

Moral Judgment and Deontology: Empirical Developments

Joshua May

Published (2014) in *Philosophy Compass* 9(11): 745-755.

[Penultimate draft; any citations should be to the [final, published version](#).]

Abstract: A traditional idea has it that moral judgment involves more than calculating the consequences of actions; it also requires an assessment of the agent's intentions, the act's nature, and whether the agent uses another person as a means to her ends. I survey experimental developments suggesting that ordinary people often tacitly reason in terms of such deontological rules. It's now unclear whether we should posit a traditional form of the Doctrine of Double Effect. However, further research suggests that a range of non-consequentialist factors influence moral judgment, including: intentions, commissions, personal harm, and agent-centered commitments. Many, if not all, such factors appear to affect how involved the agent is in bringing about an outcome.

Keywords: trolley problem, double effect, omissions, agent-centered, battery, norms, moral dilemmas, linguistic analogy, Kantianism

1. Introduction

It seems a truism that we conceive of morality in terms of rules, principles, or norms. Rules figure prominently in many cultural traditions (e.g. thou shalt not kill; do unto others...), moral theories (e.g. maximize happiness), and even personal maxims (e.g. don't be so shallow; pay it forward). In fact, moral norms are so widespread that some argue they're a cultural universal and in some sense innate (Sripada & Stich 2006).

But some moral traditions deny the apparent truism. While some question which rules drive ordinary moral thinking, others question whether any rules are causally operative. This issue has long occupied philosophical debates, but it has recently garnered a great deal of attention among scientists, such as psychologists and neuroscientists. Some empirically informed researchers posit moral rules (e.g. Mallon & Nichols 2010), but the recent trend appears to be that emotions, but not rules, are integral to ordinary moral cognition (cf. Haidt 2001; Prinz 2007). The aim of this article is to discuss current work on this topic and how relates to an ethical tradition that has placed great emphasis on rules: deontology.

Deontological views are primarily contrasted with certain consequentialist theories, such as act utilitarianism, which explain the morality of an action solely in terms of its effects. A theory counts as a form of *deontology*, roughly, if it explains the morality of an action in terms of principles that are generally non-consequentialist. (Of course, even act utilitarians believe there's at least one true moral rule: maximize happiness.) Deontology is also discernible from other non-consequentialist theories, such as *certain* forms of virtue ethics, which explain morality in terms of the actor's character traits, not rules.

These are theories about what makes an act right or wrong, yet we'll focus on empirical developments, and ones that affect deontology in particular. We can't directly draw moral conclusions from simple empirical facts, but ethical theories often make assumptions subject to empirical scrutiny. For example, many philosophers evaluate moral theories partly in terms of how well they capture our intuitive reactions to particular cases (contrast Singer 2005). The structure of common sense morality, however, is partly an empirical matter, and there is mounting evidence that some elements of prominent ethical theories are relevant to understanding ordinary moral thought (e.g. Greene 2008; Lombrozo 2009).

In what follows, we'll examine some of the many studies that bear on the role that rules play in moral cognition with a focus on work that can tease apart consequentialist and deontological elements. First, we'll consider experiments involving two basic moral distinctions: acts vs. omissions and acting intentionally vs. accidentally. Second, we'll turn to evidence that moral judgment is influenced by internalized principles involving agent-centered commitments, personal harm (including contact and battery), and harming as a means. Finally, the conclusion will be that deontological approaches to at least some aspects of moral cognition fit rather well with emerging empirical research, although some pressing issues do remain.

2. Intention & Action

Much of the recent experimental work on moral judgment asks ordinary people to evaluate behavior in hypothetical scenarios, often ones previously discussed by philosophers that involve runaway trolleys (starting with Foot 1967). Ethicists have already been developing theories about how to explain patterns of intuitions (or pre-theoretical judgments) about specific cases. But, in the mid-1990s, an experimental strategy developed in which researchers began to systematically gather data on moral intuitions about such scenarios, testing theories about what principles, if any, shape moral cognition, even if participants cannot articulate the relevant principles.

Most of the scenarios feature harmful behavior, such as homicide. Harm and care are certainly core parts of commonsense morality, but other research suggests there is much else besides, including fairness, liberty, loyalty, authority, and sanctity (Haidt 2012). There is some evidence that factors, like intention, that are central to our judgments of wrongful harm are not as important in other domains, such as purity or sanctity. For example, we tend to treat incest as morally problematic even if it is accidental, whereas we don't tend to condemn accidental harm (Young & Saxe 2011). So the moral distinctions that follow may well vary across domains.

2.1 Intentional vs. accidental

One way for a moral principle to be non-consequentialist is for it to treat the motivation behind an action, not merely its consequences, as intrinsically relevant. There's a great deal of evidence suggesting that ordinary moral intuitions are strongly influenced by such considerations, especially those mental states that make us think an action was done intentionally (for an overview, see Young & Tsoi 2013).

Consider, for example, an experiment conducted by Liane Young and her collaborators (2007), which compared the effect of outcome and intention (via belief) on moral judgment. The researchers constructed different types of vignettes that systematically varied these two factors, an example of which involved an agent, Grace, who puts a white substance in her friend's coffee. The four versions of this scenario varied whether Grace believed the substance was sugar or poison (intention) and whether it actually was poison and thus whether her friend died (outcome). See Table 1.

	<i>Outcome: Negative vs. Neutral</i>	
<i>Intention:</i> <i>Negative</i>	<i>Intentional Harm:</i> Grace believes the substance is <i>poison</i> and it is <i>poison</i> (so her friend <i>dies</i>).	<i>Attempted Harm:</i> Grace believes the substance is <i>poison</i> but it's <i>sugar</i> (so her friend <i>lives</i>).
	<i>vs.</i>	<i>Neutral:</i> Grace believes the substance is <i>sugar</i> and it is <i>poison</i> (so her friend <i>dies</i>).
<i>Neutral</i>		<i>Neutral:</i> Grace believes the substance is <i>sugar</i> and it is <i>sugar</i> (so her friend <i>lives</i>).

Table 1: Example set of cases varying intention and outcome (Young et al 2007).

Participants' moral evaluations of the actions in such cases revealed some impact of both intention and outcome, but intention played a much more prominent role. People were inclined to count both Intentional Harm and Attempted Harm as *wrong*, whereas Accidental Harm and Neutral were both regarded on the whole as *permissible*. So it appears the agent's bad intention was the primary factor in perceived wrongness, whereas an innocent intention was the primary source of perceived permissibility.

Consequentialists typically maintain that intentions are only important for blame, not wrongness. Perhaps people often confuse these or will count an action as wrong just to blame the transgressor, in which case ordinary moral intuitions might better match consequentialism after all. However, in other experiments, Fiery Cushman (2008) shows that intentionality more powerfully affects judgments of wrongness than attributions of blame and punishment, for which outcomes are comparatively more important. So it seems ordinary moral thinking simply departs from this consequentialist prescription.

Intentions, then, seem rather important for moral evaluations of harm. This should be unsurprising given the importance of *mens rea* in the legal system, as when it grounds the criminality of certain attempted crimes or the difference between types of homicide. Even the moral evaluations of children as young as four are sensitive to intentionality (Cushman et al 2013). We might thus propose that something like the following *Intentionality Principle* shapes moral cognition: All else being equal, it's morally worse to cause harm intentionally than accidentally. The "all else being equal" or "ceteris paribus" clause allows us to make a generalization that may have exceptions—a common tool in psychological and moral theorizing (cf. Fodor 1987; Wedgwood 2011).

2.2 Action vs. Omission

Another basic moral distinction is between acts (or commissions) and omissions, which connects with the "Doctrine of Doing and Allowing" (cf. Woollard 2012). This is a common theme in the law, as there are usually prohibitions on actively causing harm, say, but rarely against inaction. Some jurisdictions, however, do have "Good Samaritan laws" or a "duty to rescue," according to which it is illegal to fail to help a fellow citizen in great need. Different moral and legal attitudes toward assisted suicide also rest on the distinction between actively killing patients and passively letting them die.

Several studies have now confirmed that this is a recurring element in moral cognition, beyond legal contexts (e.g. Spranca et al 1991; Cushman et al 2006; Cushman & Young 2011). In most of the experiments, participants are presented with numerous vignettes that systematically vary factors that include action versus omission. For example, in one vignette, James decides to poison an innocent person in order to save five, while in the contrasting story Robert saves five by merely allowing another to die by not administering an antidote (Cushman et al 2006). Participants consistently treat actively harming as worse than harming by omission. However, responses are typically recorded on fine-grained scales and the differences between such pairs of moral judgments are sometimes rather small (see esp. Cushman & Young 2011, Exp. 2). Moreover, scenarios contrasting acts and omissions often also differ in terms of bodily contact—an act of battery in the law—thus confounding two variables (Mikhail 2014). Still, such differences are stable and easily detectable; there is even some evidence that children as young as five draw the distinction (Powell et al 2012).

So we might follow Cushman and colleagues (2006) and posit an *Action Principle* roughly as follows: All else being equal, harm caused by action is morally worse than harm consequent upon omission. However, there is some evidence suggesting that our moral evaluations are influenced by this distinction because we think people are more *causally* implicated in an outcome if it results from action rather than omission (Cushman & Young 2011). In that case, it may be more appropriate to say that the relevant rule is a more general *Intentional Harm Principle*: All else being equal, it's morally wrong to intentionally cause harm. This captures both the significance of commissions (via causing harm) and intention. Or, even more generally, perhaps these two factors affect what Wedgwood calls the "degree of agential involvement" in the effect (2011, 392). We can then capture the significance of the

act/omission distinction with an *Agential Involvement Principle*: All else being equal, it is morally worse for an agent to be more involved in bringing about a bad outcome. Either way, there's some reason to posit tacitly represented rules that turn on the nature of the act and its relation to the agent, not outcomes.

3. Means, Force, & Agents

Let's now consider some evidence for more complex, yet still arguably deontological, principles in ordinary moral cognition. Most of the work involves the famous "trolley" thought experiments, which again deal with harm and death. In one study involving thousands of participants (Hauser et al 2007), the researchers focused on four trolley cases. An initial pair can be briefly summarized thus:

- *Switch*: A protagonist is faced with a choice. He can allow a runaway train to run over and kill five innocent people, or he can throw a switch to divert the train onto a sidetrack away from the five but toward one innocent person who will die as a result.
- *Footbridge*: The protagonist can again allow the train to kill the five; but the alternative is to save the five by pushing one large innocent person off of a bridge onto the tracks, where he will die but his body will stop the train.

The question is: Is it morally permissible to throw the switch/man? Vast majorities believe diverting the train in Switch is morally *permissible* but that throwing the man in Footbridge is *wrong*, despite the fact that they both involve the choice of sacrificing one to save five.

The asymmetry is robust too. It was found with apparently little variation across gender, some cultures and nationalities, and age (Hauser et al 2007). It has even been identified in children as young as three (Pellizzoni et al 2010) and is quite generalizable to harms other than death—e.g. financial loss, bodily harm, and emotional distress (Gold et al 2013). There is some evidence for slight East-West cultural variation in intuitions about some trolley cases (e.g. Gold et al 2014). But, even if there is substantial variation, it doesn't count against the existence of internalized rules, only universal ones. Moreover, East Asian responses to such dilemmas appear even less utilitarian, possibly stemming from an aversion to becoming actively involved.

How should we explain the different responses to Switch and Footbridge? In attempting to answer this question, we'll examine further research that points to additional influences on moral judgment, suggesting the involvement of further non-consequentialist moral rules.

3.1 Means vs. Side-Effect

One explanation of the Switch/Footbridge asymmetry appeals to the distinction between generating an outcome as a mere *side effect* of one's plan versus bringing it about as a *means* to some goal. In Switch, it seems that harming the one is an unintended side

effect of saving the five. In Footbridge, however, the protagonist must sacrifice the one as a means to saving the five.

The means/side-effect distinction is often employed in the famous *Doctrine of Double Effect* (DDE), which has been widely discussed both empirically and theoretically. The DDE can be formulated in various and often complicated ways (e.g. Mikhail 2011, 149). The basic idea is that there's an important moral difference between intending and merely foreseeing bad outcomes (FitzPatrick 2012). Such a principle has been used to defend the moral permissibility of various killings, e.g. when an abortion is required to save the life of the mother (cf. Foot 1967). Importantly, according to several theorists, something like this principle tacitly operates in our ordinary moral thinking (e.g. Harman 1999; Mikhail 2011).

Some variations on the original pair of trolley cases apparently provide support for this idea. Consider an important second pair of cases, briefly summarized as follows:

- *Loop Track*: The first option is again to do nothing and let the five die while the alternative is to divert the train onto a sidetrack. But now there is a large innocent man there who can stop the train, dying in the process. The important difference here is that the sidetrack loops back onto the main track where the five are, such that killing the one is required to save the five. (cf. Thomson 1985)
- *Man-in-Front*: This variant is meant to be exactly like Loop except that there is a large object behind the one man on the side track, and it's enough on its own to stop the train. However, since the man is in front of this object, the protagonist knows the one must die if the switch is flipped.

If ordinary people implicitly think in terms of DDE, then presumably one would find roughly the following pattern of results. On the one hand, we'd be more likely to say that flipping the switch in Man-in-Front is morally permissible, since this would involve a bad consequence (harming an innocent person) that is not intended but merely foreseen. On the other hand, we'd tend to treat flipping the switch as impermissible in Loop, since it involves sacrificing the one as a means to saving the five, just like Footbridge.

Initial results roughly fit with the predictions of DDE. Most people tended to think throwing the switch is permissible in Man-in-Front but fewer said this about Loop (62% vs. 48% in Mikhail 2007), and the difference is statistically significant. This even holds for the large-scale online study done by Hauser, Mikhail, and their collaborators (Hauser et al 2007).

We might thus attribute a *Means Principle* to ordinary moral judgment (a la DDE): All else being equal, it is morally worse to bring about a bad outcome as a means to an end than as a merely foreseen side effect. However, there is some evidence that the means/side-effect distinction affects our moral judgments via assessments of whether the agent *intended* the outcome (Sinnott-Armstrong et al 2008; Cushman & Young 2011). That is, we're more likely to treat as intentional those outcomes caused as a means than as a side effect, and this impacts our moral judgments via something like the Intentionality Principle (§2.1). This apparently conflicts with research suggesting

the reverse: that our moral judgments influence which side effects we treat as brought about intentionally (Knobe 2003). But perhaps the means/side-effect distinction affects moral judgments most prominently when thinking about *intending* an outcome rather than just bringing it about *intentionally* (cf. Nadelhoffer 2006). In any event, the Means Principle may simply be an application of the Intentionality Principle or a more general principle that appeals to the agent's degree of involvement (§2.2).

There are several criticisms one could make of the Loop studies. One problem is that the data support DDE only weakly. For example, looking just at the percentages of people who thought flipping the switch is permissible (from Hauser et al 2007), there is a large difference between Switch and Footbridge (85% vs. 12%) but not between Loop and Man-in-Front (56% vs. 72%). The differences between each pair are statistically significant, but that just means the observed differences were not likely due to mere chance. As some commentators have noted, DDE should presumably treat the two pairs of cases similarly, so it is odd that the difference in responses to the Switch/Footbridge pair is so much larger than in the Loop/Man pair (e.g. Phelan 2012).

Another problem is that some have failed to replicate the original results. For example, Waldmann and Dieterich (2007, Exp. 2) presented participants with, among other scenarios, a version of Loop and measured responses on a scale. Their case is not identical to Mikhail's but it's similar enough that one would expect similar results. However, participants tended to treat flipping the switch in this case as *permissible*, even though it involved harming as a means. One possible explanation is that Waldmann and Dieterich's vignettes involved causing bodily harm, not death. However, if harming as a means substantially influences moral judgment, one would expect at least a slight difference in morality ratings between cases like Loop and *Switch*, but these researchers did not find one.

Waldmann and Dieterich's experiment was not fully a direct replication attempt, but other attempts have fewer if any differences between the original Loop studies. Sinnott-Armstrong and his collaborators (2008), for example, found that subjects tended to count turning the trolley in their version of Loop as *not* morally wrong. Even more strikingly, Josh Greene and colleagues (2009, Exp. 2a) used versions of Loop and Man-in-Front ("loop weight") modeled closely on Mikhail's. Yet participants tended to treat both of these cases as morally *acceptable*, and there was no statistically significant difference between responses, despite using fine-grained scales (Zimmerman 2013 reports similar results).

How can we explain these discrepancies? Waldmann and Dieterich propose that in Mikhail's studies the large man is referred to as a "heavy object" in Loop but not in Man-in-Front. This might portray the victim as being treated less humanely in Loop, which could explain why most participants would think flipping the switch is impermissible in that version of the scenario.

Of course, an effect may still exist even if others failed to detect it. Subsequent studies may lack the statistical power to accept the null hypothesis that there is no difference in responses. Moreover, perhaps the Loop-style cases simply aren't suited for revealing that harming as a means is perceived as morally worse than harming as a side effect (cf. FitzPatrick 2012). Several studies have found precisely this effect in other

contexts (cf. Cushman et al 2006; Moore et al 2008; Sinnott-Armstrong et al 2008; Cushman & Young 2011).

However, these experiments may have some limitations. First, in contrast with Mikhail's approach, these other studies measure moral judgment using scales, which can pick up small differences without reflecting much about the categorical difference between right and wrong (compare May 2014). Indeed the differences are often rather small between moral evaluations of harming as a means versus a side effect. This fits with our modest Means Principle, but it is more difficult to square with many traditional versions of DDE (but see Wedgwood 2011). Second, most participants receive similar pairs of cases that contrast the relevant distinction (e.g. means/side-effect, act/omission), which may prime people to respond the way the researchers wanted—a problem of “experimenter demand.” (This isn't true of Sinnott-Armstrong et al 2008, given their between-subjects design, but a quarter of their participants had “previously heard of trolley problems.”) Third, researchers often vary so many factors across numerous dilemmas that one might worry about confounding different variables. For example, as previously noted (§2.2), vignettes contrasting acts and omissions often also contrast something like bodily contact or battery (cf. Mikhail 2014).

Thus, while there is some evidence that the means/side-effect distinction is treated as morally significant, further research could make matters clearer. As we've seen, there's reason to think the effect is driven by treating harming as a means as more intentional (Cushman & Young 2011)—a factor that affects the agent's involvement in the outcome (Wedgwood 2011). Perhaps again we have a version of a more general principle concerning intentionally causing harm or agential involvement (§2.2).

3.2 Personal Harm: Force, Contact, and Battery

Another difference between Switch and Footbridge involves the personal way in which harm is inflicted. A number of related factors in this category appear to influence moral judgments.

Greene (2008) famously argues that we tend to think it's wrong to push someone in cases like Footbridge because this involves *up-close and personal* harm (cf. Singer 2005; Prinz 2007, 1.2.2). Greene appeals to a wide range of empirical evidence for this claim but focuses on his own brain imaging studies, which suggest greater emotional processing when judging “personal” moral dilemmas. However, there are some methodological worries about the studies (e.g. Berker 2009) and one reanalysis of his data indicates that the effects were not driven by the personal nature of the cases (McGuire et al 2009). Moreover, the personal/impersonal distinction does not appear to capture a wide range of cases. There are many situations in which inflicting personal harm is regarded as permissible, e.g. when defending one's family, honor, or life (cf. Mallon & Nichols 2010).

Still, various experiments suggest that *contact* generally affects moral intuitions. Studies led by Fiery Cushman present participants with a host of scenarios that systematically contrast a number of factors, including physical contact (Cushman et al 2006; Cushman & Young 2011). For example, one pair of cases includes a version of

Footbridge and a modified case in which the protagonist can sacrifice one in order to save five by pulling a lever that drops the victim in front of a runaway boxcar. People treat this “drop” version of the case, which does not involve physical contact, as somewhat better than pushing. (For some potential limitations of these studies, see §3.1.) Physical contact, however, is distinct from personal harm, since the latter includes harm inflicted in the absence of physical contact—e.g. poisoning Grandmother’s tea.

Greene and his collaborators (2009) have since switched to the notion of harm via *personal force*, defined as force that directly impacts another and is generated by one’s muscles. Pushing someone with one’s hands or a pole counts as using personal force, for example, but directing a drone attack doesn’t. In one experiment, participants rated the Footbridge scenario and variants on it. In one modified version, for example, the protagonist can push the large man with a pole; in another he can drop him on the tracks through a trap door by pulling a lever. The researchers found that people were less likely to say killing the one to save the five is permissible if it involved personal force. Moreover, further experiments suggest this effect is mediated by whether the agent intended the harmful outcome.

Related to personal force is the notion of *battery*, often colloquially termed “assault.” Mikhail identifies a relevant moral rule, embodied in the law, as the Prohibition of Intentional Battery, which “forbids purposefully or knowingly causing harmful or offensive contact with another individual or otherwise invading her physical integrity without her consent” (2011, 117). Of course, as with other rules, this is not without exception. For example, many think it permissible to injure a frail person by pushing him out of the way of an oncoming train in order to save his life (see the case of “Implied Consent” in Mikhail 2011). As in the law, such harm via personal contact is justified (Mikhail 2014). Similarly, if sacrificing the life of one via harmful battery saves many more people than five, such as a billion, then few will think it wrong, especially if the one would otherwise die as well (Nichols & Mallon, Exp. 3). Perhaps such exceptions trigger Pareto considerations about not making a victim any worse off, which have been shown to influence moral judgment (Moore et al 2008; Huebner et al 2011).

It’s difficult to draw firm conclusions about how contact, personal force, or battery affect moral cognition. Battery is somewhat similar to contact and personal force, but it needn’t be especially personal or physical, as when one causes bodily injury by intentionally directing a trolley toward someone (Mikhail 2014, §1.3). Perhaps these three factors can be unified into a single principle, or perhaps one factor can explain all the relevant data. Further research is clearly needed. Until then, we can’t gain much precision. I leave the matter with a recent quip from Paul Bloom: “Here is a good candidate for a moral rule that transcends space and time: If you punch someone in the face, you’d better have a damn good reason for it” (2013, 10).

3.3 Agent-centered vs. Agent-neutral

Let’s turn briefly away from harm. Keeping one’s own commitments seems to take priority over helping others keep theirs. For example, it seems wrong to cheat on one’s

spouse just to help keep two others from cheating on theirs, even if doing so would lead to slightly more happiness or more fidelity in the world. As some philosophers have argued, such moral commitments, and the reasons they generate, seem “agent-centered.” It matters who has the commitment, and we don’t simply weigh up the commitments of a similar type in an “agent-neutral” manner.

This apparently non-consequentialist idea has been demonstrated in some experiments done by Theresa Lopez and colleagues (2009). They presented non-specialists with vignettes in which a protagonist could break a commitment of his own, such as a promise, in order to help *another* person fulfill even more such commitments. Participants tended to think one *shouldn’t* do this. However, they also provided their moral evaluation of cases in which the protagonist would have to break one of his own commitments to keep more such commitments he had made to others. People tended to think one *should* break a commitment when it’s necessary to keep a greater number of one’s own commitments. This suggests a cost-benefit calculation is perceived as morally relevant only when the dilemma requires failing to satisfy one’s *own* obligations either way. The inevitability of breaking one’s own commitment seems relevant here: in terms of violating moral commitments, no option makes one worse off (compare the effect of “Pareto considerations” in Huebner et al 2011).

What should we conclude? Some of the differences in moral judgment were rather small (esp. Exp. 2). Moreover, the studies involved presenting each participant with both contrasting cases, which may prime subjects to respond as the experimenters wanted. However, we might attribute to ordinary moral cognition an *Agent-Centered Principle* that roughly states: All else being equal, one has more moral reason to keep one’s own commitments than to help others keep theirs. The specific structure of such a principle would undoubtedly need to be refined in light of more research.

4. Conclusion

As far as rigorous empirical evidence goes, there is good news and bad news for the rule-based approach to moral cognition. The bad news is that we arguably lack strong evidence that prominent versions of the Doctrine of Double Effect play a role in ordinary moral thought. The good news, however, is that there is mounting evidence for similar non-consequentialist moral rules. Ordinary moral thinking is substantially influenced by the agent’s intentions and the nature of the acts themselves, including whether they are omissions, violate the agent’s own commitments, or involve some kind of direct invasion of another’s bodily integrity. Interestingly, such considerations often seem to reflect on the agent’s involvement in bad outcomes and the corresponding principles are surprisingly insensitive to consequences, even if sometimes factored in. Perhaps our best theory will reinterpret such tactic rules as utilitarian (or generally consequentialist), but it doesn’t seem to be the natural account to give at this point.

Of course, one might worry about whether we need to posit moral rules at all to explain the impact of these various factors on moral cognition. Some such objections turn on data, e.g. whether only emotions are required for moral judgment (Haidt 2001; Prinz 2007 vs. Nichols 2004; Maibom 2010; May 2014). Other objections turn on general issues in cognitive science, such as when to posit rule-based cognition (Horgan

& Timmons 2007; Dwyer 2009 vs. Clarke 2000; Zimmerman 2013). And there are especially conceptual questions about whether genuine moral judgment can arise in the absence of reflective endorsement of tacitly represented rules (Kennett & Fine 2009).

Further issues surrounding moral judgment likewise remain. For example, some researchers believe that, while some moral intuitions arise from deontological principles, many others arise from more utilitarian calculations (Cushman et al 2010). Relatedly, we haven't addressed whether moral judgments resulting from applications of non-consequentialist rules are justified (Singer 2005; Greene 2008; Sinhababu 2013 vs. Berker 2009; Mikhail 2011; Nichols et al ms). However, illuminating non-consequentialist influences on moral judgment can undoubtedly aid such further investigations.

Acknowledgements: For feedback on earlier drafts, I'm grateful to: Mark Alfano, Victor Kumar, Andy Lamey, David Morrow, Shaun Nichols, Gregory Pence, Rob Shaver, Mark van Roojen, Farid Zaid, and Aaron Zimmerman.

References

- Berker, Selim (2009). "The Normative Insignificance of Neuroscience." *Philosophy & Public Affairs* 37, pp. 293-329.
- Bloom, Paul (2013). *Just Babies: The Origins of Good and Evil*. New York: Crown.
- Clark, Andy (2000). "Word and Action: Reconciling Rules and Know-How in Moral Cognition." *Canadian Journal of Philosophy*, Supplementary Vol. 26: Moral Epistemology Naturalized, R. Campbell and B. Hunter (eds.): 267-90.
- Cushman, Fiery (2008). "Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment." *Cognition* 108:353–80.
- Cushman, F., & Young, L. (2011). "Patterns of Moral Judgment Derive from Nonmoral Psychological Representations." *Cognitive Science* 35(6): 1052-1075.
- Cushman, F., Young, L., & Hauser, M. (2006). "The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm." *Psychological Science*, 17(12), 1082–1089.
- Cushman, F., Young, L. and J. Greene (2010). "Multi-System Moral Psychology." In *The Moral Psychology Handbook*, ed. J. M. Doris and The Moral Psychology Research Group, New York: Oxford University Press: 47–71.
- Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). "The Development of Intent-based Moral Judgment." *Cognition* 127: 6–21.
- Dwyer, Susan (2009). "Moral Dumbfounding and the Linguistic Analogy" *Mind and Language* 24 (3):274-296.
- FitzPatrick, W. I. (2012). "The Doctrine of Double Effect: Intention and Permissibility." *Philosophy Compass* 7(3): 183–196.
- Fodor, Jerry (1987). *Psychosemantics*. Cambridge, MA: MIT Press (Bradford).
- Foot, Philippa. (1967). "The Problem of Abortion and the Doctrine of the Double Effect." *Oxford Review*, 5, 5–15.
- Gold, N., Pulford, B. D., & Colman, A. M. (2013). "Your Money or Your Life: Comparing Judgements in Trolley Problems Involving Economic and Emotional Harms, Injury and Death." *Economics and Philosophy* 29: 213–33.

- Gold, N., Colman, A. M., & Pulford, B. D. (2014). "Cultural Differences in Responses to Trolley Problems." *Judgment and Decision Making* 9(1): 65-76.
- Greene, Joshua (2008). "The Secret Joke of Kant's Soul." *Moral Psychology, Vol. 3: The Neuroscience of Morality*, Walter Sinnott-Armstrong (ed.), MIT Press, pp. 35-117.
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., and Cohen, J. D. (2009). "Pushing moral buttons: The interaction between personal force and intention in moral judgment." *Cognition*, 111 (3): 364 –371.
- Haidt, Jonathan (2001). "The Emotional Dog and Its Rational Tail." *Psychological Review* 108/4: 814–834.
- Haidt, Jonathan (2012). *The Righteous Mind*. New York: Pantheon.
- Harman, Gilbert (1999). "Moral Philosophy and Linguistics." In K. Brinkmann (Ed.), *Proceedings of the 20th World Congress of Philosophy: Vol. 1. Ethics*, pp. 107–115. Bowling Green, OH: Philosophy Documentation Center.
- Hauser, M., Cushman, F., Young, L., Jin, R., J. Mikhail (2007). "A Dissociation Between Moral Judgments and Justifications." *Mind and Language* 22(1):1–21.
- Horgan, Terry & Timmons, Mark (2007). "Morphological Rationalism and the Psychology of Moral Judgment." *Ethical Theory and Moral Practice* 10(3): 279-295.
- 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.
- Kennett, Jeanette & Fine, Cordelia (2009). "Will the Real Moral Judgment Please Stand Up?" *Ethical Theory and Moral Practice* 12:77–96.
- Knobe, Joshua (2003). "Intentional Action and Side Effects in Ordinary Language." *Analysis* 63(279): 190-194.
- Lombrozo, Tania (2009). "The Role of Moral Commitments in Moral Judgment." *Cognitive Science* 33(2): 273-286.
- Lopez, T., Zamzow, J., Gill, M., & Nichols, S. (2009). "Side Constraints and the Structure of Commonsense Ethics." *Philosophical Perspectives* 23(1): 305-319.
- Maibom, Heidi (2010). "What Experimental Evidence Shows Us About the Role of Emotions in Moral Judgement." *Philosophy Compass* 5 (11):999-1012.
- Mallon, Ron and Shaun Nichols (2010). "Rules." In *The Moral Psychology Handbook*, ed. J. M. Doris and The Moral Psychology Research Group, New York: Oxford University Press, 297–320.
- May, Joshua (2014). "Does Disgust Influence Moral Judgment?" *Australasian Journal of Philosophy* 92 (1): 125-141.
- McGuire, J., Langdon, R., Coltheart, M., C. Mackenzie (2009). "A Reanalysis of the Personal/Impersonal Distinction in Moral Psychology Research." *Journal of Experimental Social Psychology* 45(3): 577–580.
- Mikhail, John (2007). "Universal Moral Grammar: Theory, Evidence and the Future." *TRENDS in Cognitive Sciences* 11(4): 143–52.
- Mikhail, John (2008). "Moral Cognition and Computational Theory: Comment on Greene." In W. Sinnott-Armstrong (ed.) *Moral Psychology, Vol. 3*. MIT Press, pp. 81–91.
- Mikhail, John (2011). *Elements of Moral Cognition*. Cambridge University Press.
- Mikhail, John (2014). "Any Animal Whatever? Harmful Battery and its Elements as Building Blocks of Moral Cognition." *Ethics* 124(4): 750-786.
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). "Who Shalt not Kill? Individual differences in working memory capacity, executive control, and moral judgment." *Psychological Science*, 19(6): 549-557.
- Nadelhoffer, Thomas (2006). "On Trying to Save the Simple View." *Mind & Language* 21(5): 565-586.

- Nichols, Shaun (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. New York: Oxford University Press.
- Nichols, S., Kumar, S. & Lopez, T. (ms). "Rational Learners and Non-utilitarian Rules."
- Nichols, S. & R. Mallon (2006). "Moral Dilemmas and Moral Rules." *Cognition* 100(3): 530–542.
- Pellizzoni, S., Siegal, M., & Surian, L. (2010). "The Contact Principle and Utilitarian Moral Judgments in Young Children." *Developmental Science* 13(2): 265-270.
- Phelan, Mark (2012) "Review of Elements of Moral Cognition." *Notre Dame Philosophical Reviews*. <<http://ndpr.nd.edu/>>
- Powell, N. L., Derbyshire, S. W., & Guttentag, R. E. (2012). "Biases in Children's and Adults' Moral Judgments." *Journal of Experimental Child Psychology* 113(1): 186-193.
- Prinz, Jesse 2007. *The Emotional Construction of Morals*. New York: Oxford University Press.
- Sinhababu, Neil (2013). "Unequal Vividness and Double Effect." *Utilitas* 25(3): 291-315.
- Singer, Peter (2005). "Ethics and Intuitions." *The Journal of Ethics* 9: 331–52.
- Sinnott-Armstrong, W. Mallon, R., McCoy, T., & Hull, J. G. (2008). "Intention, Temporal Order, and Moral Judgments." *Mind & Language* 23(1): 90-106.
- Spranca, M., Minsk, E., & Baron, J. (1991). "Omission and Commission in Judgment and Choice." *Journal of Experimental Social Psychology* 27: 76–105.
- Sripada, Chandra & Stephen Stich (2006). "A Framework for the Psychology of Norms." In Peter Carruthers, Stephen Laurence & Stephen P. Stich (eds.), *The Innate Mind, Volume 2: Culture and Cognition*. Oxford University Press.
- Thomson, Judith Jarvis (1985). "The Trolley Problem." *The Yale Law Journal* 94(6):1395-1415.
- Waldmann, M. , & Dieterich, J. (2007). "Throwing a bomb on a person versus throwing a person on a bomb: Intervention myopia in moral intuitions." *Psychological Science*, 18: 247–253.
- Wedgwood, Ralph (2011). "Defending Double Effect." *Ratio* 24(4):384–401.
- Woollard, Fiona (2012). "The Doctrine of Doing and Allowing I: Analysis of the Doing/Allowing Distinction." *Philosophy Compass* 7(7): 448–458.
- Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). "The Neural Basis of the Interaction between Theory of Mind and Moral Judgment." *Proceedings of the National Academy of Sciences of the United States of America* 104: 8235–40.
- Young, L., & Saxe, R. (2011). "When Ignorance is No Excuse: Different Roles for Intent Across Moral Domains." *Cognition* 120(2): 202-214.
- Young, Liane & Lily Tsoi (2013). "When Mental States Matter, When They Don't, and What That Means for Morality." *Social and Personality Psychology Compass* 7(8): 585–604.
- Zimmerman, Aaron (2013). "Mikhail's Naturalized Moral Rationalism." *Jerusalem Review of Legal Studies* 8: 44-65.