios described in brief vignettes. Some participants are randomly assigned to a control group who aren’t induced to feel heightened levels of disgust before evaluating the vignettes. Those in the manipulation group, however, have this emotion elevated in various ways, such as being hypnotized to feel disgust upon reading a certain word (Wheatley & Haidt 2005). Other experiments induce disgust by having participants sit at a dirty desk with remnants of food and sticky substances; smell a foul odor; watch a gruesome film clip involving human feces; or recall a disgusting experience (Schnall et al 2008). Still other researchers had some participants drink a bitter beverage, as opposed to water or something sweet (Eskine, Kacirik, & Prinz 2011). In all of these studies, and some more, incidental disgust alone tended to make moral judgments harsher. If this effect is real, widespread, and substantial, then this is powerful evidence in favor of sentimentalism.

However, these experiments are rather limited, for many reasons (see May 2014a). First, many of the effects were found only among certain types of people or *subgroups* of the sample. Subjects in Wheatley and Haidt's (2005) experiments were only those who were "highly hypnotizable." Similarly, Schnall and her collaborators (2008) primarily found the disgust effect only among those who were especially aware of their own bodily feelings (scoring high on a Private Body Consciousness scale). Second, while participants respond to many vignettes, the disgust effect was only detected among a minority of them. So the effects on moral judgment are *scarce*. Third, even when detected, the effect is rather *small* (an issue briefly noticed by others as well, such as Mallon & Nichols 2010: 317–8; Pizarro, Inbar, and Helion 2011). For example, in one of Wheatley and Haidt's (2005) vignettes, which described an act of bribery, the average morality rating was 78.73 for the control group and 91.28 for the disgust group (using a scale where 0 is "not at all morally wrong" and 100 is "extremely morally wrong"). This mean difference between the groups is statistically significant, but that only means roughly that it was not likely due to chance. More precisely, the probability is rather low (less than .05) that we'd observe this difference in a sample even assuming there's no real difference in the population. Despite this, the difference between groups still seems rather small and fails to shift the *valence* (or polarity) of the moral judgment. Disgusted or not, both groups tend to agree about whether the hypothetical action was right or wrong.

These problems have been buttressed by a recent meta-analysis of the effect of incidental disgust on moral cognition. Landy and Goodwin (2015) combed the literature for published studies and collected numerous unpublished ones, yielding a total of 51 experiments and over 5,000 participants. Using Cohen's standard, the effect size across all experiments was officially "small" ( $d = 0.11$ ). Moreover, the effect disappears when one considers only unpublished experiments, which suggests a bias against publishing the null results or replication failures. The mainstream and underground studies cleave on this point: "the published literature suggests a reliable, though small, effect, whereas the unpublished literature suggests no effect" (2015: 528). Given publication bias and possible confounds, Landy and Goodwin conclude that disgust's amplification effect on moral cognition is extremely small at best, perhaps nonexistent.

While disgust has received the most attention, some researchers have manipulated other incidental emotions, often by piping instrumental music through headphones. For example, Seidel and Prinz (2013a) manipulated incidental anger by having participants listen to irritating "noise music." In another experiment, they induced positive feelings of happiness with uplifting classical music (Seidel & Prinz 2013b). The key results in these studies were that incidental anger slightly amplified condemnation of autonomy violations and happiness slightly amplified judgments of praise and moral obligation (while anger reduced such judgments).

Do these few additional studies demonstrate the power of incidental emotions in moral judgment? One worry is that certain noises, particularly irritating ones, could significantly distract participants from fully processing morally relevant information in the vignettes. More importantly, though, all of the findings are similar to those of the disgust experiments. While the effects weren't restricted to subgroups in the samples, and sometimes the effects were found for all or most vignettes tested (not just a minority), the differences between groups are again consistently small shifts on a fine-grained scale. Now, in these studies, the emotion does sometimes shift the valence of the moral

judgment on average compared to controls (Prinz 2016: §1.2.1). But the shift is consistently from roughly the midpoint (the mean response in the control group) to slightly beyond (in the relevant manipulation group). So these few studies from one lab don't provide sufficient evidence to conclude that incidental emotions other than disgust have more than a negligible effect on moral judgment. Further research, replications, and meta-analyses are required before we can confidently conclude that the effects are stable and substantial enough to support sentimentalism over rationalism.

Of course, if there is an amplification effect, even if small, rationalists owe an explanation of why this would happen (Prinz 2016: §1.2.1). But that burden can be discharged in a number of ways (see May 2016; May 2018).

## 5. Psychopathology

We have seen that various popular experiments fail to show that incidental emotions play an integral role in mature moral judgment. Other avenues of support come primarily from psychopathology. By studying when moral judgment breaks down, we can uncover whether an emotional deficit best explains the problem.

### 5.1 Psychopathy

One route involves studying psychopaths, since their psychological profile suggests profound moral incompetence due to emotional deficits. This seems to speak in favor of the sentimentalist's claim that emotions are essential for moral cognition (Nichols 2004; Prinz 2007).

Not all psychopaths are alike, but they are typically characterized as callous, lacking in remorse and guilt, manipulative, having a superficial charm, impulsive, irresponsible, and possessing a grandiose sense of self-worth (Hare 1993). Most studied are incarcerated men and many have committed violent crimes or engaged in reckless actions that leave innocent people destitute. Psychopathy is similar, but not exactly equivalent, to antisocial personality disorder in the Diagnostic and Statistical Manual of Mental Disorders. Researchers instead typically diagnose psychopaths using Robert Hare's Psychopathy Checklist Revised (PCL-R), which has a range of criteria pertaining to the individual's psychological traits and history of past infractions.

Psychopaths don't just behave badly; some research suggests they don't properly understand moral concepts and thus can't make genuine moral judgments. Some theorists point to their poor use of moral terms, as when some psychopaths don't appear to properly understand what it means to *regret* hurting someone (Kennett & Fine 2008). More striking is the apparent failure to distinguish moral norms (e.g. don't hit another student) from mere conventions (e.g. don't wear pajamas to school), which as we've seen many theorists believe is necessary for a proper grasp of morality. In particular, some research on adult incarcerated psychopaths suggests that they treat conventions like moral rules by categorizing them as just as serious and independent of authority (Blair 1995). One hypothesis is that such inmates incorrectly categorize conventional norms as moral in a futile attempt to show that they know it's *wrong* to violate most norms.

Other evidence, however, suggests that psychopaths do not have such a substantial deficit in moral judgment. One study attempted to remove the motivation to treat all transgressions as serious by telling inmates that the community regards only half

of the transgressions as moral violations. Yet the researchers found that a higher score on the PCL-R does *not* correlate with less accurate categorization of the norms as moral versus conventional (Aharoni et al 2012). However, the researchers in this later study did find that two sub-factors of the PCL-R (affective deficits and anti-social traits) correlate with diminished ability to accurately categorize transgressions. Other research focuses on patterns of moral judgments about scenarios in which a person can harm one for the greater good in either a personal way (e.g. pushing) or an impersonal manner (e.g. remotely pulling a lever). Like ordinary individuals and non-psychopathic delinquents, incarcerated and non-incarcerated psychopaths tend to categorize personal harms as morally problematic even if they can bring about a greater good (Cima et al 2010). Moreover, Andrea Glenn and her colleagues observed no difference in non-incarcerated psychopaths' moral judgments about the personal versus impersonal dilemmas (Glenn et al 2009), despite finding that they exhibit reduced activity in the amygdala and increased activity in a brain area associated with abstract reasoning (the dorsolateral prefrontal cortex). One study, however, did detect a difference in the responses to personal dilemmas from incarcerated psychopaths compared to non-psychopathic inmates (Koenigs et al 2012).

Thus, while there's some evidence of impaired moral cognition in psychopaths, the evidence is somewhat mixed. There is most likely some deficit in their grasp and deployment of moral concepts, but the extent of it is unclear. No doubt much of a criminal psychopath's behavior can be explained by a lack of concern for others, even if knowledge of right and wrong is roughly intact (Cima et al 2010). As Glenn and her collaborators put it: "Emotional processes that are impaired in psychopathy may have their most critical role in motivating morally relevant behavior once a judgment has been made" (2009: 910). Such a conclusion may require admitting the possibility of making a moral judgment while lacking motivation to act in accordance with it. But rationalists can happily accept that the connection between moral judgment and motivation breaks down when one isn't being fully rational (Smith 1994: ch. 3).

So what in psychopaths *explains* their (slightly) impaired capacity for moral cognition? The most popular account points primarily to emotional deficits, based on various studies of the behavioral responses and the brain activity of psychopaths and people with psychopathic tendencies. For example, key brain areas of dysfunction appear to be the amygdala and ventromedial prefrontal cortex (VMPFC), both of which are implicated in processing emotion, among other things (Blair 2007). Moreover, as already noted, the sub-factors in psychopathy that have been correlated with diminished ability to draw the moral/conventional distinction involve affective deficits (e.g. lack of guilt, empathy, and remorse) and anti-social tendencies (Aharoni et al 2012). Further evidence comes from the studies which indicate that, compared to normal individuals, when psychopaths judge moral dilemmas they exhibit decreased activation in the amygdala, which is involved in emotional responses, especially fear, and emotional memory (Glenn et al 2009).

The idea that psychopathy primarily involves an emotional deficit seems bolstered when compared to autism (Nichols 2004). On many accounts, autism involves a difficulty in, not compassion, but the ability to understand the emotions and concerns of others. People with autism can be in some sense "anti-social" but they aren't particularly aggressive or immoral. Moreover, many seem competent with moral concepts, at least

because high-functioning children with autism seem to draw the moral/conventional distinction (Blair 1996). Some argue that autism, especially when severe, can impair moral judgment by limiting the understanding of others' projects, concerns, and emotional attachments (Kennett 2002). However, even if autism does impair moral cognition in this way, the deficit seems less severe compared to psychopathy.

There is some evidence, however, that psychopaths have profound deficits that are arguably to their *rational capacities*. Notoriously, they are impulsive, disorganized, easily distracted, delusional about their own abilities and importance, and poor learners in the face of negative reinforcement (see e.g. Hare 1993). Moreover, a meta-analysis of 20 studies revealed that individuals with psychopathy (and related anti-social personality disorders) have difficulty detecting sad and fearful facial expressions in others (Marsh & Blair 2008). Such deficits can certainly impair one's reasoning about both morality and prudence at least by preventing one from properly assessing the merits of various choices and resolving conflicts among them (cf. Kennett 2002; 2006; Maibom 2005). One psychopath, for example, tells the story of breaking into a house when an old man unexpectedly appears, screaming about the burglary. Annoyed that the resident wouldn't "shut up," this psychopath apparently beat the man into submission, then laid down to rest and was later awoken by the police (Hare 1993: 91).

Ultimately, there are several issues to highlight that work in concert to neutralize the threat to rationalism from psychopathy. First, as we've seen, psychopaths have significant *motivational* (not just cognitive) impairments, such as impulsivity. Emotions may aid in motivation and behavior but that isn't in conflict with the rationalist's claim about moral *judgment* and needn't conflict with a rationalist account of all aspects of our moral psychology (Huebner et al 2009; May 2018). There is admittedly good reason to believe that psychopaths are not *fully* competent with moral concepts and judgment. But the second issue is that this isn't merely due to an emotional deficit. Our best evidence suggests that much of the problem with psychopathy is *rational* or cognitive, such as grandiosity and poor processing of others' emotions. Third, and perhaps most important, since psychopathy is a developmental disorder, their profile doesn't necessarily establish that emotions are important for mature moral judgment. In fact, there is some evidence that damage to the VMPFC later in life primarily impairs motivation, not moral judgment (Roskies 2003). At best, then, psychopathy could help sentimentalists show that emotions are *developmentally* necessary for acquiring full competence with moral concepts and judgment (Prinz 2006; Kennett 2006). Yet we've seen that even this is difficult to establish convincingly.

Psychopathy does place a burden on rationalists to explain why emotions would be necessary for normal moral development but not the basic capacity for mature moral cognition. Here is one answer. Like all development, moral development begins early in one's life before one develops many concepts and the abilities to speak, read, write, and engage in complex reasoning. Yet moral development must go on. Emotions typically are a reliable resource for sparking and guiding one's thoughts and actions, prior to the development of more sophisticated abilities. We are creatures with limited attentional resources in environments with more information than we can take in. The predicament is even more dire when we're young and developing key social concepts, associations, and habits. Our default is not generally to pay attention to everything, for that is impossible. Instead, we rely on our attention being directed in the right places based on mechanisms

that are quick, rigid, automatic, often emotionally-driven, and plausibly fitness-enhancing. In short, emotions are required early in development to facilitate inference. Absent relevant automatic emotional responses, we can expect a person's moral judgment and behavior to be compromised. When the missing or attenuated emotions are the likes of compassion and guilt, and those problems are combined with rational deficits—such as delusions of grandeur, poor attention span, and poor recognition of emotions in others—we can surely expect a callous, manipulative, and aggressive person who has a limited grasp of the moral domain. Even if psychopaths do possess moral concepts, it is no surprise that their moral judgment and behavior are impaired.

In sum: Our best scientific evidence does suggest that psychopaths lack normal moral competence, but the deficit in moral cognition is often overstated in comparison to motivational deficiencies. Moreover, while psychopathy plausibly involves some emotional dysfunction, especially in moral emotions like guilt and compassion, the condition involves at least an equal impairment in reasoning. A lifetime of impulsivity, delusions of grandeur, poor attention span, diminished sensitivity to punishment, and so on can alone explain diminished competence with moral concepts, anti-social behavior, and lack of moral motivation. Finally, the lesson from psychopathy is largely developmental, providing little information about the mechanisms of mature moral judgment. This all suggests a rationalist model of psychopathy on which the psychopath's anti-social behavior is explained partly in terms of emotional and non-emotional deficits but with only the latter chiefly explaining any abnormalities in moral judgment.

## *5.2 Lesion Studies*

When patients suffer brain damage, we can correlate differences in moral responses with the dysfunctional brain areas and their usual psychological functions. Ideally, we can then determine if emotions play an important role in moral judgment. There are two key brain lesions that have been studied in relation to moral judgment: damage to the VMPFC and neurodegeneration in the frontal or temporal lobes.

Patients with lesions of VMPFC typically develop what Antonio Damasio (1994) has called “acquired sociopathy.” The famous Phineas Gage is just one example: after a rod accidentally passed through his skull, a once upstanding Gage reportedly became more crass, had difficulty keeping jobs, and so forth, despite apparently maintaining his level of general intelligence. There is some controversy about the various details of Gage's story, but now such patients are better documented.

Acquired sociopathy has some similarities with psychopathy, at least in that both involve abnormal function in the VMPFC (Blair 2007), but the two conditions are rather different. For one, since psychopathy is a developmental disorder, the brain dysfunction has much more serious effects. Adults who suffer damage to the VMPFC have already developed mature moral judgment, and there is some evidence that it remains relatively intact (again, see Roskies 2003). The disorder is instead primarily a problem with motivation or decision-making. Those with acquired sociopathy can know or cognize the various options for a given decision but do not have the proper guidance from their gut feelings. They lack what Damasio calls “somatic markers” or bodily feedback that guides decision-making. Diminished or missing somatic markers can leave patients prone to make imprudent and morally questionable choices, but unlike psychopaths they're not characteristically manipulative, violent, or grandiose.

The brain's VMPFC certainly seems crucial for decision-making that is guided by affective feedback, but is there evidence of impaired moral judgment or just motivation? Despite some evidence of normal moral cognition, some studies suggest that VMPFC damage does yield abnormal processing of scenarios involving personal harm for the greater good. Such patients seem to be more inclined to provide the abnormal "utilitarian" judgment that one should sacrifice an innocent individual for the sake of saving a greater number of other innocents (Koenigs et al 2007; Ciaramelli et al 2007).

A similar phenomenon arises in people who have related brain abnormalities. Patients with frontotemporal dementia (FTD) can have a wide variety of symptoms, including overeating and poor hygiene, since their neurodegeneration can occur in two out of the four lobes of cortex. But some common symptoms include blunted emotions and antisocial behavior, which are typical among those with lesions of the VMPFC. Importantly for our purposes, when presented with moral dilemmas requiring personal harm, FTD patients like VMPFC patients provide more "utilitarian" moral judgments than controls (Mendez et al 2005).

Even if we take these lesion studies at face value, they don't show that emotions are essential for moral judgment. Rather, as Greene (2013) would put it, this is evidence that emotions are more important for making *certain kinds* of moral judgments, namely those that condemn personal harm for the greater good. The lesion studies actually support the rationalist idea that moral cognition can proceed even with blunted emotions. The resulting moral judgments may be somewhat abnormal, but utilitarians like Greene would have us believe that these patients happen to be morally correct! As with psychopathy, emotional deficits alone don't reveal a substantial impairment in the capacity for moral cognition. Gut feelings help to guide decision-making across many domains; they don't serve as a mark of distinctively *moral* cognition.

Now, much like the experiments manipulating incidental emotions, one might argue that the lesion studies provide an alternative way of showing that damage to emotional centers in the brain at least leads to *different* moral judgments. Even if patients retain the general capacity for moral judgment, their emotional deficiencies seem to lead to some change in moral cognition. The problem of course is that associating areas of the brain with emotional processing is far from isolating the feelings from the thoughts associated with them. Blunted emotions may simply hinder the patient's ability to pay close enough attention to morally relevant information (Huebner et al 2009) or hinder other reasoning capacities.

## 6. Conclusion

Based on the science, many believe mature moral judgment crucially involves feelings. The evidence for this sentimentalist conclusion has been diverse but generally rather weak and overblown. First, while the moral/conventional distinction may partly characterize the essence of moral judgment, moral norms needn't be backed by affect in order to transcend convention. Second, priming people with incidental feelings like disgust doesn't make them moralize actions. Third, moral judgment can only be somewhat impaired by damage to areas of the brain that are generally associated with affective processing. Psychopaths, for example, exhibit both emotional and inferential

deficits, and the latter alone can explain any minor defects in the psychopath's ability to distinguish moral from conventional norms. The greatest problem in psychopathy appears to be motivational or behavioral, and emotional deficits do much more explanatory work here.

Additionally, a growing body of research fits with the rationalist idea that both conscious and unconscious reasoning heavily influence moral cognition. This reasoning is at least systematically sensitive to an action's outcomes and the agent's role in generating them. While these considerations are often processed unconsciously, we have little reason to believe that conscious reasoning is predominantly post hoc rationalization. Our dual process minds afford both kinds of influence on moral thinking. Even if unconscious processes are much more prominent than we ordinarily expect, they're often driven by complex computations, not simply blind affect.

Sentimentalists might claim that emotions typically drive the unconscious application of moral rules (e.g. Prinz 2007). But this seems unnecessary given that we lack compelling evidence that emotions, independent of reasoning, play an integral role in moral cognition. If affect alone doesn't moralize or underwrite non-pathological moral judgment, then why think it's causally necessary for unconscious moral reasoning? Moreover, emotions—or even dispositions to feel them—are not always evoked by many of the dilemmas used to uncover tacit moral computations (e.g. when a hypothetical person's property is damaged). So there often aren't any emotions or sentiments one can plausibly appeal to in the first place.

There is certainly some role for emotions in moral cognition, but we should recognize their limits. We lack evidence that their *affective* elements play a distinctive role independent of reasoning. Emotions typically aid reasoning by, for example, drawing one's attention to relevant information (Pizarro 2000; cf. Prinz 2006: 31). Deficits in emotion then might only indicate poor input to an intact capacity for reasoned moral cognition (Huebner et al 2009). So the non-cognitive elements of emotions appear to be ultimately inessential insofar as they don't contribute to the generation of beliefs fit for inference. Rationalists should admit that moral cognition is often automatic and unreflective (contra Kennett & Fine 2009) but they needn't concede that mere feelings are necessary (contra Sauer 2012). A sentimentalist could water down the characterization of emotions so much that they are conceived simply as the states of unconscious reasoning, but that would effectively abandon the sentimentalist tradition.

**Acknowledgements:** Versions of this paper were presented at the Tennessee Value and Agency Conference, the Tennessee Philosophical Association, and the annual Philosophy Symposium at Cal State Fullerton. For feedback, I'm grateful to the audiences at these events and in particular to Nageen Jalali, Karen Jones, Victor Kumar, Elizabeth Lanphier, JeeLoo Liu, Lyn Radke, François Schroeter, Luke Semrau, and Krista Thomason.



## References

- Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. A. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *Journal of abnormal psychology*, 121(2), 484–497.
- Blair, R. J. R. (1995). “A Cognitive Developmental Approach to Morality: Investigating the Psychopath.” *Cognition* 57(1): 1-29.
- Blair, R. J. R. (1996). Brief report: Morality in the autistic child. *Journal of Autism and Developmental Disorders* 26(5): 571–579.
- Blair, R. J. R. (2007). “The Amygdala and Ventromedial Prefrontal Cortex in Morality and Psychopathy.” *Trends in Cognitive Sciences* 11(9): 387-392.
- Campbell, R. & Kumar, V. (2012). “Moral Reasoning on the Ground.” *Ethics* 122 (2):273-312.
- Ciaramelli, E., Muccioli, M., Ladavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, 2(2), 84–92.
- Chapman, H. A. and A. K. Anderson, 2013. ‘Things Rank and Gross in Nature: A Review and Synthesis of Moral Disgust’. *Psychological Bulletin* 139(2): 300–327.
- Cima, M., Tonnaer, F., & Hauser, M. D. (2010). Psychopaths know right from wrong but don’t care. *Social Cognitive and Affective Neuroscience*, 5(1), 59–67.
- Craigie, J. (2011). Thinking and feeling: Moral deliberation in a dual-process framework. *Philosophical Psychology* 24(1): 53–71.
- Cushman, F., Young, L., & Hauser, M. D. (2006). The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm. *Psychological Science*, 17(12), 1082–1089.
- D’Arms, J. & Jacobson, D. (2014). Sentimentalism and Scientism. In J. D’Arms & D. Jacobson (Eds.), *Moral Psychology and Human Agency*. Oxford University Press.
- Damasio, A. (1994/2005). *Descartes’ Error*. Penguin Books. (Originally published by Putnam.)
- Dwyer, S. (2009). “Moral Dumbfounding and the Linguistic Analogy.” *Mind and Language* 24(3): 274–296.
- Eskine, K. J., Kacirik, N. A., & Prinz, J. J. (2011). A Bad Taste in the Mouth: Gustatory Disgust Influences Moral Judgment. *Psychological Science* 22(3): 295–299.
- Feltz, A. & May, J. (2017). The Means/Side-Effect Distinction in Moral Cognition: A Meta-Analysis. *Cognition* 166: 314–327.
- Glenn, A. L., Schug, R. A., Young, L., & Hauser, M. D. (2009). “Increased DLPFC activity during moral decision-making in psychopathy.” *Molecular Psychiatry*, 14(10), 909–911.
- Gold, N., Pulford, B. D., & Colman, A. M. (2013). “Your Money or Your Life.” *Economics and Philosophy*, 29(02), 213–233.
- Greene, J. (2013). *Moral Tribes*. Penguin Press.
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition* 107(3): 1144–1154.
- Haidt, J. (2001). “The Emotional Dog and Its Rational Tail.” *Psychological Review* 108/4: 814–834.
- Haidt, J. (2003). “The Moral Emotions.” *Handbook of Affective Sciences*, R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (eds.). Oxford: Oxford University Press.
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog. *Journal of Personality and Social Psychology* 65(4): 613–628.
- Hare, R. D. (1993). *Without Conscience*. Guilford Press.
- Hauser, M. D., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A Dissociation Between Moral Judgments and Justifications. *Mind & Language*, 22(1), 1–21.

- Horberg, E. J., Oveis, C., & Keltner, D. (2011). "Emotions as Moral Amplifiers." *Emotion Review* 3(3): 237-244.
- Horgan, T., & Timmons, M. (2007). Morphological Rationalism and the Psychology of Moral Judgment. *Ethical Theory and Moral Practice*, 10(3), 279–295.
- Horne, Z. & Powell, D. (2016). How large is the role of emotion in judgments of moral dilemmas? *PLoS ONE*.
- Huebner, B. (2015). Do Emotions Play a Constitutive Role in Moral Cognition? *Topoi*, 34(2), 427–440.
- Huebner, B., Dwyer, S., and M. Hauser (2009). "The Role of Emotion in Moral Psychology." *Trends in Cognitive Sciences* 13(1): 1–6.
- Huebner, B., Hauser, M. D., & Pettit, P. (2011). How the Source, Inevitability and Means of Bringing About Harm Interact in Folk-Moral Judgments. *Mind & Language* 26(2): 210-233.
- Jacobson, D. (2012). Moral Dumbfounding and Moral Stupefaction. In *Oxford Studies in Normative Ethics, Volume 2*, Mark Timmons (ed.), Oxford University Press.
- Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2012). The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience*, 7(4), 393–402.
- Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). "Harm, affect, and the moral/conventional distinction." *Mind & Language* 22(2): 117-131.
- Kelly, D. (2011). *Yuck!: The Nature and Moral Significance of Disgust*. MIT Press.
- Kennett, J. (2002). "Autism, Empathy and Moral Agency." *Philosophical Quarterly* 52 (208):340-357.
- Kennett, J. (2006). Do psychopaths really threaten moral rationalism? *Philosophical Explorations*, 9(1), 69–82.
- Kennett, J., & Fine, C. (2008). Internalism and the Evidence from Psychopaths and "Acquired Sociopaths." In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 3*. MIT Press.
- Kennett, J., & Fine, C. (2009). "Will the real moral judgment please stand up?" *Ethical Theory and Moral Practice* 12(1): 77-96.
- Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, 7(6), 708–714.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M. D., & Damasio, A. R. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, 446(7138), 908–911.
- Kohlberg, L. (1973). "The Claim to Moral Adequacy of a Highest Stage of Moral Judgment." *The Journal of Philosophy* 70(18): 630-646.
- Kumar, V. (2015). "Moral judgment as a natural kind." *Philosophical Studies* 172(11): 2887-2910.
- Landy, J. F. and G. P. Goodwin (2015). "Does Incidental Disgust Amplify Moral Judgment? A Meta-analytic Review of Experimental Evidence." *Perspectives on Psychological Science* 10(4): 518-536.
- Maibom, H. L. (2005). "Moral Unreason: The Case of Psychopathy." *Mind and Language* 20(2):237-57.
- Maibom, H. (2010). "What Experimental Evidence Shows Us About the Role of Emotions in Moral Judgement." *Philosophy Compass* 5 (11):999-1012.
- Mallon, R. & Nichols, S. (2010). "Rules." In *The Moral Psychology Handbook*, ed. J. M. Doris and The Moral Psychology Research Group, New York: Oxford University Press: 297–320.
- Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews* 32(3): 454–465.

- May, J. (2014a). "Does Disgust Influence Moral Judgment?" *Australasian Journal of Philosophy* 92(1): 125–141.
- May, J. (2014b). "Moral Judgment and Deontology: Empirical Developments." *Philosophy Compass* 9(11): 745-755.
- May, J. (2016). "Repugnance as Performance Error: The Role of Disgust in Bioethical Intuitions." *The Ethics of Human Enhancement: Understanding the Debate*, Stephen Clarke et al (eds.), Oxford University Press.
- May, J. (2018). *Regard for Reason in the Moral Mind*. Oxford University Press.
- May, J. & Kumar, V. (2018). "Moral Reasoning and Emotion." K. Jones, M. Timmons, & A. Zimmerman (eds.) *The Routledge Handbook of Moral Epistemology*. Routledge.
- Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An Investigation of Moral Judgement in Frontotemporal Dementia. *Cognitive and Behavioral Neurology*, 18(4), 193–197.
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(02), 57–74.
- Mikhail, John, 2011. *Elements of Moral Cognition*. Cambridge University Press.
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). "The Neural Basis of Human Moral Cognition." *Nature Reviews Neuroscience* 6(10): 799-809.
- Nado, J., Kelly, D., & Stich, S. (2009). "Moral Judgment." *The Routledge Companion to Philosophy of Psychology*, J. Symons & P. Calvo (eds.). Routledge, pp. 621-633.
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition* 84(2): 221–236.
- Nichols, S. (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. New York: Oxford University Press.
- Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition* 100(3): 530–542.
- Nucci, L. P. (2001). *Education in the Moral Domain*. Cambridge University Press.
- Nussbaum, M. C. (2001). *Upheavals of Thought*. New York: Cambridge University Press.
- Paxton, J. M., Ungar, L., & Greene, J. D. (2011). Reflection and Reasoning in Moral Judgment. *Cognitive Science* 36(1), 163–177.
- Pizarro, D. A. (2000). Nothing More than Feelings? *Journal for the Theory of Social Behaviour*, 30(4), 355–375.
- Pizarro, D., Inbar, Y., and C. Helion (2011). "On Disgust and Moral Judgment." *Emotion Review* 3(3): 267–268.
- Plakias, Alexandra (2013). "The Good and the Gross." *Ethical Theory and Moral Practice* 16(2): 261–78.
- Prinz, J. (2006). "The Emotional Basis of Moral Judgments." *Philosophical Explorations* 9(1): 29-43.
- Prinz, J. 2007. *The Emotional Construction of Morals*, Oxford University Press.
- Prinz, J. (2016). "Sentimentalism and the Moral Brain." S. Matthew Liao (ed.), *Moral Brains: The Neuroscience of Morality*. New York: Oxford University Press.
- Railton, P. (2014). The Affective Dog and Its Rational Tale: Intuition and Attunement. *Ethics*, 124(4), 813–859.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* 118(3): 219–235.
- Roskies, A. L. (2003). Are Ethical Judgments Intrinsically Motivational? Lessons From "Acquired Sociopathy." *Philosophical Psychology* 16(1): 51–66.
- Royzman, E. B., Goodwin, G. P., & Leeman, R. F. (2011). "When Sentimental Rules Collide: "Norms with Feelings" in the Dilemmatic Context." *Cognition* 121(1): 101-114.
- Royzman, E. B., Leeman, R. F., & Baron, J. (2009). "Unsentimental ethics: Towards a content-specific account of the moral–conventional distinction." *Cognition* 112(1): 159-174.

- Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, 10(4), 296–313.
- Sauer, H. (2012). Psychopaths and Filthy Desks. *Ethical Theory and Moral Practice* 15(1): 95–115.
- Schnall, S., Haidt, J., Clore, G. L., and A. H. Jordan (2008). “Disgust as Embodied Moral Judgment.” *Personality and Social Psychology Bulletin* 34: 1096-1109.
- Seidel, A., & Prinz, J. (2013a). Sound morality: Irritating and icky noises amplify judgments in divergent moral domains. *Cognition* 127(1): 1-5.
- Seidel, A., & Prinz, J. (2013b). Mad and glad: Musically induced emotions have divergent impact on morals. *Motivation and Emotion* 37(3): 629-637.
- Sinhababu, Neil (2017). *Humean Nature: How Desire Explains Action, Thought, and Feeling*. Oxford University Press.
- Smith, M. (1994). *The Moral Problem*. Blackwell.
- Stich, S. (1978). Beliefs and Subdoxastic States. *Philosophy of Science*, 45(4), 499–518.
- Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge University Press.
- Young, L., & Tsoi, L. (2013). When Mental States Matter, When They Don't, and What That Means for Morality. *Social and Personality Psychology Compass*, 7(8), 585–604.
- Wheatley, T. & Haidt, J. (2005). “Hypnotic Disgust Makes Moral Judgments More Severe.” *Psychological Science* 16(10): 780-784.
- Zimmerman, Aaron (2013). “Mikhail’s Naturalized Moral Rationalism.” *Jerusalem Review of Legal Studies* 8: 44-65.