

Reliable but not home free? What framing effects mean for moral intuitions

Article

Accepted Version

Andow, J. (2016) Reliable but not home free? What framing effects mean for moral intuitions. *Philosophical Psychology*, 29 (6). pp. 904-911. ISSN 1465-394X doi: 10.1080/09515089.2016.1168794 Available at <https://centaur.reading.ac.uk/51203/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1080/09515089.2016.1168794>

Publisher: Taylor & Francis

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Reliable but not home free? What framing effects mean for moral intuitions*

James Andow

January 8, 2016

Abstract

Various studies show moral intuitions to be susceptible to framing effects. Many have argued that this susceptibility is a sign of unreliability and that this poses a methodological challenge for moral philosophy. Recently, doubt has been cast on this idea. It has been argued that extant evidence of framing effects does not show that moral intuitions have a unreliability problem. I argue that, even if the extant evidence suggests that moral intuitions are fairly stable with respect to *what* intuitions we have, the effect of framing on the *strength* of those intuitions still needs to be taken into account. I argue that this by itself poses a methodological challenge for moral philosophy.

1 Introduction

In recent decades, a number of researchers have reported that moral intuitions are subject to various types of framing effects. These studies have contributed to the much discussed methodological worries about the use of intuitions in philosophical methodology which draw on framing effects and other problematic influences in various philosophical domains (see, e.g., [Swain et al., 2008](#)). The typical argument given is that the framing effects demonstrate that the intuitions are unreliable and this puts pressure on the use of intuitions in moral philosophy (e.g., [Sinnott-Armstrong, 2008](#)). This challenge has spurred a backlash. For example, it has recently been argued that the relevant evidence does not show that the extent to which moral intuitions fall short of perfect reliability is sufficient to cause problems for moral philosophy (e.g., [Demaree-Cotton, 2015](#)). It has been argued that we are entitled to assume that moral philosophers' intuitions are reliable even if we concede that ordinary folks' intuitions are unreliable (see my [Andow, 2015a](#)). In other words the issue at the heart of this backlash has been *reliability*.

In this paper, I take a step back from the issue of reliability. I will concede that, vis-à-vis reliability, extant evidence of framing effects maybe does not present much of a worry for moral philosophy. I will argue that even if moral intuitions are not sufficiently *unreliable* to generate a problem, framing effects nonetheless demonstrate that the use of intuitions in moral philosophy is problematic. The structure of the paper is as follows. §2 reviews the evidence of framing effects, outlines and evaluates the unreliability argument. §3 then argues that the framing effects nonetheless highlight a problem with the role that intuitions play in philosophical methodology. §4 makes a number of remarks in light of the preceding discussion.

At the heart of my discussion is a distinction between the intuition one has and the strength of that intuition. It might be helpful to say something to motivate that distinction upfront. Consider your favourite philosophical thought experiment. Suppose you have an intuition about it. You think

*This paper is forthcoming in *Philosophical Psychology*. Please refer to the published version for citations, page numbers, and so on.

that p (some proposition concerning the case). My attitude to the same case might be different. I might think that not- p . In this case, you would have a different intuition to me. But that is not the only way our attitudes might differ. We might agree in *polarity* (i.e., as to whether p , or not- p) but nonetheless have drastically different attitudes in terms of the *strength* of our intuitions. For example, while I might generally agree with you, my intuition might be fairly weak, suppose, while yours is very strong.

Of course, there are various ways in which an intuition might be said to differ in degree from another if not polarity. For example, one intuition might be clearer than another, or more emotionally invested, or more entrenched, or more intellectually compelling, and so on. These are important differences and the role of such differences in philosophical methodology is yet to be explored in any detail. However, for my purposes in this paper, I won't be dwelling on these differences and will instead use the word 'strength' simply to bundle all these potential types of difference in degree together. What is important for my purposes is the fact that (a) intuitions can differ in degree even if not in polarity, (b) this has significant implications within standard philosophical methodology, and, perhaps most importantly, (c) some such differences in degree might show up in survey responses.

This latter point is worth illustrating with an example. Suppose that a survey measured participants' intuitions on 6-point scales, anchored at either end by '1 – strongly disagree' and '6 – strongly agree'. Suppose that this survey asked participants the same question given two different framings. Suppose that for framing A the mean score comes back as 1.2 and for framing B the mean score is 2.5. Suppose that both these means are significantly lower than the midpoint of the scale, 3.5, and that framing A is significantly lower than B. All else being equal, the natural interpretation is that these results reflect the fact that, while participants tend to have the *same intuition* given each framing, that intuition tends to be *stronger* in some way given framing A.

2 The unreliability argument

Moral intuitions have been claimed by numerous researchers to exhibit framing effects. This means that moral intuitions about cases are sensitive to morally irrelevant differences in the ways in which those cases are presented. Typically these framing effects are order effects. For example, there are numerous studies which find that participants' intuitions about various cases are sensitive to which cases they considered beforehand (Lanteri et al., 2008; Lombrozo, 2009; Nichols and Mallon, 2006; Petrinoich and O'Neill, 1996; Schwitzgebel and Cushman, 2015; Wiegmann et al., 2012). However, there are other types of framing effects which have been claimed to affect moral intuitions and philosophical intuitions more widely, e.g., how vividly a case is presented (Bartels, 2008), whether a case is high affect or low affect (Nichols and Knobe, 2007), how mechanistically/psychologically cases are described (Nahmias et al., 2007), the presence of unconscious cleanliness cues (Tobia et al., 2013), or even what typeface a case is presented in (Tobia, 2012; Weinberg et al., 2012, who cites the unpublished Gonnerman et al. 2011). (This is not intended to be an exhaustive list.)

Are framing effects a sign of unreliability? The answer depends on what we mean by 'unreliable'. The relevant sense of 'reliability' is a measure of how likely an intuition is to be incorrect. If intuitions are liable to say that p (e.g., acting in a certain trolley case is permissible) in around 80% of cases and p is true, then intuitions would be 80% reliable (with respect to that proposition). (Things get a little more complicated when we scale up and ask questions like how reliable are peoples' intuitions generally about moral matters generally, but we don't need to get into that.) But that leaves the question of *unreliability* unanswered.

What is it then for intuitions to be *unreliable*? Here there are various senses one might have in mind. One relevant sense might be that intuitions are unreliable if they are more likely to be false than true. But framing effects typically don't show that intuitions are this unreliable. Indeed, given

that most moral intuitions are about the truth or falsity of some particular claim, there is no way that evidence of framing effects could by itself demonstrate that intuitions were less than 50% reliable (one would need some external source of evidence). However, I take it that no one ever thought that framing effects would have to show that intuitions were unreliable in this sense in order to highlight a reliability problem with moral intuitions. Rather, it seems problematic for moral philosophy even if intuitions are right only 51% of the time. This said, it is not as if *any* degree of unreliability is a worry for moral philosophers, since very few (if any) philosophers treat moral intuitions as being 100% reliable. So, the relevant sense of ‘unreliable’, then, must be somewhere between *less than perfectly reliable* and *actually more likely to be wrong than right*.

Demaree-Cotton (2015) articulates an interesting way to get a handle on the relevant sense of unreliability. The effect of irrelevant factors is worrying, is her thought, when the framing of a scenario is likely to affect which answer we give. For example, one might not be worried about the influence of irrelevant factors if, regardless of framing, one’s intuitions gave the same verdict in 99.99% of cases. On the other hand, one might be pretty worried about the reliability of your intuition that *p*, if in all likelihood you would have had the intuition that not-*p* given a different framing of the relevant question. She surveys a large number of studies which have reported framing effects in moral intuitions and, for each, reports the probability that subjects would have made the same judgment regardless of which frame they were exposed to. The resulting figures cluster at around the 80% mark.¹

Demaree-Cotton considers this figure to be quite an optimistic finding for the proponent of the use of intuitions in moral philosophy. Now, I think this optimism might not be completely warranted.² I am tempted by the thought that this figure actually shows that framing a case one way rather than another, where that difference is irrelevant to the moral facts, actually seems to be able to push our intuitions around a surprising amount (bear in mind that I antecedently expect that framing effects are just one type of a bunch of irrelevant effects on my intuitions). However, it is true that there is a kind of self-doubt which Demaree-Cotton has shown to be unwarranted by the evidence. Given that the probability you would have had a different intuition given a different framing is only around 20%, one can still be pretty confident that one’s answer would have been the same.

It is worth noting that there are various reasons one might doubt the claim that the available evidence suggests that intuitions have a reliability problem, in addition to those which Demaree-Cotton provides. For example, many extant studies use between-subjects designs one might think are non-ideal for the task of investigating the influence of framing effects. But, for my purposes, it is not important to do a complete survey of all the reasons one might doubt whether the relevant evidence shows signs of problematic unreliability. Instead, I want to argue that even if Demaree-Cotton and other’s doubts hold out, and people’s intuitive judgments are much more stable (in terms of yes/no answers) than has been claimed, this does not mean that the available evidence presents no methodological challenge to the use of intuitions in philosophy.

¹Despite being careful elsewhere, in her conclusion Demaree-Cotton (2015) cites this figure as suggesting that intuitions are around 80% *reliable*. We should be careful in making this leap. We shouldn’t confuse reliability with consistency. It is perfectly possible to give consistently incorrect answers. However, this 80% figure does give us a measure of the causal contribution of framing effects to the instability of intuitions.

²Another reason for some caution is that a number of the ‘intuitions’ which Demaree-Cotton investigates are either explicitly or implicitly comparative. Haidt and Baron (1996) ask participants for an explicitly comparative assessment of ‘An action on the part of a friend is morally worse than on the part of a stranger’. Schwitzgebel and Cushman (2012) ask their participants about two scenarios separately and the figure Demaree-Cotton examines is the proportion of participants who gave the same numerical answer to each (so measuring an ‘implicit’ comparative judgment). The reason for caution is that there is reason to think that comparative judgments like this would be more resilient to order effects than the precise strength of intuitions about single cases.

3 Intuitions are not home free

So, it seems that the extant evidence on framing effects in moral intuitions may not be too worrying where reliability is concerned; there may be no great reliability problem per se. Nonetheless, although moral intuitions may be reliable enough (for all the relevant evidence shows), I argue that extant evidence of framing effects still presents a problem for moral philosophers. Indeed, my final observation in the previous section is a sign of the nature of the problem. Intuitions come in different strengths.

Indeed, philosophers sometime make the strength of certain intuitions explicit, especially when the relevant intuitions play important roles in the relevant debates or dialectic. Here are a couple of examples (in which philosophers variously describe intuitions as strong, weak, and consider the comparative strength of intuitions – my emphasis):

There is *a strong and sensible intuition* that the laws of nature are logically contingent. (Mumford, 1998)

According to many philosophers, there is *a strong intuition* that an agent in a standard manipulation case is not free... (Sripada, 2012)

We have *a strong intuition* that a person's moral standing should not be affected by luck. (Brogaard, 2003)

...the Langton/Lewis analysis of “intrinsic” faces two problems. The first is a straight counterexample—if you believe *our weak intuition* that being accompanied by a cube is no more natural than being such that there is a cube, their analysis will make the latter intrinsic. The second is that there doesn't seem to be any particular theoretical reason to overthrow that intuition. (Marshall and Parsons, 2001)

...the strength of the reason created by (negative) desert is not constant across all offenses. The strength of this reason varies with the seriousness of the crime... This characteristic of desert accounts for *a (weak) intuition* I tend to share and that many philosophers hold to represent the ultimate test of retributive sympathies: the affront to retributive justice many of us would feel if the appropriate institution had the opportunity but failed to give Hitler or Stalin the punishments they deserve. (Husak, 2013)

I will conclude, however, by suggesting some factors that must be considered when deliberating over which of the *two intuitive beliefs* we have more reason to reject.

1. *Which belief is stronger?* Which is more difficult psychologically to give up?

(Sommers, 2009)

Nonetheless, I do think that *the intuition is much stronger* S fails to know when S possesses an intersubjectively evident defeater than when S possesses an intersubjectively opaque defeater. (Cohen, 1987)

The fact that intuitions can vary in strength is significant as the effect of framing observed in the empirical studies isn't necessarily just a matter of *what* intuition participants have. Even when participants would have had the same intuition given each framing, the framing might affect the strength of that intuition. Of course, some empirical studies cannot pick up on the strength of participant's intuitions. If a study uses a form of question which asks participants only to select 'yes' or 'no' to the question 'Is it morally permissible to X?' then the data one obtains would not tell one anything

about the strength of intuitions. If on the other hand, as is typically more common in empirical moral psychology, the study uses a form of question which asks participants to indicate to what extent they agree with the statement ‘It is morally permissible to X’ on a scale from 1 to 7, the data one obtains allows one to make some inferences about the strength of participants’ intuitions.

I am not denying that framing effects which change the polarity of one’s intuitions inspire a certain special kind of methodological worry. I think that studies often rightly make particular note of irrelevant effects which *change the modal response* (by this I mean cases where the manipulation changes which answer is the most popular) or which, for a study using a scale, move the mean response from one side of the midpoint of the scale to the other (see, e.g., [Gold et al., 2014](#)). Rather, my claim is that significant effects in participants’ responses on a scale also raise methodological worries regardless of whether the effects threaten to make one’s intuition ‘agree’ rather than ‘disagree’ or vice versa. To see why, we need to consider some facts about how intuitions factor into moral theorizing. Once we give this a moment’s thought, it should be clear that it is not only what intuitions we have (or some relevant population has) but how strong those intuitions are which factor into moral theorizing.

Suppose John has a very strong intuition that it is wrong to claim benefits to which one has no legal entitlement and a pretty weak intuition that those with certain disabilities should be provided with financial assistance by the state. Suppose that Jane has a pretty weak intuition that it is wrong to claim benefits to which one has no legal entitlement but a very strong intuition that those with certain disabilities should be provided with state help. John and Jane have the same intuitions. They find the same propositions intuitive and the same negations counterintuitive. And in both their cases, those same intuitions are likely going to factor into their theorizing when they consider what a just organization of the benefits system is going to look like. However, we can foresee that the different relative strengths of those intuitions is likely to mean that the upshot of John and Jane’s theorizing will be rather different.

Frequently in philosophical debates one can feel intuitive pulls in multiple directions. The reason moral dilemmas are called moral dilemmas is that we feel at least somewhat torn. Effects of irrelevant factors on the comparative strength of these pulls could have pretty radical effects on our final position. [Stratton-Lake \(2014\)](#) expresses a similar idea when considering evidence of framing effects on intuitions about trolley cases. He expresses doubt that there will be any participants who (a) think that one should push the man under the trolley to save the five bystanders, and yet (b) wouldn’t nonetheless agree that the fact that this act will kill someone counts against performing that action. Everyone feels the pull, that’s his point. What is important then—what makes the difference between Kantian and Utilitarian—is the way that the competing pulls stack up against each other (as well as with other considerations which may be in play).

Consequently, the influence of morally irrelevant factors on the strength of intuitions is thus very important even if it doesn’t change which intuitions one has. It is capable of radically altering the moral position which one ends up endorsing. This might not be obvious; the kinds of manipulation typically used to push people’s intuitions around in moral psychology are not generally thought to permanently alter people’s intuitions. However, even a momentary effect on the strength of intuitions can have a dramatic effect.

Suppose that, as a philosophy student, week-after-week for a whole term, you have the class on intergenerational justice in the case of climate change right after the class on reproductive ethics. Moreover, suppose this framing means that when you first encounter each topic on the course you have weaker (or stronger) intuitions about the demands of future generations on our climate policy decisions. This could radically affect how seriously you take certain objections, how persuasive you find certain arguments, or which positions you simply dismiss—in that class. And the effects of this could well be longlived, even if the direct effect of framing is not. It could affect which ideas one chooses to write on, what one chooses to work on at graduate school, and so on.

Likewise, suppose that a key paper in an important debate exploits a certain framing in order to

pump intuitions in favour of a particular position. Suppose it doesn't affect what people generally find intuitive, but merely how intuitive they find it while assessing the relevant argument. This is a framing which will potentially affect everyone involved in that particular debate and could have a significant influence on the trajectory of scholarship in that particular field, with many people taking the relevant position more seriously than they would have done had that single paper used an alternative framing.

If you are a moral philosopher, evidence that the strength of moral intuitions is sensitive to irrelevant factors should shake you considerably, whether or not which moral intuitions you have is similarly sensitive. So, what does the extant evidence say? Can we quantify the risk that the *strength* of our intuitions would be rather different given a different framing (even if the content of our intuitions would not)?

Three of the studies [Demaree-Cotton](#) examines allow us to do something close:³

1. [Nadelhoffer and Feltz \(2008\)](#) report that participants asked to state 'how much control you think [X] has over the outcome' in a standard switch trolley case had a mean response of 4.28 (on a scale from 1 to 7) when the protagonist was 'you' and 5.12 when the protagonist was 'John' ($t(83) = -2.217, p = .029$).⁴ We can get a sense of the size of this effect; from the df and t-value we can calculate Cohen's d which reveals this to be a medium sized effect $d = .49$).
2. [Liao et al. \(2011\)](#) report that participants asked about the statement 'It is morally permissible to redirect the trolley onto the second track' in a Loop trolley case had a mean response of 3.10 (on a scale from 1 to 6) when preceded by a push trolley case and 3.82 when preceded by a standard trolley case ($p = 0.029$).
3. [Wiegmann et al. \(2012\)](#) report a number of significant effects. All concern responses on a scale from 1 to 6 to 'Karl shouldn't, in terms of morality, [perform the relevant action]'. The two orders were MAF (most agreeable first) and LAF (least agreeable first). The results are in [Table 1](#). Again we can calculate effect sizes – this time from reported means and standard deviations – for all order effects observed. All are large.

Table 1: Means and effect sizes for Wiegmann et al.

Scenario	Mean (MAF)	Mean (LAF)	Cohen's d
Trap	3.24	2.12	.74
Redirect	3.84	2.52	.90
Run Over	3.84	2.52	.90
Standard	4.08	2.68	.95

We can see that whether or not these framing effects make a worrying contribution to *what* intuitions one has, they threaten to make an appreciable contribution to the *strength* of one's intuition. Given one framing, one's intuition might be fairly firm; given another, it might be pretty non-committal. This is no trivial difference given the way that intuitive judgments factor into moral theorizing.

One might resist this idea that framing could make the difference between having a strong intuition and having a fairly weak intuition. The thought might be that strong intuitions will be less

³Unfortunately, the others either (i) do not report mean ratings in relation to the order effects ([Haidt and Baron, 1996](#)), (ii) concern implicitly comparative judgments which do not represent a single answer given by participants (so participants were never asked a question which would allow them to indicate the strength of their intuition) ([Schwitzgebel and Cushman, 2012](#)), or else (iii) used 'yes/no' questions only ([Lanteri et al., 2008](#); [Tobia et al., 2013](#)).

⁴This is a different effect than the one [Demaree-Cotton](#) discusses. [Nadelhoffer and Feltz \(2008\)](#) measure moral intuitions using a 'yes/no' question.

susceptible to framing effects on strength. However, there's no positive reason to hold out hope that it is only the weak middle-of-the road intuitions which get pushed around in terms of strength. There is, it is true, some much discussed evidence that more confident/strong intuitions are more stable (Wright, 2013, 2010). Zamzow and Nichols (2009), for example, argue the defender of intuitions may find some relief in the fact that confidence ratings track the extent to which intuitions are sensitive to order effects. However, one must remember the nature of the relevant evidence. The evidence shows that confidence/strength ratings are strong predictors of stability of intuitions in the following sense: they are strong predictors of whether the case participants are considering is a case for which 'yes'/'no' answers are stable or unstable (Wright, 2010). The evidence does not show that framing effects' influence can push the strength of weak intuitions around but not that of strong intuitions. Moreover, if intuitions' susceptibility to framing effects in terms of their strength was uniform (strong intuitions were affected as much as weak intuitions), then one would predict that framing effects in terms of intuitions' polarity would vary between strong intuitions and weak intuitions (because weak intuitions need a smaller change in order to push them over the midline).

4 Final remarks

Before closing, let me make three final remarks.

Evaluative methodology needs to keep descriptive methodology in mind Because debates about the epistemic standing of intuitions have quite understandably been parsed in terms of reliability, it has been easy for efforts in evaluative methodology to focus on an overly simplified model of the way that intuitions factor into moral theorizing. There is an important sense in which descriptive methodology is properly considered prior to evaluative methodology (as I have argued elsewhere Andow, 2015b). There is a danger that the implications for the evaluation of philosophical methods of empirical studies on moral intuitions and indeed other intuitions will not be properly understood if the working model of intuitions' role in philosophy is one according to which it is only *which* intuitions participants have that is relevant.⁵

A suggestion about how to use intuitions unproblematically My main aim in this paper has been to argue that framing effects might show moral intuitions to be problematic even if they don't show moral intuitions to be problematically unreliable. I do not want to paint a too pessimistic picture of the use of intuitions in moral philosophy, however. The reason is as follows. In the event that intuitions are pretty *reliable* but have a *strength problem* there is a natural ameliorative strategy which suggests itself. The strength problem stems from facts about the ways that philosophers actually use intuitions. However, philosophers could change the ways that they used intuitions. There might be better ways to use intuitions which are available. Were it the case that intuitions were unreliable as well, there would be no obvious solution. However, suppose that Demaree-Cotton (2015) is correct that intuitions are not problematically unreliable. If this is the case, we can trade on the fact in order to avoid the strength problem. So the suggestion is this: philosophers stop giving weight to the strength of their intuitions and just rely on the answers that intuition endorses regardless of the strength.

There are things we can say to make this sound a little more palatable, e.g, (i) philosophers do not need to rely only on their own intuitions – one can still have a measure of *how intuitive* a certain proposition is by using experimental ethics which means that one is not at an evidential stalemate every time that intuitions apparently conflict; and (ii) another way to break stalemates is to ask whether

⁵Indeed, as I have argued elsewhere, recent work in descriptive methodology suggests that the role of intuition in philosophy seems likely not to be restricted to evidential use but also play an important explanatory role which raises new issues in evaluative methodology (Andow, 2015b).

either of the conflicting intuitions can be explained away, or whether reasons can be given for assigning a certain intuition a lesser weight in the theory choice.

A note of caution Even though there is this obvious ameliorative strategy available, there might be some reason to worry that we are not up to the task of putting that ameliorative strategy into action. Although philosophers do sometimes appeal to intuitions and explicitly appeal to their being ‘particularly strong’, most of the time the ways that the strength of intuitions influences philosophers is more subtle and not part of the explicit reasoning process. It may be that we are rarely consciously aware of factoring in the strength of intuitions. It may be that the weighing up of intuitions of different strengths takes place with a certain level of automaticity, meaning that it is difficult to rein in. Nonetheless, even if we are not able to completely bring this use under control, it is still a step forward to exercise greater caution when using intuition in guiding our moral theorizing (if not about the direction in which intuition points, then about the vigour with which it points).

Acknowledgements

Thanks to two referees for this journal for swift and helpful comments.

References

- Andow, J. (2015a). Expecting moral philosophers to be reliable, *Dialectica* 69(2): 205–220.
- Andow, J. (2015b). How distinctive is philosophers’ intuition talk?, *Metaphilosophy* 46(4–5): 515–538.
- Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making, *Cognition* 108(2): 381–417.
- Brogaard, B. (2003). Epistemological contextualism and the problem of moral luck, *Pacific Philosophical Quarterly* 84(4): 351–370.
URL: <http://dx.doi.org/10.1046/j.1468-0114.2003.00178.x>
- Cohen, S. (1987). Knowledge, context, and social standards, *Synthese* 73(1): 3–26.
- Demaree-Cotton, J. (2015). Do framing effects make moral intuitions unreliable?, *Philosophical Psychology* online first: 1–22.
URL: <http://dx.doi.org/10.1080/09515089.2014.989967>
- Gold, N., Colman, A. M. and Pulford, B. D. (2014). Cultural differences in responses to real-life and hypothetical trolley problems, *Judgment and Decision Making* 9(1): 65–76.
- Gonnerman, C., Reuter, S. and Weinberg, J. (2011). More oversensitive intuitions: Print fonts and could choose otherwise.
- Haidt, J. and Baron, J. (1996). Social roles and the moral judgement of acts and omissions, *European Journal of Social Psychology* 26(2): 201–218.
- Husak, D. (2013). Retributivism in extremis, *Law and Philosophy* 32(1): 3–31.
URL: <http://dx.doi.org/10.1007/s10982-012-9145-2>
- Lanteri, A., Chelini, C. and Rizzello, S. (2008). An experimental investigation of emotions and reasoning in the trolley problem, *Journal of Business Ethics* 83(4): 789–804.

- Liao, S. M., Wiegmann, A., Alexander, J. and Vong, G. (2011). Putting the trolley in order: Experimental philosophy and the loop case, *Philosophical Psychology* 25(5): 661–671.
- Lombrozo, T. (2009). The role of moral commitments in moral judgment, *Cognitive Science* 33(2): 273–286.
- Marshall, D. and Parsons, J. (2001). Langton and Lewis on 'intrinsic', *Philosophy and Phenomenological Research* 63(2): 347–351.
- Mumford, S. (1998). Laws of nature outlawed, *Dialectica* 52(2): 83–101.
URL: <http://dx.doi.org/10.1111/j.1746-8361.1998.tb00043.x>
- Nadelhoffer, T. and Feltz, A. (2008). The actor–observer bias and moral intuitions: Adding fuel to Sinnott-Armstrong's fire, *Neuroethics* 1(2): 133–144.
- Nahmias, E., Coates, D. J. and Kvaran, T. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions, *Midwest Studies in Philosophy* 31: ?
- Nichols, S. and Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions, *Nous* 41: 663–685.
- Nichols, S. and Mallon, R. (2006). Moral dilemmas and moral rules, *Cognition* 100(3): 530–542.
- Petrinovich, L. and O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions, *Ethology and Sociobiology* 17(3): 145–171.
- Schwitzgebel, E. and Cushman, F. (2012). Expertise in moral reasoning? order effects on moral judgment in professional philosophers and non-philosophers, *Mind & Language* 27(2): 135–153.
- Schwitzgebel, E. and Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection, *Cognition* 141: 127–137.
- Sinnott-Armstrong, W. (2008). Framing moral intuitions, in W. Sinnott-Armstrong (ed.), *Moral Psychology*, Vol. 2: The Cognitive Science of Morality, MIT Press, pp. 47–76.
- Sommers, T. (2009). More work for hard incompatibilism¹, *Philosophy and Phenomenological Research* 79(3): 511–521.
URL: <http://dx.doi.org/10.1111/j.1933-1592.2009.00291.x>
- Sripada, C. S. (2012). What makes a manipulated agent unfree?, *Philosophy and Phenomenological Research* 85(3): 563–593.
URL: <http://dx.doi.org/10.1111/j.1933-1592.2011.00527.x>
- Stratton-Lake, P. (2014). Intuitionism in ethics, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2014 edn.
- Swain, S., Alexander, J. and Weinberg, J. (2008). The instability of philosophical intuitions: Running hot and cold on truetemp, *Philosophy and Phenomenological Research* 76(1): 138–155.
- Tobia, K. (2012). Expert intuition, *38th Meeting of the Society for Philosophy and Psychology*. University of Colorado at Boulder.
- Tobia, K., Buckwalter, W. and Stich, S. (2013). Moral intuitions: Are philosophers experts?, *Philosophical Psychology* 26: 629–638.

- Weinberg, J., Alexander, J., Gonnerman, C. and Reuter, S. (2012). Restrictionism and reflection: Challenge deflected or simply redirected?, *The Monist* 95(2): 200–222.
- Wiegmann, A., Okan, Y. and Nagel, J. (2012). Order effects in moral judgment, *Philosophical Psychology* 25(6): 813–836.
URL: <http://dx.doi.org/10.1080/09515089.2011.631995>
- Wright, J. (2013). Tracking instability in our philosophical judgments: Is it intuitive?, *Philosophical Psychology* 26(4): 485–501.
URL: <http://dx.doi.org/10.1080/09515089.2012.672172>
- Wright, J. C. (2010). On intuitional stability: The clear, the strong, and the paradigmatic, *Cognition* 115(3): 491–503.
- Zamzow, J. L. and Nichols, S. (2009). Variations in ethical intuitions, *Philosophical Issues* 19(1): 368–388.