

This is an Accepted Manuscript of an article published by Taylor & Francis in *Philosophical Psychology*, first published online on 12 December 2014, available at: .

<http://www.tandfonline.com/10.1080/09515089.2014.989967>

Citation information: Demaree-Cotton, J. (2016). Do Framing Effects Make Moral Intuitions Unreliable? *Philosophical Psychology*, 29, 1-22.

DO FRAMING EFFECTS MAKE MORAL INTUITIONS UNRELIABLE?

Joanna Demaree-Cotton

1. Introduction: Framing Effects in Moral Psychology

The term ‘framing effects’ can be traced to the work of Tversky & Kahneman (e.g. 1981) and their work on the psychology of decision-making which showed that ways of ‘framing’ options affect people’s choice behaviour without actually altering any relevant information about the options presented to them. Similarly, framing effects in *moral* psychology refer to evidence that morally irrelevant differences in the way a scenario is presented affect people’s moral intuitions regarding that scenario.

For example, if presented with the famous “Switch” version of the Trolley Problem, where a runaway trolley is heading towards five workmen, subjects are more likely to agree with pulling a lever so that the trolley switches to a side-track with one person on it when that same action is framed in terms of *saving* the five workmen than when that action is framed in terms of *killing* the one on the side-track. This is so even though no information about the scenario is altered and even though all subjects are aware that the action will result in five living and one dying (Petrinovich & O’Neill, 1996). What is crucial is that some apparently morally irrelevant factor, such as wording or the order in which different scenarios are presented, changes people’s intuitions.¹

Over the last two decades, a number of philosophers and psychologists have argued that the justificatory status of moral intuitions are undermined precisely because they are subject to framing effects.² I will focus in particular on Sinnott-Armstrong's argument (2008a, 2008b; see also 2011). According to Sinnott-Armstrong, moral intuitions which are subject to framing effects are unreliable and, therefore, they are not justified without independent inferential confirmation, that is, they are not noninferentially justified (I follow Sinnott-Armstrong's treatment of intuition as a type of belief). Furthermore, he takes the empirical evidence demonstrating particular framing effects on moral intuitions to support the claim that our set of moral intuitions in general are subject to framing and are therefore similarly unreliable and not noninferentially justified.³

He lays out his argument as follows (2008a, p.52):

1. If our moral intuitions are formed in circumstances where they are unreliable, and if we ought to know this, then our moral intuitions are not justified without inferential confirmation.
2. If moral intuitions are subject to framing effects, then they are not reliable in those circumstances.
3. Moral intuitions are subject to framing effects in many circumstances.
4. We ought to know (3).
5. Therefore, our moral intuitions in those circumstances are not justified without inferential confirmation.

His argument, if successful, undermines *moral intuitionism*, according to which moral intuitions are noninferentially justified. Different intuitionist theories differ with regards to the source of this justification. For example, according to Tolhurst's (1990, 1998) "experientialist" intuitionism, moral intuitions are noninferentially justified when they strike

us as true in the absence of a reason to believe they are false; Shafer-Landau's intuitionism (2003) holds that moral intuitions are noninferentially justified simply by being the product of a reliable process; and Audi's view (2004) is that moral intuitions are noninferentially justified when they follow certain sorts of reflection. Sinnott-Armstrong's argument would undermine all such views. Worryingly, it follows from the failure of moral intuitionism that our everyday moral intuitions are unjustified unless we are able to infer them from other beliefs.

I will accept Sinnott-Armstrong's central epistemological claims (premise 1) for the sake of this paper. However, I wish to call into question premises 2 and 3, thereby calling into doubt Sinnott-Armstrong's conclusion regarding the justification of moral intuitions.

First I will discuss the nature of epistemic unreliability, and then analyse how framing effects render moral intuitions unreliable. This analysis calls for a modification of premise 2, which in turn requires a modification of premise 3 if the argument is to remain valid. I then re-examine the empirical evidence which is supposed to support the modified premise 3. In doing so, I provide a novel suggestion for how to analyse the strength of framing effects in empirical studies as this pertains to unreliability. My analysis of the empirical evidence leads me finally to suggest that it is unclear that it provides adequate support for the modified premise 3 in order for the argument to go through. I therefore show that Sinnott-Armstrong has failed to demonstrate that our moral intuitions are not noninferentially justified.

2. Framing and Epistemic Reliability

2.1. Unreliable Beliefs: Likelihood of Error

Belief-forming processes are unreliable, and the resultant beliefs not epistemically justified, if there is a sufficiently large likelihood that the beliefs are in error (e.g. Goldman, 1986).⁴ Sinnott-Armstrong claims that moral intuitions are unreliable in this way.

Few epistemologists require *perfect reliability* for epistemic justification. If we were only justified in holding perfectly reliable beliefs, most of our everyday beliefs would not be justified. For example, most of our perceptual beliefs would not be justified, since our perceptual systems are subject to a certain margin of error and therefore do not generate *perfectly* reliable judgments. This sort of scepticism is generally seen as sufficient to reject a perfect reliability requirement on epistemic justification (see e.g. Cohen, 1984, or Klein, 2013).

This suggests that when discussing the epistemic reliability of beliefs we are concerned with whether a belief is *sufficiently likely to be in error so as to defeat epistemic justification*. It is difficult to demarcate precisely what level of error is sufficiently unreliable to defeat epistemic justification. Sinnott-Armstrong refers to those processes which are “likely to lead to error” or which result in beliefs which “will often be false” (2008a, pp.50–51), and he emphasises that epistemic justification is defeated just in case “it is reasonable for a person to assign a *large probability that a certain belief is false*” (2008b, p.98, my emphasis). Although Sinnott-Armstrong explicitly states that he does not need nor want to commit to an exact cut-off (2008b, p.101) he appeals to intuitive examples that “a process that is inaccurate at least half of the time” (2008a, p.53) is not reliable, and “a belief is not epistemically justified if it is reasonable for the believer to assign a probability of error as high as .45” (2008b, p.100).⁵

So, to sum up this section: a belief is unreliable, and its justification is defeated, if it is sufficiently likely to be in error.

2.2. *Unreliable Beliefs: The Moral Domain*

To evaluate the reliability of moral beliefs, we need to identify when moral belief-forming processes are likely to lead to error. This is no straightforward matter, for we do not

have broad consensus as to which moral propositions are true. However, by appealing to cases of inconsistency we can unambiguously identify some cases of error. This is because if someone has inconsistent beliefs—such as the belief that p and the belief that $\neg p$ —at least one of those beliefs must be in error. Equally, if someone is *prone* to producing beliefs which are inconsistent with each other, the belief-generating process responsible for these beliefs is unreliable.

Belief-forming processes subject to framing effects are modally inconsistent in the following way. When people make moral claims, we take them to commit, at least implicitly, to a further set of moral beliefs according to a particular generalization whereby they hold equivalent moral beliefs regarding all relevantly similar scenarios⁶ i.e. those in which there is no morally relevant difference. This is because if there are no morally relevant differences the truth-value of a particular moral proposition must remain constant. Take the aforementioned example of the “kill” vs. “save” framing effect in the trolley problem. Let us call the “kill” framing of the moral scenario M_K and the “save” version M_S . We take someone who judges “Turning the trolley is wrong in M_K ” to commit themselves to “Turning the trolley is wrong in M_{KI} ” where M_{KI} is exactly like M_K except that in M_{KI} the scenario is described in a Scottish accent. This is because the accent in which the scenario is described is obviously irrelevant to the truth about the wrongness of turning the trolley. Similarly, we take someone who judges “Turning the trolley is wrong in M_K ” to be committed to judging “Turning the trolley is wrong in M_S ,” because how a moral scenario is worded is morally irrelevant.⁷ However, there is evidence that “kill” framing tends to lead to the judgment that turning the trolley is wrong but “save” framing is more likely to lead to the judgment that it is right. If this is the case, we might expect a certain subject S to judge “Turning the trolley is wrong in M_K ” if she is presented with “kill” framing, thereby committing herself to:

“Turning the trolley is wrong in M_S ”

But we might also have expected her to judge:

“Turning the trolley is not wrong in M_S ”

had she in fact been presented with the “save” framing. S 's belief-forming process is thus modally inconsistent since the two expected beliefs regarding turning the trolley in M_S are inconsistent.⁸ This inconsistency means that S can only be correct in one out of two of the framing conditions, and we said above that being inaccurate “at least half of the time” seems too high an error rate.⁹ The process is thus unreliable and the resultant beliefs not justified.

Given this understanding of unreliability in framing cases, I want to stipulate the following technical sense in which a belief may be *determined by a frame*:

A subject's belief is *determined* by a frame just in case *they would have had a different belief, inconsistent with the one actually formed, had nothing but the frame changed.*

And we should additionally say:

A subject's belief is *unreliable* if there is a *sufficiently high likelihood* that their belief is determined by a frame.

For example, if someone in fact judges that it is wrong to turn the trolley when they are presented with the “kill” version of the scenario, but there is a sufficiently high likelihood that they *would have judged that it isn't wrong to turn the trolley if they had been presented with the “save” version*, their judgment is unreliable.

3. Amending Sinnott-Armstrong's Argument

The following points have come out of the last two sections. Firstly, a belief is unreliable—and not epistemically justified—if there is a *sufficiently high probability* of error.

Secondly, we can take there to be a sufficiently high probability of error for a moral belief when there is a sufficiently high probability that it was *determined* by a frame.

If this is right, then, in order to establish that moral intuitions are unreliable, it is not sufficient to show merely that moral intuitions are *subject* to framing effects. Rather, one must also show that framing effects are a powerful enough force in the belief-forming process such that they are sufficiently likely to *determine* the resulting content of moral intuitions.

This suggests that Sinnott-Armstrong's Premise 2, that:

(2) If moral intuitions are subject to framing effects, then they are not reliable in those circumstances.

needs qualification. Based on the above lines of argument, I suggest:

(2*) If moral intuitions are *sufficiently likely to be determined by framing effects* then they are not reliable in those circumstances.

If this modification is accepted, we must in turn modify premise 3 if the argument is to remain valid. The premise:

(3) Moral intuitions are subject to framing effects in many circumstances.

must be modified to:

(3*) Moral intuitions are *sufficiently likely to be determined by framing effects* in many circumstances.

Is there a large likelihood that intuitions are determined by framing? Sinnott-Armstrong writes that “[e]vidence of framing effects makes it reasonable for informed moral believers to assign *a large probability of error* to moral intuitions in general” (2008b, p.99, my emphasis). He goes on (p.101): “What shows that the probability of error in moral

intuitions is too large to meet an appropriate standard [of justification] is the size and range of framing effects in the studies.” However, nowhere in his discussion of the framing studies does he actually provide evidence concerning the size of the probability that moral intuitions are affected by framing, let alone that the probability is “large.” He only assumes it. But the size of the probability of error is crucial, for in order for justification to be defeated he must show that framing effects are sufficiently likely to determine people’s moral beliefs.

4. Operationalizing Unreliability Due to Framing Effects

I assume that Sinnott-Armstrong did not attempt to extract the probability of error because it is not *prima facie* obvious how to identify the probability of error from the evidence provided by the studies.

However, I have suggested that a belief is unreliable if there is a sufficiently high probability that the belief is determined by the frame; in other words, there is a sufficiently high probability that someone’s belief produced under Frame F1 with regards to moral scenario S1 is inconsistent with the belief they would have had under Frame F2 with regards to moral scenario S2 where the only difference between S1 and S2 is the morally irrelevant frame.

How can we know the probability that a belief is determined by framing? Framing experiments typically report the results of two conditions, where the only difference between the two conditions is some frame. Framing effects, when they occur, are reported in terms of a numerical difference between the responses given by subjects in each condition. For example, the proportion of subjects rating a particular action as right is $x\%$ in one condition, but only $y\%$ in the other condition. When the difference between these two results is statistically significant, we can infer that the difference was causally affected by the framing and therefore that that particular judgment is subject to a framing effect.

Say that an experiment shows that under Frame A, 90% of subjects rate an action as right, but under Frame B, only 20% of subjects do so. The difference in the proportion of people rating the action as right under each frame is 70%. This example is illustrated in Figure 1.

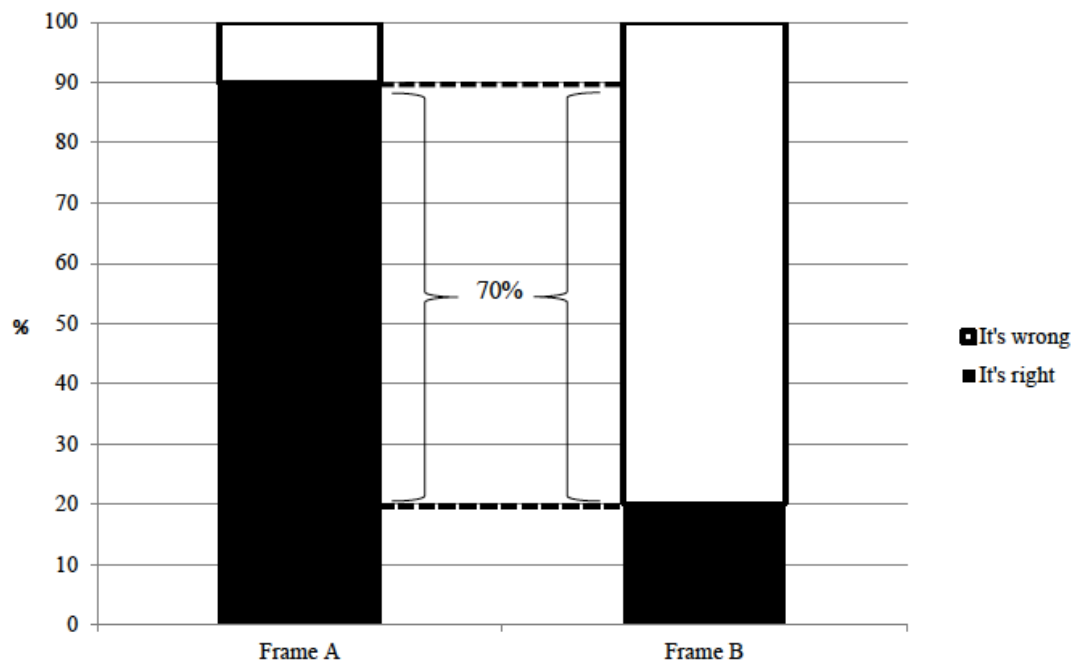


Figure 1. Fictitious example of a framing effect on the intuition that a particular action is right or wrong.

This means that, for every 100 people who were in fact subjected to Frame A, we would expect 70 of them to have given the opposite judgment if they had instead been subjected to Frame B. In other words, for 70 of them we expect their judgment to be determined by the frame. This is represented by the area of the graph in between the two thick black lines. We would only expect 30 of them to have made the same judgment irrespective of the frame (represented by the areas of the graph above the top thick black line and below the bottom thick black line). And the same goes the other way around: for every 100 subjects that were in fact subjected to Frame B, we could expect 70 of them to have given the opposite judgment had they instead been subjected to Frame A.

This means that if we pick a subject who is making such a judgment under frame A or B at random, we can say there is a 70% chance that this subject's response would have been different if we had changed the frame. So if *I* am a participant being subjected to either frame, and I make a judgment concerning whether or not the action is right, there is a 70% chance I would have made a different judgment—inconsistent with the one I in fact made—just in case some morally irrelevant factor about framing had changed. This looks like pretty bad news; it is more likely that my judgment would have changed due to framing than not. In the terminology of the reformulated premises 2* and 3*, there is a 70% likelihood that my belief is being determined by a framing effect. It would seem fair in this case to conclude that my belief-forming process regarding this sort of judgment in this sort of context is sufficiently unreliable so as to defeat epistemic justification for that belief.

A brief aside on a mathematical simplification: it is true that for any subject picked *at random* there is a 70% chance they would have given a different judgment had the frame been different. However (as readers familiar with Bayesian statistics will notice), the probability that your judgment is determined by a frame in fact depends on which frame you are exposed to and which judgment you make. To illustrate: we can say that Frame A favours “it's right” judgments in the sense that it tends to cause more respondents to make that judgment. The same goes for Frame B and “it's wrong.” For the responders who give judgments which are *not* favoured by the particular frame to which they are exposed—i.e. the *Frame A-Wrong* responders and the *Frame B-Right* responders—we don't actually expect there to be *any* chance that they would have changed their response had the frame been the one which *does* favour that response. Conversely, if you have given a response which *is* the one favoured by the frame—i.e. if you are in group *Frame A-Right* or *Frame B-Wrong*—there is an even higher probability that your response is being determined by the frame than the 70% average.

However, the *mean* probability across all responders is 70%. And I suggest that, for the purposes of drawing conclusions about the reliability of moral judgments *in general*, it is appropriate for us to treat any particular individual's placement in the range of responders as random and thus to assign them the average probability value. This is because, for any person who may at some point give a judgment exposed to a potential framing effect, we cannot say ahead of time which group they will fall in to. Moreover, outside of the lab it is highly unlikely that at any given time that a person makes a moral judgment that they, or anyone else, will be aware of a) whether a factor which may exert a framing effect is present, b) what type of frame is present (e.g. Frame A or Frame B) and c) whether the person is making a judgment which is favoured by that frame or not favoured.

I will now take a closer look at the studies reporting evidence of framing effects. I will outline the evidence concerning the probability that our moral beliefs are determined by frames using the method just outlined. I will then assess whether the studies do indeed support premise (3*) by making "it reasonable... to assign a large probability of error to moral intuitions," "too large to meet an appropriate standard" of justification.

5. An Analysis of the Evidence: How Unreliable are Moral Intuitions?

I will go through one set of results from Haidt & Baron (1996) to give an idea of the type of findings we are dealing with, and then I will present a concise summary of results from other framing studies.

Haidt & Baron's first experiment used a moral scenario concerning the selling of a Mazda. In the story, Nick wants to sell his 1984 Mazda MPV. There are 40,000 miles on it. He knows that due to a manufacturing defect particular to that year, a third of 1984 Mazda MPV's break down when they get to 50,000 miles. Kathy, who is considering buying the car,

asks Nick which year was the one which produced defective cars. Nick decides to lie, and he tells her it was 1983, not 1984. Kathy buys the car for \$4700.

Three versions of the scenario were presented to participants: one where Nick and Kathy are strangers, one where they are acquaintances, and one where they are friends. Half of the participants were presented with the scenarios in the following order: strangers, then acquaintances, then friends. Call this Mazda Order 1. The other participants received the stories in the opposite order: friends, acquaintances, strangers. Call this Mazda Order 2. Immediately after each version of the story, subjects were explicitly asked to give moral evaluations by rating Nick's behaviour on a scale where 0 represented "morally neutral" and - 100 represented "the most immoral thing a person could do." These ratings were used to determine the following crucial variable: do subjects think Nick's behaviour is worse if he is a friend compared to if he is a stranger, or not? If a subject gave a lower rating on the scale in the friend condition than they did in the stranger condition, the result was recorded as "yes"; the result was "no" if they gave the same rating for the friend compared to the stranger condition.

Subjects tended to consider Nick's behaviour as more immoral if he was Kathy's friend compared to when they were strangers. But this comparison was affected by the order in which the cases were presented. In Mazda Order 1, 88% of subjects rated Nick's behaviour as morally worse when he was a friend compared to a stranger. However, in Order 2, only 66% of subjects did so i.e. a greater proportion of subjects thought that both cases were equally immoral.

The upshot is that ordering effects may determine whether or not someone believes that an action is morally worse if it is performed against a friend compared to against a stranger.

Given the evidence, *how* unreliable are such judgments rendered by ordering effects? The percentage of participants that judged that an action against a friend is morally worse changed by 22% across each ordering condition. This means that if a person makes a judgment regarding the moral comparability of an action against a friend vs. a stranger, we can say that there is a 22% chance that their judgment was determined by ordering; i.e. there is a 22% chance they would have given a different judgment had certain irrelevant facts about ordering changed. This also means that there is a 78% chance they would not have given a different judgment had the ordering changed.

Table 1 details results from a range of psychological studies concerning framing effects on moral judgments. All of the studies concern the framing of moral scenarios or dilemmas, and the effects of those frames on moral judgments concerning the scenario or judgments concerning general moral principles. The table states: the moral judgment at issue; percentage of participants who express belief in that judgment under one framing condition F1; percentage of participants who express belief in that judgment under a second framing condition F2; the resulting probability we can assign to a random participant that they would have expressed a *different* belief had they been subjected to a different frame (“% Dif.”); and the resulting probability that a randomly selected participant would have expressed the *same* belief even if they had been subjected to the other frame (“% Same”).

I omit framing effects if the authors did not, unfortunately, report the results in such a way that made it possible to extract the requisite numerical values (e.g. Lombrozo, 2009; Wiegmann & Waldmann, 2014). I also exclude studies which, rather than asking subjects to form a moral judgment, simply asked them to make a choice, how they would act, about their preferences, or how someone should act without disambiguating pragmatic from moral considerations (e.g. Tversky & Kahneman, 1981; Petrinovich & O’Neill, 1996). Although the study of framing in decision-making is an important and interesting area, it cannot directly

bear evidence on the question of the reliability of specifically moral beliefs; to ask participants to choose a course of action is a different matter from asking them to evaluate its moral status (Christensen & Gomila, 2012) and it is likely that these questions are processed differently (Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006). Other than studies which I have excluded for reasons stated here, I have included all relevant studies of which I am aware. I include cases where *no* statistically significant evidence of a framing effect was found by inserting “n/a” in place of statistical values.

Table 1
Framing effects on moral judgments

Study/ Experiment	Frame Type, F1/F2	Judgment	Results, F1%/F2 %	% Dif	% Same
Haidt & Baron, 1996:					
Exp. 1, Mazda	Order, stranger first/friend first.	An action on the part of a friend is morally worse than on the part of a stranger.	88/66	22	78
	Order, omission first/act first.	An act is morally worse than an omission.	80/50	30	70
Exp. 1, Crane	Order, subordinate first/authority first.	An action on the part of an authority is morally worse than on the part of a subordinate.	78/56	22	78
Exp. 2, Crane	Order, omission first/act first.	An act is morally worse than an omission.	66/39	27	73
Schwitzgebel & Cushman, 2012:¹⁰					
Double Effect dilemmas, Q1–2	Order, means cases first/ side effect cases first.	Killing the one as a side effect to save the five is equally morally good/bad to killing the one as a means.	70/54	16	84
Double Effect dilemmas, Q14–17	Order, means cases first / side- effect cases first.	Killing the one as a side effect to save the five is equally morally good/bad to killing the one as a means.	73/60	13	87
Oxygen Dilemma: Non-Ac	Order, act first/omission first.	The act is equally morally good/bad as the omission.	21/37	16	84

Ac Non-Phil	Order, act first/omission first.	The act is equally morally good/bad as the omission.	18/34	16	84
Phil	Order, act first/omission first.	The act is equally morally good/bad as the omission.	17/29	12	88
Eth PhD	Order, act first/omission first.	The act is equally morally good/bad as the omission.	20/29	n/a	n/a
Vest Dilemma:					
Non-Ac	Order, act first/omission first.	The act is equally morally good/bad as the omission.	44/33	11	89
Ac Non-Phil	Order, act first/omission first.	The act is equally morally good/bad as the omission.	42/23	19	81
Phil	Order, act first/omission first.	The act is equally morally good/bad as the omission.	38/25	13	87
Eth PhD	Order, act first/omission first.	The act is equally morally good/bad as the omission.	38/17	21	79
Moral Luck dilemmas, Q10–11	Order, good luck first/bad luck first.	The morally lucky act is equally morally good/bad to the morally unlucky act.	68/60	8	92
Moral Luck dilemmas, Q12–13 ¹¹	Order, good luck first/bad luck first.	The morally lucky act is equally morally good/bad to the morally unlucky act.	62/66	4	96
Moral Luck Principle:					
Non-Ac	Order, good luck first/bad luck first.	People who do the same thing but have different outcomes by chance should receive different amounts of punishment.	20/20	n/a	n/a
Ac Non-Phil	Order, good luck first/bad luck first.	People who do the same thing but have different outcomes by chance should receive different amounts of punishment.	18/18	n/a	n/a
Phil	Order, good luck first/bad luck first.	People who do the same thing but have different outcomes by chance should receive different amounts of punishment.	29/45	16	84
Eth PhD	Order, good	People who do the same thing	38/38	n/a	n/a

	luck 1st/bad luck 1st.	but have different outcomes by chance should receive different amounts of punishment.			
Act- Omission Principle, Q19:					
Non-Ac	Order, Take Oxygen or Not Give Vest first/ Not Give Oxygen or Take Vest first.	Saving people by actively killing one person is morally worse than by allowing one person to die.	52/44	8	92
Ac Non- Phil	Order, Take Oxygen or Not Give Vest first/ Not Give Oxygen or Take Vest first.	Saving people by actively killing one person is morally worse than by allowing one person to die.	59/53	n/a	n/a
Phil	Order, Take Oxygen or Not Give Vest first/ Not Give Oxygen or Take Vest first.	Saving people by actively killing one person is morally worse than by allowing one person to die.	54/58	n/a	n/a
Eth PhD	Order, Take Oxygen or Not Give Vest first/ Not Give Oxygen or Take Vest first.	Saving people by actively killing one person is morally worse than by allowing one person to die.	56/56	n/a	n/a
Doctrine of Double Effect Principle, Q20:					
Non-Ac	Order, side- effect scenarios first/means scenarios first.	Killing one as a means to saving several is morally worse than killing one as a side effect in saving several.	46/53	7	93
Ac Non- Phil	Order, side- effect scenarios first/means scenarios first.	Killing one as a means to saving several is morally worse than killing one as a side effect in saving several.	51/55	n/a	n/a
Phil	Order, side- effect scenarios first/means	Killing one as a means to saving several is morally worse than killing one as a	62/46	16	84

Eth PhD	scenarios first. Order, side-effect scenarios first/means scenarios first	side effect in saving several. Killing one as a means to saving several is morally worse than killing one as a side effect in saving several.	59/40	n/a	n/a
Tobia, Buckwalter & Stich, 2013:					
Jim and the Indians:					
Non-Phil	Wording, “you”/“Jim”.	One is morally obligated to shoot the native to save the others.	19/53	34	66
Philosophers	Wording, “you”/“Jim”.	One is morally obligated to shoot the native to save the others.	36/9	27	73
Trolley: Switch	Wording, “you”/“John”.	It is morally permissible to kill one to save five.	89/64	25	75
Nadelhoffer & Feltz, 2008:					
Trolley: Switch	Wording, “you”/“John”.	It is morally permissible to kill one to save five.	65/90	25	75
Liao, Wiegmann, Alexander, & Vong, 2012:					
Trolley: Loop	Context, Footbridge first/Switch first. ¹²	Disagree with: “It is morally permissible to redirect the trolley onto the second track”.	56/34	22	78
Lanteri, Chelini, & Rizzello, 2008:					
Trolley: Switch vs. Footbridge	Order, Switch first/Footbridge first. ¹³	Turning the trolley in Switch is morally obligatory.	34/11	23	77
Trolley: Switch vs. Footbridge	Order, Switch first/Footbridge first.	Turning the trolley in Switch is morally acceptable.	94/78	16	84
Trolley: Switch vs. Footbridge	Order, Switch first/Footbridge first.	Pushing the man in Footbridge is morally obligatory.	3/7	n/a	n/a
Trolley: Switch vs. Footbridge	Order, Switch first/Footbridge first.	Pushing the man in Footbridge is morally acceptable.	46/48	n/a	n/a

Wiegmann, Okan, & Nagel, 2012:					
Trolley, Switch	Order, most agreeable first/least agreeable first.	Karl shouldn't, in terms of morality, turn the trolley.	32/68	36	64
Trolley, Footbridge	Order, most agreeable first/least agreeable first.	Karl shouldn't, in terms of morality, push the man.	76/80	n/a	n/a
Trolley, Trap	Order, most agreeable first/least agreeable first.	Karl shouldn't, in terms of morality, open the trapdoor.	52/80	28	72
Trolley, Redirect	Order, most agreeable first/least agreeable first.	Karl shouldn't, in terms of morality, turn the trolley.	40/72	32	68
Trolley, Run Over	Order, most agreeable first/least agreeable first.	Karl shouldn't, in terms of morality, turn the trolley.	32/72	40	60

As can be seen from this hasty survey of the literature, the numbers that come up fall within a relatively consistent range. In particular, the probability that participants would *not* have expressed a different belief had the frame been different is relatively high, clustering predominantly in the eighties and seventies to give an average of 80%.¹⁴ The spread of values, excluding “n/a” cases where no framing effect was found (which can actually be imagined as approaching 100) is shown in Figure 2.

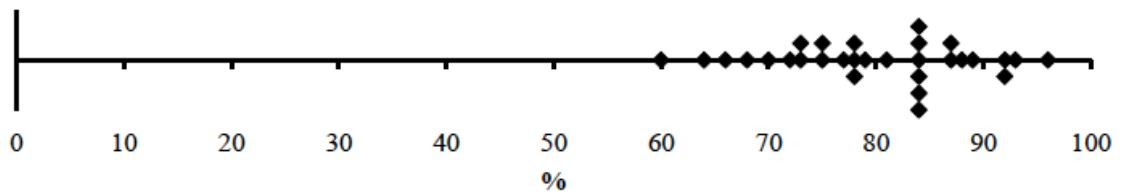


Figure 2. Scatter plot showing the range of probabilities from different studies that subjects would have made the same judgment regardless of which frame they were exposed to.

As mentioned earlier, Sinnott-Armstrong takes it as intuitive that “a belief is not epistemically justified if it is reasonable for the believer to assign a probability of error as high as .45” (2008b, p.100). In other words, 55% reliable is not good enough.

But what about 80% reliable? Exactly half of these values, even when taken only from cases where framing effects *are* significant, are higher than 80%. 96%—the highest value—is *surely* sufficiently reliable. Although my intuitions here are admittedly not very strong, I am inclined to suggest that 80% is good enough for at least initial, defeasible epistemic justification. What about 70–80% reliable (where the majority of the rest of the data points are located)? The intuition is even less certain here. But it is at least not *obvious* whether this is insufficiently reliable for epistemic justification. It is certainly significantly higher than the 55% reliable suggested by Sinnott-Armstrong as a purportedly uncontroversial case of insufficient reliability. I therefore contend that he has failed to show that these moral intuitions are unreliable.¹⁵

Furthermore, even if these particular intuitions *were* taken to be too unreliable, they don’t provide the evidential basis from which Sinnott-Armstrong can infer that our moral intuitions in general are unreliable and unjustified. This is because the experimental literature on framing effects gives us some reason to think that the effects discussed in this paper are particularly strong, and that many other sorts of moral intuitions would be *even less*

susceptible to framing. Firstly, regarding moral judgments in particular, research is beginning to show that certain types of judgments are immune from certain types of framing effects. For example, only certain versions of the trolley problem, such as the classic “Switch” version, are susceptible to order effects, and even then only when preceded by scenarios with similar content (Petrinovich & O’Neill, 1996). Other versions of the trolley problem, such as “Footbridge” (where the only way of stopping the trolley from killing the five is to push a very large man off of a footbridge into the trolley’s path), are immune from order effects (Lanteri, Chelini, & Rizzello, 2008; Wiegmann, Okan, & Nagel, 2012; Wiegmann & Waldmann, 2014). This may be because only causally ambiguous moral scenarios (Wiegmann & Waldmann, 2014) which are normally regarded as impermissible (Wiegmann, Okan & Nagel, 2012) are amenable to order effects.¹⁶

Secondly, the general psychological literature on framing effects outside of the moral domain indicates that framing effects are mitigated or abolished when the case at hand involves strongly held attitudes (e.g. abortion decisions), high personal involvement (e.g. positive/negative framing of one’s own performance, as opposed to another’s performance), or extreme rather than intermediate attributes (e.g. where a product to be chosen is in very bad condition, rather than in somewhat poor condition) (see Levin, Schneider, & Gaeth, 1998, for a review). But many of the studies conducted so far, and which have been cited here, concern moral scenarios which are both highly out of the ordinary and where it is very difficult to see what the correct moral judgment is (that is, many of them concern *dilemmas*); so we would expect them to elicit relatively weakly held attitudes, have low personal involvement and display intermediate moral attributes. Although research is needed to confirm it, this gives us some reason to expect framing effects to be mitigated or even eliminated in the case of everyday, less ambiguous moral issues, about which participants are likely to have strong opinions (such empirical confirmation would vindicate Shafer-Landau’s

insistence, 2008, in a response to Sinnott-Armstrong, that his belief about the moral wrongness of the deliberate torture, humiliation and rape of a child would be invulnerable to changes in wording or context). We therefore have some reason to suspect that the modest framing effects examined here, which tended to lead to approximately 80% reliability, are particularly powerful, and that they do not support the claim that moral intuitions in general are unreliable.¹⁷

6. Objections and Responses

6.1. Data Points Under 70%

It is true that four of the data points fall under 70%. However, I do not believe they threaten my argument, for the following reasons.

Firstly, *even if* these data points represent unreliable moral intuitions, they form only a small minority of all of the framing effects recorded here (four out of thirty of the significant effects), and as such do not constitute strong evidence that a large proportion of our moral intuitions are unreliable.

Secondly, it is questionable that these data points in particular *do*, in fact, represent unreliable moral intuitions.

One of the data points under the 70% mark (at 66%), as well as three more of the lower points (at 73% and two at 75%), are arguably not recording differences in judgments due to morally irrelevant frames at all, but rather indicate responses to morally relevant factors which vary between the conditions. This is because the “frames” in these studies involve altering the *agent* in the moral dilemma from a fictional third-person character (“Jim” or “John”) to “you” (Tobia, Buckwalter, & Stich, 2013, and Nadelhoffer & Feltz, 2008). Sinnott-Armstrong (2011) believes that, as well as differences in order and equivalent

wording, differences in the identity of the person are morally irrelevant and fall under his idea of frames which make our judgments unreliable. I am, in fact, highly sceptical that “person frames” *are* morally irrelevant. Changing the agent in a scenario from an unknown person to the respondent is likely to generate a host of morally relevant differences, since in the version in which the respondent is the agent they will reasonably infer a lot of information regarding the attitudes, emotional states, characteristics and capabilities of the moral agent, information which they lack in the case of “Jim.” This information might affect such morally relevant things as the motives and intentions of the agent, the likely psychological consequences of the act on the agent, the ability of the agent to successfully carry out the act in question, and so on. If this is right, then these moral judgments have not been shown to be unreliable in the relevant sense.

Furthermore, the other three data points under the 70% mark (at 60%, 64%, and 68%), as well as another of the lower data points (at 72%), all come from the same study (Wiegmann, Okan & Nagel, 2012), and given the methodology of the study in question, these data points may a) misrepresent the level of reliability, and may b) not represent the reliability of *moral* intuitions. Regarding (a), the study had a very small sample size (the smallest of all the studies recorded here) with just twenty-five participants in each order condition, making this study more prone to sampling error than the others. If we wish to make inferences about the level of reliability of moral intuitions subject to framing, it is sensible to put more weight on the figures produced by studies with larger samples. Regarding (b), there is a possibility that their methodology meant that the results reflect the unreliability of non-moral considerations rather than the unreliability of moral intuitions. In each case, they simply asked participants, “Should Karl perform the proposed action?” without reference to judgments about specifically moral features. General preference or decision-making questions seem to generate rather large effect sizes (e.g. Tversky &

Kahneman, 1981) compared to questions regarding specifically moral beliefs and intuitions. The authors did instruct participants *at the beginning* of the study to answer questions with regard to *what should be done in terms of morality*, which is why I did not exclude it, but I wonder whether the lack of explicit mention of this aim at the time of answering questions about each scenario led participants to be influenced by more general decision-making factors, explaining the particularly strong result without showing that moral intuitions are highly unreliable.

6.2. *Independence of Multiple Framing Effects*

I have said that even for the judgments which *have* been shown to be affected by framing that there is only approximately a 20% chance that it will be determined by a particular frame. A potential worry is that, in the uncontrolled world outside of the laboratory, such judgments may be subject to several potential frames at once. And if the probability of the judgment being determined by one of those frames is independent of the probability of it being determined by another frame, the probability that it will be determined by *some* frame will become increasingly high the more frames one is exposed to. Note that even if the empirical assumptions underlying the objection were accurate, it would not yet show that moral intuitions *in general* are unreliable, even if it showed some particular moral intuitions to be highly unreliable. And, unfortunately, assessing the plausibility of this objection depends on empirical work not yet conducted. To my knowledge, there has been no work to date directly investigating the interaction of multiple frames of the sort with which we are concerned here, and it is very difficult to assess the likelihood that multiple frames would behave in this way without more comprehensive empirical knowledge of the range of possible framing effects there are and the nature of the psychological mechanisms that drive different effects.

6.3. *Proportions of People vs. Proportions of Judgments*

I have argued that even when our judgments are subject to framing effects there is only an approximately 20% chance that they will be determined by frames. However, there are two ways that this could be true. Firstly, it could be the case that on each occasion where I make such a judgment there is a 20% chance that the judgment will be determined by framing (the result being that about 20% of all of my frame-exposed and frame-vulnerable judgments will end up being determined by framing). There is an alternative possibility: whilst many people are immune from framing effects, twenty percent of people are such that their judgments are *always* determined by framing effects whenever they are making judgments at the same time as being exposed to the frames which can affect those judgments.

The latter possibility might feel much more worrying. I might be happy to accept the possibility that my moral judgments are off-track 20% of the time; but could I accept a 20% chance that I am off-track *all* of the time?¹⁸

It is not clear whether the latter possibility would affect my argument that the *epistemic* justification of moral intuitions is not undermined by framing effects. Nevertheless, it is unfortunate that empirical work to date on moral intuitions and framing effects has not distinguished between these two possibilities. Drawing again on empirical work on framing effects outside of the domain of moral belief (in particular on the large literature on influencing preferences by “risky choice” framing), I suspect that the answer is probably somewhere in between: individuals with certain characteristics may tend to be more susceptible to certain types of effects (but not others), and other characteristics tend to protect individuals from certain effects (but not others); but whether and in what way various personal characteristics influence one’s susceptibility to framing depends on the context, and no group is totally vulnerable or totally invulnerable to all types of framing effects (e.g.

Shiloh, Salton, & Sharabi, 2002; Levin, Gaeth, Schreiber, & Lauriola, 2002; Simon, Fagley, & Halleran, 2004; Mahoney, Buboltz, Levin, Doverspike, & Syvanteck, 2011). Even when one has a characteristic that “protects” one from a certain framing effect, this merely somewhat reduces the probability that one’s judgment will be determined by it, but does not eliminate it (Björklund, & Bäckström, 2008). Sinnott Armstrong has pointed out the desirability of a large research program investigating which features of scenarios or people (e.g. certain emotional states) lead to moral framing effects and hence unreliability (2011). On this point I am in firm agreement with him.

6.4. *Noninferential Justification*

Recall that Sinnott-Armstrong denies that moral intuitions are justified *noninferentially*. If we infer that moral intuitions can be justified because they are not sufficiently unreliable, have we only given those intuitions *inferential* justification (as Sinnott-Armstrong implies, 2008a, pp.70–72)? No; to see this, consider Sinnott-Armstrong’s own definitions of inferential and noninferential justification (2008a, p.48):

“A belief is justified *inferentially* if and only if it is justified only because the believer is able to infer it from some other belief.

A belief is justified *noninferentially* if and only if it is justified independently of whether the believer is able to infer it from any other belief.”

Now, say I have the following beliefs:

B1: “It is wrong to turn the trolley.”

B2: “Framing effects are not too strong so as to defeat justification for moral intuitions.”

B3: “My intuition that ‘it is wrong to turn the trolley’ is justified.”

Given my knowledge of framing effects, B3 is justified only because I am able to infer it from B2 (and perhaps other beliefs concerning the immediate defeasible justification of intuitions). B3 is thus *inferentially* justified. However, the same does not go for intuition B1. If it is justified, this is not because I am able to infer it from B2. It is not even clear that such an inference is possible. Our reasoning concerning the strength of framing effects merely shows that framing effects can be ruled out as defeaters of noninferential justification. Ruling out a defeater of B1's justification is not the same as justifying B1 or providing a belief from which one can infer B1.¹⁹

7. Conclusion

I must conclude, at the very least, that Sinnott-Armstrong's conclusion is mistaken in one way, and too hasty in another. It is mistaken in the claim that the evidence demonstrates a *large* probability of error; and it is too hasty in concluding that the rate of unreliability is high enough to defeat justification. Intuitions regarding exact cut-offs for reliability are hazy at best, but I do believe that the burden lies with Sinnott-Armstrong to engage with the data on these probabilities and to convince us why we must agree that we cannot have noninferential justification for our moral intuitions.

Acknowledgements

This paper has benefited greatly from insightful discussion and comments from Regina Rini, Roger Crisp, Maxime Lepoutre, and two anonymous reviewers.

Notes

1. I follow Sinnott-Armstrong in this relatively broad usage of 'framing effects' to include order effects as well as effects of wording. This is appropriate for the purpose of analysing the philosophical arguments to be discussed in this paper, but it might be noted that this

departs from standard usage in the psychological and behavioural economics literatures, where the term ‘framing effects’ is often restricted to effects of wording.

2. E.g. see Horowitz (1998). For counter-claims that framing effects fail to undermine the justificatory status of moral intuitions, see Kamm (1998) and van Roojen (1999). For an overview of contemporary empirical challenges to intuitions in moral as well as non-moral domains, and an argument that such challenges can be undercut and that empirical work can even strengthen our confidence in and use of intuitions, see Nagel and Mortensen (forthcoming). For related arguments that empirical challenges to the reliability of *epistemic* intuitions fail, see, for example, Nagel (2012), and Boyd & Nagel (2014). See also Kauppinen (2007) for the argument that psychological studies in general fail to elicit robust intuitions, which are the sort intuitions of principal philosophical interest, and, a fortiori, cannot undermine robust intuitions.

3. I thank an anonymous reviewer for pushing me to distinguish the claim that particular moral intuitions are unreliable from the related claim that the set of our moral intuitions as a whole are unreliable. These are related, for experiments which test particular sets of intuitions can provide the evidential basis from which to infer that our moral intuitions in general are unreliable. But note that the two claims are not the same. There could be evidence which supports one of these claims but not the other: in the same way that there is evidence that perceptual judgments regarding very distant objects are highly unreliable, without thereby showing that perceptual judgments *in general* are unreliable, there might be evidence which indicates that particular moral intuitions are highly unreliable, but which does not support the claim our moral intuitions in general are unreliable. This distinction will become important in sections 5 and 6.

4. According to internalist theories of justification, we must also be aware of the unreliability in order for it to undermine justification (or it must be the case that we ought to be aware of it), hence Sinnott-Armstrong's premise 4 and the clause "if we ought to know this" in premise 1. I will not dispute such claims here, and their acceptance does not affect my argument. I therefore ignore this subtlety in what follows for the sake of brevity.

5. Vague appeal to "large" or "small" likelihoods of error is symptomatic of discussions of epistemic reliability in general. For example, see Goldman (2011) who contrasts unreliable processes which "produce false beliefs a large proportion of the time" with reliable processes where "most of the beliefs" produced are true.

6. In Harman's words (1986, p.13), some beliefs are "implicit in one's believing something else. For example... it may be that in believing *P* one is committed to and so implicitly believes the proposition that one is justified by believing *P*."

7. The idea of a "frame", therefore, depends on normative assumptions regarding moral relevance. I take it to be uncontroversial that differences in wording of equivalent descriptions is morally irrelevant, and similarly for differences in scenario order, but see Sinnott-Armstrong, 2008a, pp.53-54, on the moral irrelevance of frames. See also McKenzie, 2004, for doubts that frames are normatively irrelevant. In section 6, I will dispute Sinnott-Armstrong's belief that a certain type of "frame" which alters the identity of the *agent* truly represents a morally irrelevant factor. An anonymous reviewer has pointed out that depending on one's normative assumptions about what properties are morally relevant, the class of frames which affect our moral judgments might extend much further than changes in wording and order – an interesting and important point which I unfortunately cannot explore here.

8. I thank Regina Rini for pushing me to clarify that the source of the inconsistency when it comes to framing effects relies on commitments to implicit generalizations in the way outlined in this section; for notice that the judgments “Turning the trolley is wrong in M_K ” and “Turning the trolley is not wrong in M_S ” are not by themselves logically inconsistent.

9. Any anti-realists who do not accept that moral judgments are true or false in even some minimal way can still accept that moral judgments which are subject to such effects are unreliable in a justification-undermining way by deeming such processes “unacceptable” even if not strictly speaking *error-prone* (Sinnott-Armstrong, 2011).

10. Schwitzgebel & Cushman (2012) and Tobia, Buckwalter & Stich (2012) tested different sub-samples of participants based on the extent of their academic philosophical training. If significant results are reported for participants considered as a whole group, I include merely that overall result. Where such an overall result is not significant or not reported, I provide the data points for each sub-sample individually.

11. This approached significance at $p=0.06$.

12. Liao, Wiegmann, Alexander, & Vong (2012) refer to the Switch and Footbridge dilemmas as Standard and Push, respectively.

13. Lanteri, Chelini, & Rizzello (2008) refer to Switch and a version of the Footbridge dilemma as “lever” and “stranger” respectively.

14. Mean = 79.8, median = 80.

15. An anonymous reviewer points out that thresholds for epistemic reliability may not be equal for all intuitions: just as many evidentialists require that more important beliefs are supported by more evidence in order to be justified, more important beliefs may be required to meet higher standards of reliability in order to be justified. Moral beliefs (being extremely

important) may then require higher levels of reliability in order to be justified than, say, the intuitions underlying everyday perceptual claims. This is certainly an interesting proposal and would provide a *possible* basis on which Sinnott-Armstrong might argue that our moral beliefs are unjustified at a level of 80% reliability (but see note 16), although he has previously indicated that he is unwilling to commit himself to a strong claim about thresholds for reliability (and, indeed, precisely because he interpreted the data on framing effects to indicate very *high* levels of unreliability he did not previously seem to need to make such a commitment). In any case, he either needs to avail himself of this strategy by committing himself to such a claim and arguing for it, *or* he has failed to show that framing effects on moral intuitions make them insufficiently reliable. The argumentative burden remains on his side of the debate.

16. It would be particularly interesting if it turned out that only moral intuitions about actions normally regarded as permissible were affected by framing, for if framing effects only serve to make us more conservative regarding otherwise permissible actions, and do not cause us to run the risk of mistakenly licensing acts which are in fact impermissible, we *might* be more willing to accept slightly greater levels of unreliability. This might go some way towards abating the worries raised in note 15, and similarly Sinnott-Armstrong's concern that those who accept moral beliefs even in the knowledge of the existence of framing effects "must be willing to take big chances with their moral beliefs" (2008b, p.101).

17. If framing effects do not, in the end, pose a challenge to the justificatory status of our moral intuitions, then why has it seemed to many, such as Sinnott-Armstrong, that they do? As an anonymous reviewer has pointed out, one might desire an error theory for why this mistake has been made. I have suggested that this may be due to a lack of attention to relevant numerical details of the empirical studies in question. This may be combined with the mistaken assumption that, if frame *x* can cause intuition *p* and frame *y* can cause intuition

$\neg p$, then our intuitions are only 50% reliable, because only one out of two of those intuitions can be correct. The mistake here is to take it as certain that frame x will cause intuition p and that frame y will cause intuition $\neg p$. The reviewer has made a fascinating alternative suggestion that the mistake might be due to paying attention to the *decline* in confidence in the reliability of our moral intuitions following the discovery of framing, rather than the observation that the reliability of moral intuitions has dipped below a certain threshold.

18. I thank Regina Rini for bringing this issue to my attention.

19. Tolhurst (2008, pp.81–82) makes an analogous point.

References

Audi, R. (2004). *The good in the right: A theory of intuition and intrinsic value*. Princeton, NJ: Princeton University Press.

Björklund, F., & Bäckström, M. (2008). Individual differences in processing styles: validity of the Rational-Experiential Inventory. *Scandinavian Journal of Psychology*, 49, 439–446.

Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, Action, and Intention as Factors in Moral Judgments: An fMRI Investigation. *Journal of Cognitive Neuroscience*, 18, 803–817.

Boyd, K., & Nagel, J. (2014). The Reliability of Epistemic Intuitions. In E. Machery & E. O'Neill (Eds.), *Current Controversies in Experimental Philosophy* (pp.109–127). New York, NY: Routledge.

Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience and Behavioral Reviews*, 36, 1249–1264.

- Cohen, S. (1984). Justification and Truth. *Philosophical Studies*, 46, 279–295.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- Goldman, A. (2011). Reliabilism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), URL = <http://plato.stanford.edu/archives/spr2011/entries/reliabilism/>.
- Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology*, 26, 201–218.
- Harman, G. (1986). *Change in View*. Cambridge, MA: MIT Press.
- Horowitz, T. (1998). Philosophical Intuitions and Psychological Theory. *Ethics*, 108, 367–385.
- Kamm, F. (1998). Moral Intuitions, Cognitive Psychology, and the Harming-Versus-Not-Aiding Distinction. *Ethics*, 108, 463–488.
- Kauppinen, A. (2007). The Rise and Fall of Experimental Philosophy. *Philosophical Explorations*, 10, 95–118.
- Klein, P. (2013). Skepticism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition). URL = <http://plato.stanford.edu/archives/sum2013/entries/skepticism/>
- Lanteri, A., Chelini, C., & Rizzello, S. (2008). An Experimental Investigation of Emotions and Reasoning in the Trolley Problem. *Journal of Business Ethics*, 83, 789–804.

- Levin, I. P., Gaeth, G. J., Schreiber, J., & Lauriola, M. (2002). New Look at Framing Effects: Distribution of Effect Sizes, Individual Differences, and Independence of Types of Effects. *Organizational Behavior and Human Decision Processes*, 88, 411–429.
- Levin, I. P., Schneider, S. L., Gaeth, G. J. (1998). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes*, 76, 149–188.
- Liao, S. M., Wiegmann, A., Alexander, J., & Vong, G. (2012). Putting the trolley in order: Experimental philosophy and the loop case. *Philosophical Psychology*, 25, 661–671.
- Lombrozo, T. (2009). The Role of Moral Commitments in Moral Judgment. *Cognitive Science*, 33, 273–286.
- Mahoney, K. T., Buboltz, W., Levin, I. P., Doverspike, D., Svyantek, D. J. (2011). Individual differences in a within-subjects risky-choice framing study. *Personality and Individual Differences*, 51, 248–257.
- McKenzie, C. R. M. (2004). Framing effects in inference tasks—and why they are normatively defensible. *Memory & Cognition*, 32, 874–885.
- Nadelhoffer, T., & Feltz, A. (2008). *Neuroethics*, 1, 133–144.
- Nagel, J. (2012). Intuitions and Experiments: A Defense of the Case Method in Epistemology. *Philosophy and Phenomenological Research*, 85, 495–527.
- Nagel, J., & Mortensen, K. (forthcoming). Armchair-Friendly Experimental Philosophy. Forthcoming in J. Sytsma & W. Buckwalter (Eds.), *A Companion to Experimental Philosophy*. Blackwell.

Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology*, *17*, 145–171.

Schwitzgebel, E., & Cushman, F. (2012). Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind & Language*, *27*, 135–153.

Shafer-Landau, R. (2003). *Moral realism: A defense*. Oxford: Clarendon Press.

Shafer-Landau, R. (2008). Defending Ethical Intuitionism. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality: Intuition and Diversity* (pp. 47–76). Cambridge, MA: MIT Press.

Shiloh, S., Salton, E., & Sharabi, D. (2002). Individual differences in rational and intuitive thinking styles as predictors of heuristic responses and framing effects. *Personality and Individual Differences*, *32*, 415–429.

Simon, A. F., Fagley, N. S., & Halleran, J. G. (2004). Decision Framing: Moderating Effects of Individual Differences and Cognitive Processing. *Journal of Behavioral Decision Making*, *17*, 77–93.

Sinnott-Armstrong, W. (2008a). Framing Moral Intuitions. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality: Intuition and Diversity* (pp. 47–76). Cambridge, MA: MIT Press.

Sinnott-Armstrong, W. (2008b). How to Apply Generalities: Reply to Tolhurst and Shafer-Landau. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality: Intuition and Diversity* (pp. 97–105). Cambridge, MA: MIT Press.

- Sinnott-Armstrong, W. (2011). Emotion and Reliability in Moral Psychology. *Emotion Review*, 3, 288–289.
- Tobia, K., Buckwalter, W., & Stich, S. (2013). Moral intuitions: Are philosophers experts? *Philosophical Psychology*, 26, 629–638.
- Tolhurst, W. (1990). On the epistemic value of moral experience. *Southern Journal of Philosophy*, 29(Supp.), 67–87.
- Tolhurst, W. (1998). Seemings. *American Philosophical Quarterly*, 35, 293–302.
- Tolhurst, W. (2008). Moral Intuitions Framed. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality: Intuition and Diversity* (pp. 77–82). Cambridge, MA: MIT Press.
- Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211, 453–458.
- Van Roojen, M. (1999). Reflective Moral Equilibrium and Psychological Theory. *Ethics*, 109, 846–857.
- Wiegmann, A., & Waldmann, M. R. (2014). Transfer effects between moral dilemmas: A causal model theory. *Cognition*, 131, 28–43.
- Wiegmann, A., Okan, Y., & Nagel, J. (2012). Order effects in moral judgment. *Philosophical Psychology*, 25, 813–836.