

A Note on Error Theory and the Refined Moral Problem

Alexander Miller and Seth Whittington

Abstract

In this paper, we argue that “The Moral Problem” identified by Michael Smith in his book of that name as “the central organizing problem” of metaethics needs to be refined in order to accommodate moral error theories (in the style of J.L. Mackie), and we suggest a refinement that allows it to do this. We conclude by drawing out some consequences for the formulation of internalism about moral motivation.

1. Introduction

One of the most important and deservedly most cited and discussed books in metaethics in the past three decades is Michael Smith’s *The Moral Problem* (Smith 1994). In this book, Smith outlines what he takes to be “the central organizing problem in contemporary metaethics” (1994: 11), what he calls “The Moral Problem”. According to Smith, the main views in recent and contemporary metaethics can be regarded as making distinctive attempts at solving this problem. We are fans of Smith’s way of organizing the landscape of recent and contemporary metaethics, but we think that as it stands Smith’s formulation of “The Moral Problem” isn’t yet capable of accommodating all of the central metaethical views that Smith himself regards as important. The purpose of this note, in a suitably tentative spirit, is to suggest for discussion a refined version of “The Moral Problem” that possibly overcomes this limitation. In section 2 we outline Smith’s version of the “The Moral Problem” and show how he uses it to provide a topography of metaethical views. In section 3 we argue that as it stands, it is unclear how the topography provided by Smith’s version of “The Moral Problem” can accommodate error theories of the sort most famously advocated by J.L. Mackie (Mackie 1977), and we trace this limitation to a failure to distinguish between two distinct types of what Smith terms “Moral Nihilism”. In section 4 we present “The Refined Moral Problem”, an augmented version of Smith’s problem that does allocate a clear place to the moral error theory. In section 5, we conclude with

some brief comments on the consequences our refinement of Smith's "Moral Problem" has for the formulation of internalist views of moral judgement and motivation.

2. Smith's "Moral Problem"

In Smith's view, the problem is that the following three propositions are individually very plausible but appear to generate an inconsistency when taken together (Smith 1994: 12):

- (1) Moral judgements of the form 'It is right that I Φ ' express a subject's beliefs about an objective matter of fact, a fact about what is right for her to do.
- (2) It is a conceptual truth that if an agent judges that it is right to Φ and is practically rational then, *ceteris paribus*, she is motivated to Φ .¹
- (3) An agent is motivated to act in a certain way just in case she has an appropriate desire and means-end belief, where belief and desire are, in Hume's terms, "distinct existences"².

According to (1), when I judge that staying at home if I'm sick is right, I express the *belief* that staying at home if I'm sick is right. If that belief accords with the moral facts – if it actually is the case that it is right to stay home if you're sick - my judgement is true, and false otherwise. (1) certainly seems to be *prima facie* plausible.

Can we say the same thing about (2)? When someone judges that it is right to stay home if sick we would normally expect that, other things being equal, they will be motivated to stay home if sick. If they don't, we would expect an explanation why not. Moreover, it seems plausible to think that this is a distinctive feature of specifically moral judgement: other sorts of judgement appear not to have this kind of close tie to motivation. So (2) seems *prima facie* plausible also.

Finally, on the face of it, (3) seems plausible. Beliefs cannot produce actions on their own. Beliefs aim to tell us how things stand in the world. So although beliefs aspire to tell us how the world might be altered to make it

¹ Although the requirement that the agent is practically rational is not present in Smith's initial presentation of the problem, he introduces it (in our view rightly) in chapter 3 of his (1994), so we've included it here. Nothing turns on this for the main argument of this note.

² To say that a belief B and a desire D are "distinct existences" is to say that there is no necessary connection between them: it is possible to have B in the absence of D and D in the absence of B.

different from how it actually is, they don't aspire to tell us how it should be changed. Similarly, desires cannot produce actions on their own. Desires aim to tell us how things should stand in the world. So although desires aspire to tell us how the world should be changed, because they don't aim to tell us how things actually stand in the world as it is, they don't tell us how the world needs to be altered to make it the way they tell us that it should be. So neither beliefs nor desires on their own are capable of motivating someone to act. But beliefs and desires together can produce action: the desire tells us how things in the world should be, and the belief tells us how we need to change the world so that it is that way. And the fact that beliefs and desires have these different "directions of fit" rules out there being a necessary connection between them. So (3) seems *prima facie* plausible.

Obviously, these are just preliminary remarks and much more remains to be said regarding the individual plausibility of (1), (2) and (3). But we can already see some recognisable metaethical views on the table. (1) amounts to *cognitivism* about moral judgement, (2) amounts to *internalism* about moral judgement and motivation, while (3) is what is known as the *Humean theory of motivation* (or, as below, *Humeanism*).

Despite the fact that (1), (2) and (3) are individually plausible, as Smith points out, it is far from clear how they can be held simultaneously. Consider the judgement that it is right to stay home when sick, made by a practically rational agent, Jones. From (1), Jones expresses the belief that it is right to stay home when sick. From (2), it follows as a matter of conceptual fact that Jones is motivated to stay home when sick. From (3), it follows that Jones has a desire to stay home when sick, so that there is a conceptual, and hence necessary, connection between the belief expressed in virtue of (1) and the desire present in virtue of (2) and (3). But (3) itself requires that there be no necessary connection between any given belief and any given desire.

This apparent inconsistency is what Smith calls "The Moral Problem", and the main metaethical positions are represented in terms of how they attempt to respond to it. *Non-Cognitivism*, *Externalism* and *Anti-Humeanism* all concede that there is a problem generated by the conjunction of (1), (2) and (3), and attempt to resolve the problem by rejecting one member of the trio.

Non-Cognitivism attempts to preserve (2) and (3) by giving up (1). In its purest forms (e.g. Ayer (1946), chapter 6) when Jones judges that it is right to stay home when sick he's not expressing a belief at all. Rather, he is expressing a non-cognitive *sentiment* of moral disapproval towards acts of not staying home when sick. So the non-cognitivist gives up (1). But moral judgements can certainly retain a necessary connection with motivation,

because someone who makes a moral judgement is expressing a non-cognitive sentiment, a psychological state more akin to a desire than a belief. And we can retain the Humean view that requires motivation to involve, *inter alia*, the possession of a suitable desire. The Non-Cognitivist thus solves “The Moral Problem” by rejecting (1) and retaining (2) and (3). Views in the broad Non-Cognitivist family have been developed by Ayer (1946), Stevenson (1937, 1944) and – more distantly – Hare (1952), Blackburn (1984, 1993) and Gibbard (1990, 2003).³

Externalism retains (1) and (3) – Cognitivism and Humeanism – by giving up (2). According to the externalist, moral judgements have at most a contingent and external connection to motivation: even in the case of practically rational agents there is no necessary or conceptual connection between moral judgement and motivation. Examples of Humean and Externalist Cognitivism include Sturgeon (1988), Brink (1989), Railton (1986), and Boyd (1988).

In contrast with Non-Cognitivism and Externalism, *Anti-Humeanism* retains (1) and (2) by giving up (3): it rejects the Humean theory of motivation. Moral judgements express beliefs capable of motivating action in the absence of (independently intelligible) desires. Purveyors of Anti-Humean Cognitivist Internalism include McDowell (1979, 1981), Wiggins (1993), and Platts (1979, 1981), and more recent developments in the same tradition as McDowell, Wiggins and Platts might include Shafer-Landau (2003) and Huemer (2005).

It is no part of our brief in this paper to argue in favour of one of these solutions or to argue that there is a genuine inconsistency between (1), (2) and (3): we do not wish to foreclose the possibility that there might be an analysis of moral judgement that shows that all three propositions can be held without inconsistency.⁴ Our main point is that there is a very important metaethical theory that appears not to have a place on Smith’s map.

3. Error Theory

One position that doesn’t appear to have a place on Smith’s map is the error theory developed by J. L. Mackie (1946, 1977), despite the fact that in his opening chapter it is the first metaethical view mentioned by Smith in his

³ This is of course very broad brush, as some of the philosophers on this list would disavow the “Non-Cognitivist” label. But it is standard to include Hare, Blackburn and Gibbard in the same strand of the metaethical tradition as Ayer and Stevenson. The complications involved don’t matter for our present purposes.

⁴ Smith himself argues against each of the Non-Cognitivist, Externalist and Anti-Humean solutions, and proposes – via a “response-dependence” analysis of moral judgement, a view which is Cognitivist, Internalist and Humean. See Chapter 6 of Smith 1994.

section “Meta-ethics Today”: according to Mackie’s error theory “engaging in moral practice presupposes that there exist moral facts, and that this presupposition is an error or mistake akin to the error of presupposition made by someone who engages in a religious practice when there is in fact no God” (1994: 3). Mackie holds that our concept of a moral fact is a concept of a categorical reason for action: a moral fact would be a fact capable of motivating any rational agent independently of their contingent desires. But the existence of such a fact would mean that the universe itself contained “to be done-ness”, or intrinsically normative states of affairs, something rejected by Mackie in his famous “Argument from Queerness” (1977: 38-42). Moral judgements therefore express beliefs whose truth would require the existence of a type of fact which does not actually exist and are therefore systematically false.⁵ Despite this, Mackie is not an eliminativist about moral judgement: although moral judgements are systematically false, some moral judgements are justified in the sense that their adoption assists groups of humans to garner the benefits of social co-operation (1977: Chapter 5).

Mackie’s error theory thus contains a negative thesis (moral judgements are systematically false) and a positive thesis (despite the systematic falsity of moral judgements there is still a point to the practice of making moral judgements). Now Mackie’s view has been extensively criticized in the literature (see Miller (2013: chapter 6) for an overview). We are not concerned here with the question of its plausibility or otherwise, just with the question as to where it appears in the taxonomy generated by Smith’s “Moral Problem”. For there appears to be nowhere for it go. The error theory is cognitivist (moral judgements express false *beliefs*), but also internalist and Humean. Smith himself sees the error theorist as committed to the internalist claim (2), in virtue of the fact that he accepts the claim that our concept of a moral fact is a concept of a reason for action, and that (according to Smith) this claim implies (2) (Smith 1994: 62). Moreover, moral facts are problematic for the error theorist precisely because an agent who judged that a moral fact obtained would be motivated to act in the absence of appropriate desires, something ruled out by Humeanism about motivation.⁶ The error theorist thus accepts (1), (2) and (3) and so cannot be represented as attempting to solve the “Moral Problem” by rejecting one of its constituent propositions. This severely limits Smith’s claim that the “Moral Problem” is the “central organizing problem in contemporary metaethics”.

⁵ Strictly speaking, the claim of falsity is restricted to atomic, positive moral judgements (since the negation of a false moral judgement will be true).

⁶ We’ll see below that the claim about judgement and motivation is subject to a restriction that we don’t mention here. See sections 4 and 5.

Smith himself describes the error theory characterized as above as involving “moral nihilism” (1994: 11), and he also characterizes moral nihilism as involving the idea that no proposition in the trio of (1), (2) and (3) can be justifiably rejected to solve the “Moral Problem” (1994: 13).⁷ Perhaps, then, the error theory does appear in Smith’s cartography: it’s the view that the “Moral Problem” can’t in fact be solved by justifiably rejecting one of the propositions which gives rise to it.

This won’t do, however, as the “moral nihilism” that concedes that the “Moral Problem” can’t be solved would amount to the view that “the very idea of morality [is] altogether incoherent” (1994: 5), incoherent in the sense of containing a conceptual contradiction or irresolvable tension of the sort apparently generated by the conjunction of (1), (2) and (3). This is not a good fit for Mackie’s error theory: Mackie’s view, after all, is not the view that moral practice is in some deep sense incoherent, but rather the view that although moral practice is coherent, nothing in fact corresponds to it in reality. If moral practice were “altogether incoherent” as per Smith’s characterization of moral nihilism, it would presumably be impossible to be anything other than an *eliminativist* concerning it. And as we saw above, Mackie’s error theorist is not an eliminativist. One way of framing this point would be to characterise Mackie’s error theory as a form of *weak moral nihilism* and the view that the Moral Problem cannot be solved as a form of *strong moral nihilism*. The criticism of Smith would then be that he conflates the strong and weak forms of moral nihilism.⁸

What this shows is that if we want the “Moral Problem” to encompass a position as important as Mackie’s error theory – which it surely should do to count as the “central organizing problem in contemporary metaethics” – we will need to refine Smith’s version of the problem to allow the error theory to occupy a place alongside Non-Cognitivism, Externalism and Anti-Humeanism. We attempt this in the next section.

4. The Refined Moral Problem

We suggest that Smith’s version of the “Moral Problem” be expanded and refined along the following lines:

⁷ Smith writes, of Non-Cognitivism, Externalism and Anti-Humeanism: “no matter which proposition these philosophers choose to reject, they are bound to end up denying something that seems more certain than the theories they themselves go on to offer. Moral nihilism quite rightly looms” (1994: 13).

⁸ Perhaps another way of making this point would be to say that Smith treats Mackie’s view of moral judgements as more akin than it actually is to Churchland’s view of propositional attitude discourse (Churchland 1981). And the point is that Churchland’s view is eliminativist in a way in which Mackie’s is not.

1. There is at least one moral judgement of the form “It is right that I Φ ” such that this judgement
 - a. Expresses a subject’s belief about an objective matter of fact, a fact about what is right for her to do.
 - b. Is true.
 - c. Is justified.
2. It is a conceptual truth that: if someone knows that it is right to Φ and is practically rational then, *ceteris paribus*, she is motivated to Φ .
3. An agent is motivated to act in a certain way just in case she has an appropriate desire and means-end belief, where belief and desire are, in Hume’s terms, “distinct existences”.

For the purposes of this discussion, we will construe *knowledge as justified true belief*. This position is now not commonly held as doctrine thanks to the development of *Gettier counterexamples* (Gettier, 1963; 122-123). However, despite Gettier’s objection, on most characterisations of knowledge belief and truth are at least necessary conditions of knowledge. And in any event, those who prefer can substitute their favoured definition of knowledge.⁹

Our changes to Smith’s triad of propositions also includes a revision of his internalist proposition (2). While Smith formulates (2) in terms of *judgement* internalism, we instead prefer a formulation in terms of *knowledge* internalism. Knowledge internalism is the thesis that in a practically rational agent there is a conceptual (and hence necessary) connection between possessing moral knowledge and being motivated to act accordingly. Only when the judgement constitutes knowledge is there a necessary connection between that judgement and motivation. There is no commitment to anything less than knowledge having that necessary connection.

The quintet of propositions we now have still retains Smith’s *prima facie* tension. According to proposition 1, there is a moral judgement such that this moral judgement will express a belief, be true, and be justified. So, there is at least one moral judgement that will be an instance of moral knowledge. From proposition 2, moral knowledge has a necessary connection with being motivated. So, by proposition 3, moral knowledge – and hence the relevant

⁹ This allows us a degree of flexibility around the definition of knowledge in formulating the Refined Moral Problem. For example, those who might defend a reliabilist account of knowledge could adapt proposition 1 to include a requirement (instead of, or in addition to c) that the belief expressed by the relevant judgement is formed by some reliable process.

belief - must have a necessary connection with some desire. But also, by 3, this desire state must be distinct from the belief, even if that belief constitutes an instance of knowledge. So, Smith's *prima facie* tension is preserved in the Refined Moral Problem.

Importantly, since the "Refined Moral Problem" consists of *five* rather than three propositions, space is opened up for two further ways of defusing the problem. One of these is the error theory, which specifically rejects 1b.¹⁰ Thus, Mackie's view can now appear alongside Non-Cognitivism, Externalism, and Anti-Humeanism on the map of responses to the refined version of the problem. We suggest that this constitutes an important improvement to Smith's original formulation.

5. Why Knowledge Internalism?

Why formulate internalism in terms of a necessary connection between moral knowledge and motivation rather than in terms of a necessary connection between moral judgement and motivation (irrespective, in other words, of whether the judgement constitutes knowledge)? And is knowledge internalism philosophically defensible? We will not attempt to answer these important questions here.¹¹ Instead, we limit ourselves to two observations.

First, Mackie himself explicitly deploys what appears to be a form of knowledge internalism in the course of his Argument from Queerness. Mackie says

The form of the Good is such that *knowledge* of it provides the knower with both a direction and an overriding motive, something's being good both tells the person who knows it to pursue it and makes him pursue it... The need for an argument of this sort can be brought out by reflection on Hume's argument that 'reason' - in which at this stage he includes all sorts of knowing as well as reasoning - can never be an 'influencing motive of the will' (Mackie, 1977: 40, emphasis added).

The error theorist accepts that the conceptual profile of moral facts implies that knowledge of them has a necessary influence on motivation. So, the error theorist is very willing to accept knowledge internalism as a conceptual thesis. It poses no problem for the error theorist to reconcile this with the

¹⁰ The other way of defusing the problem would be to reject 1c. and deny that any moral beliefs are justified - to adopt, in other words, a form of epistemological scepticism about moral judgement. We won't attempt to discuss this here, but note that this makes for a new and potentially interesting connection between metaethics and moral epistemology.

¹¹ For a preliminary attempt at doing so, see Whittington 2023.

Humean theory of motivation because they believe that no moral judgement is true, so no moral judgement would ever constitute knowledge. Due to its conditional nature, knowledge internalism is just a benign thesis about non-instantiated facts. The fact that it is knowledge internalism that plays a central role in Mackie's Argument from Queerness – possibly the most famous argument in metaethics in the second half of the 20th Century – arguably provides at least a *prima facie* case for giving it a central role in any problem with aspirations to count as the “central organizing problem of contemporary metaethics”.

Second, the focus on knowledge as opposed to judgement internalism allows us to close an apparent gap in one of Smith's arguments. Smith (1994: 62) characterises *rationalism* as the view that our concept of a moral fact is a concept of a reason for action, and he argues that rationalism entails (judgement) internalism, but not vice-versa (Smith 1994: 62). Smith's argument that rationalism entails (judgement) internalism goes as follows.

Suppose that rationalism is true: that our concept of a moral fact is a concept of a reason for action. Suppose that Octavia judges that it is right for her to stay home when sick. Then, by the rationalist claim, she judges that she has a reason to stay home when sick. Plitudinuously, Octavia will have a reason to stay home when sick just in case she would be motivated to stay home when sick were she rational. Thus, Octavia judges that she would be motivated to stay home when sick if she were rational. Now if she fails to be motivated to stay home when sick she is irrational “by her own lights” (1994: 62). Thus, if Octavia judges that it is right for her to stay home when sick then, unless she is practically irrational, she will be motivated to stay home when sick. And this last claim is just a statement of the internalist thesis (2) in Smith's “Moral Problem”. So rationalism entails internalism.

This is interesting, because it seems to allow Smith to distinguish between stronger and weaker commitments to internalism. Although (according to Smith) rationalism entails internalism, the converse implication does not hold. Non-cognitivists accept internalism, but they hold that our concept of a moral requirement is not a concept of a reason for action. According to non-cognitivists “rational agents may ... differ in their moral judgements ... without being in any way subject to rational criticism” (1994: 86).

However, as it stands, Smith's argument that rationalism entails (judgement) internalism contains a gap. The fact that I am irrational “by my own lights” only entails that I am genuinely irrational if my judgement about what I would be motivated to do if I were rational is *true*. If my judgement about what I would be motivated to do if I were rational is false, nothing follows as to my status as genuinely rational or not if I'm not suitably motivated.

Clearly, this problem would not afflict the corresponding argument to the effect that rationalism entails knowledge internalism: if I know that if I were rational I'd be motivated to stay at home when sick, then it is true that if I were rational I'd be motivated to stay home when sick. My failing to be so motivated would indeed be a mark of irrationality.

So, switching to knowledge internalism allows us to preserve Smith's claim about the relationship between rationalism and our favoured form of internalism. Afficionados of chapter 3 of Smith's book will appreciate the potential importance of this fact.¹²

We suggest that the Refined Moral Problem, and knowledge internalism, would be well worthy of more metaethical scrutiny than they have hitherto received.¹³

References

- Ayer, A.J. 1946: *Language, Truth, and Logic*. London: Gollanz (Second Edition).
- Blackburn, S. 1984: *Spreading the Word*. Oxford: Oxford University Press.
- Blackburn, S. 1998: *Ruling Passions*. Oxford: Clarendon Press.
- Boyd, R. 1988: How to be a moral realist. In G. Sayre-McCord (ed), *Essays on Moral Realism*, Ithaca: Cornell University Press.
- Brink, D. 1989: *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Churchland, P. 1981: 'Eliminativist Materialism and the Propositional Attitudes', *Journal of Philosophy* 78: 67–90.
- Gettier, E. 1963: "Is Justified True Belief Knowledge?", *Analysis*, 23(6): 121–123.
- Gibbard, A. 1990: *Wise Choices, Apt Feelings*. Oxford: Clarendon Press.
- Gibbard, A. 2003: *Thinking How to Live*. Cambridge MA: Harvard University Press.
- Hare, R. 1952: *The Language of Morals*. Oxford: Oxford University Press.
- Huemer, M. 2005: *Ethical Intuitionism*. Basingstoke: Palgrave Macmillan.
- Mackie, J.L. 1946: "A Refutation of Morals", *Australasian Journal of Philosophy* 24: 77–90.
- Mackie, J.L. 1977: *Ethics: Inventing Right and Wrong*. New York: Penguin.

¹² Note that the central point of this section does not undermine the overall argument of sections 3 and 4: representing the error theorist as solving the "Moral Problem" by rejecting judgement-internalism would fail to distinguish the error theory from the likes of the externalism favoured by Sturgeon, Boyd and Brink. So space would not yet have been found on the map as a distinctive place to be occupied by the error theory.

¹³ Thanks to Charles Pigden and an anonymous referee for *Filosofiska Notiser* for helpful comments, and to the Daniel Taylor Scholarship for partial funding of this research.

A Note on Error Theory and the Refined Moral Problem

- McDowell, J. 1979: "Virtue and Reason", in his *Mind, Value and Reality* (Cambridge MA: Harvard University Press 1998).
- McDowell, J. 1981: "Non-Cognitivism and Rule-Following", in his *Mind, Value and Reality* (Cambridge MA: Harvard University Press 1998).
- Miller, A. 2013: *Contemporary Metaethics: An Introduction*, Revised and Expanded 2nd Edition, Cambridge: Polity Press.
- Platts, M. 1979: *Ways of Meaning*. London: Routledge and Kegan Paul.
- Platts, M. 1981: "Moral Reality and the End of Desire", in M. Platts (ed.) *Reference, Truth and Reality*. Routledge and Kegan Paul.
- Shafer-Landau, R. 2003: *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Smith, M. 1994: *The Moral Problem*. Oxford: Blackwell.
- Stevenson, C. 1937: "The emotive meaning of ethical terms". Reprinted in A.J. Ayer (ed) *Logical Positivism*. Glencoe: Free Press.
- Stevenson, C. 1944: *Ethics and Language*. New Haven: Yale University Press.
- Sturgeon, N. 1988: "Moral Explanations". In G. Sayre-McCord (ed), *Essays on Moral Realism*. Ithaca: Cornell University Press.
- Whittington, S. 2023: *The Refined Moral Problem: A Case for Knowledge Internalism*. University of Otago, MA Dissertation.
- Wiggins, D. 1993: "Cognitivism, naturalism, and normativity". In J. Haldane and C. Wright (eds.) *Reality, Representation and Projection*. Oxford: Oxford University Press.

Alexander Miller and Seth Whittington
University of Otago, New Zealand
alex.miller@otago.ac.nz (Alexander Miller)
whise654@student.otago.ac.nz (Seth Whittington)