The Intelligibility of Moral Intransigence: A Dilemma for

Cognitivism about Moral Judgment

Final Draft. Forthcoming in Analysis

Many have argued that various features of moral disagreements create problems for cognitivism about moral judgment, but these arguments have been shown to fail. In this paper I articulate a new problem for cognitivism that derives from features of our responses to moral disagreement. I argue that cognitivism entails that one of the following two claims is false: (1) a mental state is a belief only if it tracks changes in perceived evidence; (2) it is intelligible to make moral judgments that do not track changes in perceived evidence. I explain that there is a good case that (1) holds such that we should prefer theories that do not entail the negation of (1). And I argue that the seeming intelligibility of entirely intransigent responses to peer disagreement about moral issues shows us that there is a good case that (2) holds.

Ι

Many including Simon Blackburn (1984: 168), R.M. Hare (1952: 146-149), Terence Horgan and Mark Timmons (1993), Mark Kalderon (2005: ch. 1), Charles Stevenson (1963: 48–51), and Folke Tersman (2006: ch. 5-6) have argued that various features of moral disagreement create problems for

cognitivism about moral judgment. But many have argued that all these arguments fail. In this paper I make a new argument for the view that claims about moral disagreement seem to create a problem for cognitivism about moral judgment.

In order to make this argument I need to explain several concepts. Firstly, intelligibility. A mental or physical act is intelligible if and only if we can make sense of it. Many use this notion of intelligibility. For instance, according to Donald Davidson (1963: 693-695) and Michael Smith (1994: 95), claims about A's reasons for ϕ -ing render A's ϕ -ing intelligible. Un-intelligible acts are acts that it seems could not happen at all. Consider someone who freely spends all day counting blades of grass but who claims to get no pleasure out of doing this, to not desire doing this with their day, to see nothing of value in counting grass all day, and to not feel any obligation to do this. It is not intelligible that this person could be doing what they claim to be doing. For it is hard to believe that the grasscounter could really freely spend all day counting blades of grass but get no pleasure out of doing this, not desire to do this with their day, see no other value in doing this, and feel no obligation to do this. Note that for ϕ -ing to be intelligible is not for ϕ -ing to be *justifiable*. Suppose that Amy is explaining to Becky why she did not complete a task. It would be intelligible for Amy to worry that Becky thinks that she is presenting a poor excuse even if Amy has no reason, such as a grimace or a probing question, for thinking that Becky thinks that she is presenting a poor excuse.² (Sometimes it might seem intelligible that someone could ϕ even though it is not intelligible that they could knowingly ϕ . For this paper, take 'it is unintelligible that A could ϕ ' to mean 'it is unintelligible that A could knowing φ').

1

¹ See Kramer (2009: 187-190), Fisher (2007), Enoch (2009: 31-32), Plunkett and Sundell (2013), Dowell (2015), and Dunaway and McPherson (2016).

² Audi (2011: 29)

Secondly, non-derivative moral judgments. A non-derivative moral judgment is a judgment about the moral status of ϕ -ing that is not *solely* inferred from any other judgment. An example of a non-derivative moral judgment is that if ϕ -ing has better consequences than not- ϕ -ing, then we should ϕ rather than not- ϕ . A committed act-consequentialist, for instance, only makes one non-derivative moral judgment about the moral wrongness of actions, namely that performing actions that do not maximize the good is morally wrong. Thirdly, epistemic peerhood. Assume that for A to be B's epistemic peer regarding p is for B to have the same evidence, sensitivity to the evidence, intelligence, freedom from bias, and cognitive functioning regarding p as A does and/or for B to be as reliable as A regarding whether p.³ For instance, those with whom we took high school math classes and who were about as good as we were at mental math in these classes are plausibly our epistemic peers regarding mental math questions. And we might see those with whom we went to grad school as our epistemic peers about any randomly chosen question about contemporary philosophy or the history of philosophy.

Consider

_

³ See Christensen (2009: 756-757). This definition is ecumenical between different accounts of epistemic peerhood; see Killoren (2010: 15-20), Frances (2014: 45), and Matheson (2015: §2.1).

Moral Peer Intransigence is Intelligible. For any agent A, it is intelligible that both (a) A's non-derivative moral judgment about the moral status of ϕ -ing and (b) A's judgment about who her epistemic peers about the moral status of ϕ -ing are could be simultaneously entirely intransigent in light of her judging that a significant number of her epistemic peers about the moral status of ϕ -ing disagree with her about the moral status of ϕ -ing.

In §2 I argue that we have good reasons to accept *Moral Peer Intransigence is Intelligible*. In §3 I argue that we can derive a strong case from the recent epistemological literature that a mental state is a belief only if it tracks changes in perceived evidence. In §4 I argue that my arguments in §2-3 produce a dilemma for cognitivism about moral judgment because the combination of the view that a mental state is a belief only if it tracks changes in perceived evidence and *Moral Peer Intransigence is Intelligible* entails that non-derivative moral judgments are not beliefs.

II

There are three good reasons to hold *Moral Peer Intransigence is Intelligible*. Firstly, it seems that our intuitions about cases favour *Moral Peer Intransigence is Intelligible*. For instance, suppose that Anna is a moral philosopher who also spends a lot of time doing humanitarian work for torture victims and engaging in anti-torture advocacy. She judges that it is always wrong to torture. But Anna comes to believe that at least half of those who have spent their lives open-mindedly considering, thinking about, and writing arguments about the moral status of torture, and whose moral capacities, sensitivity, and reasoning skills she respects, judge that there are some circumstances in which torture is morally permissible. So, Anna finds herself making a judgment about the moral status of torture that she

knows half of those whom she believes to be her epistemic peers about the moral status of torture disagree with. It seems intelligible for Anna to carry on judging that half of those whose moral reasoning capacities she respects and who have spent their lives thinking about the moral status of torture are her epistemic peers about the moral status of torture but to also not alter her moral judgment about whether torture is ever morally permissible at all. With her strong anti-torture commitment, it would not be that surprising if she didn't change her judgment at all even if she acknowledged that so many of those whose epistemic capacities regarding this particular issue are at least as good as her own disagree with her. But if this is correct, then *Moral Peer Intransigence is Intelligible*.

Secondly, for A's ϕ -ing to be unintelligible is for A's ϕ -ing to have a property instantiated by B's (i) freely counting blades of grass all day when they—B—(ii) get no pleasure out of counting blades of grass all day, do not desire to do this, see no value in doing this, and do not feel any obligation to count blades of grass all day. This is because it is unintelligible that someone could be (i) when they (ii). But there does not seem to be any property shared by, for instance, Anna's continuing to hold exactly the same moral judgment regarding the wrongness of torture and failing to downgrade the epistemic status of those whom she judges to be her epistemic peers about this issue and B's counting blades of grass all day. So, Moral Peer Intransigence is Intelligible.

It might be objected to this second argument that Anna's intransigent moral judgment does not seem to share any properties in common with *B*'s counting blades of grass all day just because these acts and judgments and their contents are so different. However, consider

-

⁴ Beyond trivial and irrelevant properties that these putative acts have in common such as the property they both have of being putative acts.

Unicorns. Believing that unicorns exist and that if unicorns exist, then there are some mythical creatures, but not believing that there are some mythical creatures; and

New York. Desiring to go to New York but not desiring any possible means to go

to New York.

It is widely held to be intuitive that *Unicorns* and *New York* instantiate irrationality or breach a rational requirement in the same sense.⁵ And if we can observe similarities and differences in the rationality of different types of judgments and acts with very different contents, then I see no reason to believe that we cannot also observe similarities and differences in the intelligibility of very different types of judgments and acts with very different contents.

Thirdly, it seems that if many people who are competent users of a concept C believe that it is conceptually possible that X could fall under concept C, then, at least when other things are equal, prima facie, we should hold that it is conceptually possible that X could fall under C rather than unintelligible that X could fall under C (for otherwise we would attribute conceptual confusion to competent users of C for no good reason). But Kieran Setiya (2012: 19-20) has claimed that Moral Peer Intransigence is Intelligible. Mark Kalderon and Matthew Kramer also seem to judge that Moral Peer Intransigence is Intelligible. Kalderon claims that the following example is not only intelligible but one in which Edgar responds to a disagreement epistemically reasonably: Kalderon imagines that Edgar and Bernice are in a disagreement in non-derivative moral judgments about the moral status of abortion. Kalderon (2005: 35) says that

⁵ See Way (2010) and the references therein.

⁶ See Rowland (2016, pp. 1374-1376).

Bernice might strike Edgar as an otherwise rational and reasonable, informed human being who coherently accepts a reason that, if genuine, would undermine his acceptance of the permissibility of abortion. Nevertheless, Edgar feels no embarrassment about this. His persistence in his liberal morality is unflinching.

And in response to Kalderon, Kramer (2009: 186) bemoans the 'tiresome familiarity' of people who respond to disagreements like Edgar. If people who respond like Edgar are tiresomely familiar, then this intransigent response to peer disagreement is intelligible. So, we should hold *Moral Peer Intransigence is Intelligible* in order to refrain from attributing excess conceptual confusion to moral philosophers such as Kalderon and Kramer. (The claim that *Moral Peer Intransigence is Intelligible* is not the same as Kalderon's claim that intransigent responses to peer disagreement about moral issues are *epistemically reasonable* or *justified*; intelligible responses can be responses that are not justified).

_

⁷ It might be objected that Kalderon does not stipulate that Edgar judges that Bernice is his peer about the moral status of abortion. However, presumably Kalderon means to stipulate that Edgar judges that Bernice is his peer about the moral status of abortion and not just about morality more generally as Kalderon (2005: 21, 27, 34) means to imagine a case in which if the matter that Edgar and Bernice disagreed about were a cognitive matter, then they would be under a lax obligation to investigate whether they are really right about the matter further. But when D1 is a cognitive matter, the fact that A whom we judge to be our peer about matters in domain D, which is the domain of D1, disagrees with us about D1 does not establish that we are under a lax obligation to investigate D1 further. Suppose that we judge that A is our peer about metaphysics generally but not about metaphysical grounding which is our speciality and not theirs. The fact that A disagrees with us about metaphysical grounding does not necessarily give rise to any obligation for us to investigate the matter further. For if it is our specialty and not theirs, then we have thought about and puzzled over it more than they have, and our judgment is—given that they are not our peers on this particular issue—more informed than theirs. So, it seems that reading Kalderon charitably requires us to add to Kalderon's stipulations in the Bernice and Edgar case that Bernice strikes Edgar as his epistemic peer regarding the moral status of abortion in particular.

III

Consider

Beliefs Track Perceived Evidence. If one gains what one believes to be new evidence that bears on whether p, and one's judgment regarding p (whether p) is a belief, then either (i) one adjusts one's judgment regarding p in light of and in line with this evidence or (ii) one adjusts one's judgment that what one gained was in fact evidence bearing on whether p.

I can't give a decisive case for *Beliefs Track Perceived Evidence* in this paper. But in this section, I'll explain why *Beliefs Track Perceived Evidence* seems plausible to me and many others such that I'd prefer not to accept views that entail its falsity.

Firstly, Tamar Gendler (2008: 565-566) has argued that *Beliefs Track Perceived Evidence* fits well with and explains our intuitions about the following cases:

I used to believe that stomach ulcers were caused primarily by stress and diet; but when Warren and Marshall's research on the *Helciobacter pylori* bacterium became widely known, I revised my belief to reflect this information. Williamson's "N.N."—"who has not yet heard the news from the theatre where Lincoln has just been assassinated"—believes that Lincoln is President; but as soon as he learns that Lincoln has been shot, he will make the corresponding adjustment in his belief.

Secondly, many have argued that if *Beliefs Track Perceived Evidence* did not hold, then it would be intelligible for us to occurrently believe that *p* whilst simultaneously occurrently believing that our evidence does not justify our believing that *p*. But

No Akrasia. It is not intelligible that we might consciously believe that p whilst simultaneously consciously believing that our evidence does not justify us in believing that p.

One way of seeing the plausibility of *No Akrasia* is to see that the claim, 'p but my evidence does not support p' seems to be an instance of a Moorean paradoxical claim.⁸ Another way of seeing the plausibility of *No Akrasia* is to consider cases like the following:

Matt is extremely afraid of flying. When professional obligations require him to travel (even thousands of miles), he either drives or takes a train. He does not travel overseas. When his friends and loved ones travel by air, he obsessively checks the status of their flights online, and calls them as soon as possible after landing to make sure that they're OK. When asked about all this behavior, he doesn't defend it. Instead, he says things like the following: "Of course the evidence shows that flying is not particularly dangerous—certainly less dangerous than driving comparable distances, but I just can't shake the belief that if I fly, my plane will crash and I will die. What's holding it up there anyway?"

According to many, such as Tamar Gendler, John Greco, Nishi Shah, and David Velleman, it is hard to believe that what Matt says could be literally true. Of course, *worrying* about flying and yet still believing that all the evidence shows that flying is not particularly dangerous *is* perfectly intelligible to us. But, according to these philosophers, consciously believing that flying is dangerous and

9

⁸ See Chisholm (1989: 6), Scanlon (1998: 25), Adler (2002: 21), Horowitz (2013: 725), and Greco (2014).

⁹ Greco (2014: 202).

simultaneously consciously believing that all the evidence shows that flying is not particularly dangerous is very difficult to understand; it seems impossible to knowingly simultaneously have both these occurrent beliefs.¹⁰

Thirdly, *Beliefs Track Perceived Evidence* is entailed by 'transparency'. According to 'transparency', the view that beliefs are necessarily transparent to the truth, 'the deliberative question *whether to believe that p* inevitably gives way to the factual question *whether p*'. Transparency has been held and defended by many including Bernard Williams (1973: 148), Richard Moran (1988: 146), Richard Foley (1993: 16), and Nishi Shah (2006: 481). And transparency seems to hold. For instance, when we think about what to believe regarding the winner of the next Presidential election we only think about considerations that bear on who the winner will in fact be, that is, evidence regarding who the winner is likely to be. But if transparency holds, then we cannot take ourselves to have gained new evidence that *p* without adjusting our belief that *p* in line with this evidence; this is a fact about what it is for a state to be a belief. So, if transparency holds, then *Beliefs Track Perceived Evidence*.

Beliefs Track Perceived Evidence is a descriptive thesis about how belief-states can and cannot behave. It might be objected that instead of this descriptive thesis we should accept the normative thesis that we ought to adjust our beliefs in line with our perceived evidence. However, such a normative thesis cannot explain why it seems that Matt (in the example above) cannot genuinely believe that all the evidence shows that flying isn't dangerous but also genuinely simultaneously and occurrently believe that if he flies his plane will crash and he will die; for such a normative thesis only entails conclusions about what we ought to believe, not what we can and cannot believe. Furthermore, if we accept that

¹⁰ See Gendler (2008) and *supra* note 8.

¹¹ Shah (2006: 481)

¹² See Conee and Feldman (2004) and Sylvan (2016: 370-371).

beliefs are transparent to the truth, then we must accept that *Beliefs Track Perceived Evidence* regardless of whether we should accept that we *ought* to adjust our beliefs in line with our perceived evidence.¹³

IV

Consider

Perceived Peer Disagreement is Perceived Evidence. When one comes to believe that someone one believes to be one's epistemic peer regarding p disagrees with one regarding p, (other things equal) one comes to believe that there is (strong) evidence against one's view regarding p that one did not previously have regarding p.

Perceived Peer Disagreement is Perceived Evidence is extremely plausible. Suppose that Alice and Beth go to dinner with several of their friends. At the end of the meal Alice decides to use her mental capacities to calculate everyone's equal share of the bill. Suppose that Alice believes, with good reason, that Beth is her epistemic peer regarding mental math questions. Alice calculates everyone's share of the bill and comes to believe that everyone owes \$30. But then Beth turns to Alice and tells her that she has just mentally calculated everyone's share of the bill and according to her calculation everyone owes \$35. Alice has now acquired evidence that she did not previously possess regarding whether everyone owes \$30 or not. (This evidence is higher-order evidence, for it is evidence that she erred in her calculation.

_

¹³ Stanley (2015: ch. 5) argues that we should reject *Beliefs Track Perceived Evidence* because moral, political, philosophical, and religious beliefs that are resistant to evidence breach *Beliefs Track Perceived Evidence* but still count as beliefs. Some who reject *Beliefs Track Perceived Evidence* give positive alternative accounts of how beliefs are necessarily (descriptively) related to truth. For instance, Vahid (2009: ch. 1-2, esp. 25-26) argues that believing that p only necessarily involves regarding that p is true (for its own sake). For discussion, see Fassio (2015: §2.c).

But such evidence is still evidence that bears on whether she should believe, and the degree to which she should believe, that everyone owes \$30). 14 So, it seems that Perceived Peer Disagreement is Perceived Evidence. And it is generally presumed by all parties in the peer disagreement literature that Perceived Peer Disagreement is Perceived Evidence. 15

However, in §3 I discussed the view that

Beliefs Track Perceived Evidence. If one gains what one believes to be new evidence that bears on whether p, and one's judgment regarding p (whether p) is a belief, then either (i) one adjusts one's judgment regarding p in light of and in line with this evidence or (ii) one adjusts one's judgment that what one gained was in fact evidence bearing on whether p.

If Beliefs Track Perceived Evidence and one's judgment regarding p is a belief, then it will not be intelligible to judge that one has gained evidence that bears on whether p but to fail to either (i) adjust one's judgment regarding p in light of this evidence or (ii) adjust one's judgment that what one gained was evidence bearing on whether p (once one has realized that there is this interaction between the evidence that one has acquired and one's belief). But if Perceived Peer Disagreement is Perceived Evidence, then to be entirely intransigent in one's judgments in light of perceived peer disagreement is just to fail to do either (i) or (ii). So, the combination of Perceived Peer Disagreement is Perceived Evidence and Beliefs Track Perceived Evidence entails that

¹⁴ See Kelly (2010) and Christensen (2010).

¹⁵ *Supra* note 14.

Peer Intransigent Judgments are not Beliefs. For any agent A and judgment J, if it is intelligible for both A's judgment J and A's judgments about who her epistemic peers about the content of J are to be simultaneously entirely intransigent in light of her judging that a significant number of her epistemic peers about the content of J disagree with her regarding the content of J, then J is not a belief.

In §2 I argued that we have strong reasons to accept

Moral Peer Intransigence is Intelligible. For any agent A, it is intelligible that both (a) A's non-derivative moral judgment about the moral status of ϕ -ing and (b) A's judgment about who her epistemic peers about the moral status of ϕ -ing are could be simultaneously entirely intransigent in light of her judging that a significant number of her epistemic peers about the moral status of ϕ -ing disagree with her about the moral status of ϕ -ing.

But the combination of Moral Peer Intransigence is Intelligible and Peer Intransigent Judgments are not Beliefs entails that non-derivative moral judgments are not beliefs. So, if cognitivism holds—and so non-derivative moral judgments are beliefs—then either Moral Peer Intransigence is Intelligible or Peer Intransigent Judgments are not Beliefs must be false.

I'm strongly attracted to cognitivism. But I find the fact that it entails the falsity of either Moral Peer Intransigence is Intelligible or Peer Intransigent Judgments are not Beliefs worrying. This is because, as I discussed earlier in this section, (i) Peer Intransigent Judgments are not Beliefs follows directly from the combination of Perceived Peer Disagreement is Perceived Evidence and Beliefs Track Perceived Evidence. But (ii),

as I also discussed earlier in this section, it seems hard to deny that Perceived Peer Disagreement is Perceived Evidence. And, (iii) although Beliefs Track Perceived Evidence is not uncontroversial, as I explained in §3, we can derive a strong case that Beliefs Track Perceived Evidence from the recent epistemological literature and in this case it seems to me that it would be better not to have to accept a view that entails the falsity of Beliefs Track Perceived Evidence. And, given (i-iii) we have strong reasons not to reject Peer Intransigent Judgments are not Beliefs. But, as I explained in §2, we have strong reasons to accept Moral Peer Intransigence is Intelligible. So, it seems that those attracted to cognitivism, such as myself, face a dilemma: to maintain cognitivism we must either reject Moral Peer Intransigence is Intelligible or Peer Intransigent Judgments are not Beliefs but we have good reasons not to reject either of these claims.¹⁶

References

Adler, Jonathan (2002). 'Akratic Believing'. Philosophical Studies 110: 1-27.

Audi, Robert (2011). Rationality and Religious Commitment. OUP.

Blackburn, Simon (1984). Spreading The Word. Clarendon.

Chisholm, Roderick (1989). Theory of Knowledge Third Edition. Prentice Hall.

Christensen, David (2009). 'The Epistemology of Controversy'. *Philosophy Compass 4*: 756-767

_____ (2010). 'Higher-Order Evidence'. *Philosophy and Phenomenological* Research 81: 185-215.

Conee, Earl and Richard Feldman (2004). Evidentialism. OUP.

Davidson, Donald (1963). 'Actions, Reasons, and Causes'. *Journal of Philosophy* 60: 685-700.

Dowell, Janice (2015). 'The Metaethical Insignificance of Moral Twin Earth'. In Russ Shafer-Landau (ed)., Oxford Studies in Metaethics Volume 11. OUP.

Dunaway, Billy, and Tristram McPherson (2016). 'Reference Magnetism as a Solution to the Moral Twin Earth Problem'. *Ergo 3*: 639-679

Enoch, David (2009). 'How is Moral Disagreement a Problem for Realism?' *Journal of Ethics 13*: 15-50.

Fassio, Davide (2015). 'The Aim of Belief'. Internet Encylopedia of Philosophy. www.iep.utm.edu

Fisher, Andrew (2007). 'Review of Folke Tersman, Moral Disagreement'. Notre Dame Philosophical Reviews.

¹⁶ I would like to thank audiences at the University of Reading, La Trobe University, the Pacific APA, and the 4th Australasian Workshop in Moral Philosophy at Kioloa for their comments. I would also like to thank David Killoren, Tyler Paytas, Charles Urban, and two anonymous reviewers for their comments.

Foley, Richard (1993). Working Without a Net. OUP.

Frances, Bryan. (2014). Disagreement. Polity.

Gendler, Tamar Szabó (2008). 'Alief in Action'. Mind and Language 23: 552-585.

Greco, Daniel (2014). 'A Puzzle about Epistemic Akrasia'. *Philosophical Studies 167:* 201–219.

Hare, R.M. (1952). The Language of Morals. Clarendon.

Horgan, Terrence and Mark Timmons (1993). 'New Wave Moral Realism Meets Moral Twin Earth'. In John Heil (ed.), Rationality, Morality, and Self-Interest. Rowman and Littlefield.

Horowitz, Sophie (2014). 'Epistemic Akrasia'. Noûs 48, 4: pp. 718-744.

Kalderon, Mark (2005). Moral Fictionalism. Clarendon.

Killoren, David. (2010). Moral intuitions, reliability, and disagreement. *Journal of Ethics and Social Philosophy*, 4: 1–35.

Kelly, Thomas (2010). 'Peer Disagreement and Higher-Order Evidence'. In Richard Feldman and Ted Warfield (ed.), *Disagreement*. OUP

Kramer, Matthew (2009). Moral Realism as a Moral Doctrine. Blackwell.

Matheson, Jonathan. (2015). 'Disagreement and epistemic peers'. Oxford Handbooks Online.

Moran, Richard (1988). 'Making Up Your Mind: Self-Interpretation and Self Constitution'. Ratio 1: 135-151.

Plunkett, David, and Timothy Sundell (2013). 'Disagreement and the Semantics of Normative and Evaluative Terms'. *Philosophers' Imprint 13*: 1-37.

Rowland, Richard (2016). 'In defence of good simpliciter'. Philosophical Studies 173, 5: pp 1371–1391.

Scanlon, T.M. (1998). What We Owe To Each Other. Harvard.

Setiva, Kieran (2012). Knowing Right from Wrong. OUP.

Shah, Nishi (2006). 'A New Argument for Evidentialism'. *Philosophical Quarterly 56*: 481-498.

Smith, Michael (1994). The Moral Problem. Blackwell.

Stanley, Jason (2015). How Propaganda Works. Princeton.

Stevenson, Charles (1963). Facts and Values. Yale.

Sylvan, Kurt (2016). 'Epistemic Reasons 1: Normativity'. Philosophy Compass 11, 7: pp. 364–376.

Tersman, Folke (2006). Moral Disagreement. CUP.

Vahid, Hamid (2009). The Epistemology of Belief. Palgrave.

Way, Jonathan (2010). 'The Normativity of Rationality'. *Philosophy Compass 5*: 1057-1068.

Williams, Bernard (1973). Problems of the Self. CUP.