



BRILL



Obligations of Conscience

Shane N. Glackin

Department of Sociology, Philosophy, and Anthropology,
University of Exeter, Exeter, United Kingdom
s.n.glackin@exeter.ac.uk

Abstract

In this paper, I outline and defend a commonly-held moral view which has received surprisingly little sustained philosophical attention. This view, which I call the ‘authority of conscience,’ states that believing ourselves to have moral obligations to act in a certain way does in fact create an obligation to act in that way. Although I do not provide a positive case for the principle of authoritative conscience, beyond its popularity and intuitive force, I defend it against several *prima facie* objections. I then go on to demonstrate that the principle does not entail any anti-realist metaethical commitments, and is therefore compatible in particular, and contrary to appearances, with plausible formulations of moral realism.

Keywords

conscience – deontic logic – moral realism – moral dilemma – obligation

1 Introduction

In this paper, I elaborate and defend a widespread, but philosophically little-discussed, intuition about morality and practical reason more generally: the view that we are morally bound by the dictates of conscience. That is, the fact that we believe ourselves to have particular moral obligations imposes on us a moral obligation to act on those beliefs, even if the beliefs in question are false.

The paper has two main sections. In the first, I will elaborate the intuition I will call ‘the authority of conscience’ – henceforth ‘AC’ – and defend it against

a number of common *prima facie* objections. Although I do not advance a positive argument for the position, the detailed elaboration of its features in the course of these defenses will hopefully make its attractions clearer than they have hitherto been. In the second section, I will show that AC is, in particular, compatible with the metaethical position known as *moral realism*, which states that facts about moral obligations are independent of anything we may believe about them. It is therefore possible to hold both that moral realism is correct, and that we are ‘really’ – in whatever sense – morally bound by the dictates of conscience. I outline one version of moral realism which shows particularly clearly the relevant features of realism: I demonstrate that AC is compatible with it, before arguing that the point generalizes, and that any version of moral realism which does not permit AC carries the implausible consequence that ‘tragic dilemmas’ cannot exist. By way of a ‘proof of concept’ I then briefly consider two further particular, and recently influential, versions of moral realism, and show that they fall under the analysis given.

2 The Authority of Conscience

According to a widespread¹ intuition about practical reason, which has nevertheless received surprisingly little philosophical attention, we are morally bound by the dictates of conscience. That is, if we sincerely believe ourselves to be under a moral obligation to behave in a certain way, then we do indeed

1 I cannot give a comprehensive defense here of the empirical claim that the intuition is ‘widespread.’ However, public controversies over, *e.g.* the right of pacifists to refuse conscription, the right of pro-life doctors to refuse to perform abortions in jurisdictions where they are legal, the right of religious pharmacists to refuse to stock or sell emergency contraceptives, or the right of a conservative baker to refuse to cater for a same-sex wedding, evidently demonstrate a common – not to say universal – discomfort with the notion that someone can or should be forced to perform actions which they sincerely believe to be seriously morally wrong. It should be stressed that support for AC does not entail unqualified support for any of these stances; one might hold, for instance, that pharmacists who do not wish to dispense ‘morning-after’ pills to those who need them need not be compelled to do so, but should certainly be made to find other employment. The reader may nonetheless find it useful to bear examples like these in mind during the discussion which follows. John Gardner (‘Wrongs and Faults,’ *The Review of Metaphysics*, 59 (2005), pp. 95–132), moreover, identifies a closely-related intuition as motivating the *mens rea* principle in criminal law: by the Rule of Law, ‘we should be put on stark notice’ that we are about to commit a crime if we are to be justly punished for doing so, since ‘(t)he criminal law should not ambush us unexpectedly.’ Thus, he argues, ‘according to the *mens rea* principle, criminal wrongs should be such that one does not commit them unless one intends or is aware of at least one wrong-making feature of what one is about to do’ (p. 119).

have a moral obligation to act accordingly. Whatever competing moral-theoretical claims may tell us, we must refrain from actions which we believe to be wrong, and perform those which we believe to be right.² It is commonly accepted that we can performatively impose (defeasible) moral obligations on ourselves by, for instance, making a promise;³ AC – the principle of *the Authority of Conscience* – merely adds the act of making certain sorts of judgement, or forming certain sorts of belief, to the list of deeds by which we bind ourselves in this way. Of course, I do not mean by this to suggest that judgements of conscience operate in just the same way as acts of promising. As with the uttering of promises, though, the fact that the judgement is made or the belief is formed for any number of defective reasons does not lessen the force of the (again, defeasible) obligation it imposes.

‘Conscience,’ as it is understood here, does not denote a metaphysically distinct faculty; it is not envisioned as an inerrant voice of Divine Guidance of the sort associated with Bishop Butler.⁴ Nor is it even necessarily psychologically distinctive, like the ‘phenomenological conscience’ discussed by C.D. Broad,⁵ or what Twain’s Huckleberry Finn takes himself to be acting against when he decides not to turn the escaped slave Jim in.⁶

2 I shall simply assume here that the principle operates symmetrically. One reader has objected that it is much less plausible that one acts rightly if one acts in a way that one considers to be right, than that one acts wrongly if one acts in a way that one considers to be wrong (*cf.* Philippa Foot, *Moral Relativism: The Lindley Lecture, 1978* (Lawrence KS: University of Kansas Press, 1979)). It seems clear to me, however, that – familiar worries about the ‘distinction between doing and allowing’ notwithstanding – the same moral imperatives may as easily be phrased in negative as in positive terms; and standard deontic logic appealed to elsewhere in this paper, which assumes that ‘right’ and ‘wrong’ (or ‘permissible’ and ‘impermissible’) are modal operators which take propositions as objects, proceeds on the same assumption (*cf.* David K. Lewis, ‘Semantic Analyses for Dyadic Deontic Logic,’ in D. K. Lewis, *Papers in Ethics and Social Philosophy* (Cambridge: Cambridge University Press, 2000), pp. 5–19). For a contrasting view, however, see Susan Wolf, ‘Asymmetrical Freedom,’ *The Journal of Philosophy* 77 (1980), pp. 151–66.

3 See, *e.g.* J. L. Austin, ‘Performative Utterances,’ in J. L. Austin, *Philosophical Papers* (Oxford: Oxford University Press, 1979), pp. 233–252; Thomas Pink, ‘Promising and Obligation,’ *Philosophical Perspectives* 23 (2009), pp. 389–420.

4 See, *e.g.* Joseph Butler, ‘Dissertation II: Of the nature of virtue,’ in L.A. Selby-Bigge (ed.), *British Moralists* (Oxford: Clarendon Press, 1897), §§244–250. Most Butler scholars now regard the ascription of this view to him as erroneous; see Bela Szabados, ‘Butler on Corrupt Conscience,’ *Journal of the History of Philosophy* 14 (1976), pp. 462–469.

5 C. D. Broad, ‘Conscience and Conscientious Action,’ *Philosophy* 15 (1940), pp. 115–130.

6 See, *e.g.* Jonathan Bennett, ‘The Conscience of Huckleberry Finn,’ *Philosophy* 49 (1974), pp. 123–134.

Rather, conscience in this common, secular sense represents our (potentially mistaken, of course) psychological awareness – the etymological link with ‘consciousness’ is telling – and response to the moral status of our actions.⁷ ‘It is most commonly thought of,’ moreover, as Christine Korsgaard writes, ‘as the source of pains we suffer as a result of doing what we believe is wrong – the pains of guilt, or “pangs of conscience.”’⁸ And the Authority of Conscience as I discuss it here – and the sort of normative pull it exerts – can therefore perhaps be best understood as a subset or part of a rational requirement against *akrasia*, or ‘weakness of will.’ On standard accounts of *akrasia*, if I believe that I ought to take some course of action, I have the opportunity to take it, and I nevertheless fail to do so, my conduct is thereby open to at least some sort of censure; I have failed in at least one way to do as I ought. The Authority of Conscience will therefore apply when the belief in question is a moral one, an expression of the moral status of my actions; AC states that where I believe that I ought *morally* to take some course of action, have the opportunity, and nevertheless fail to do so, my failure is a specifically moral one.⁹

Perhaps it is not so clear after all that *akrasia* regarding moral beliefs involves a moral failure, rather than a mere failure of practical rationality; it could be objected, for example, that *akrasia* represents a negative mark on the agent’s character, but not necessarily on anything that the agent does. So if someone mistakenly believes that she is obligated to attend church on Sunday, and fails to do so out of laziness, we might want to say that this shows bad character on her part even though her failure to attend church is not even a little bit wrong.¹⁰ But the key point here, I think, is the nature of the felt obligation. In this example, it is not clear that the obligation in question is a specifically moral – rather than, say, a social or liturgical – one. In unambiguously moral cases, we have less difficulty. Suppose our agent believes that her elderly neighbor will freeze unless she switches his heating on, but fails to do so out of laziness. There are all sorts of reasons she might be mistaken about her obligation, some more epistemically culpable than others; someone else might have switched

7 Christine Korsgaard, ‘Conscience,’ in J.J. Chambliss (ed.), *Philosophy of Education: An Encyclopedia* (London: Routledge, 1996).

8 *ibid.*

9 One notable historical precursor of this view is that of St. Bonaventure; see his ‘Conscience and Synderesis,’ in A.S. McGrade, J. Kilcullen, and M. Kempshall (eds.), *The Cambridge Translations of Medieval Philosophical Texts, Volume Two: Ethics and Political Philosophy* (Cambridge: Cambridge University Press, 2001), pp. 169–199 and Douglas C. Langston, *Conscience and Other Virtues: From Bonaventure to MacIntyre* (University Park PA: The Penn State University Press, 2001), Ch. 2. I thank an anonymous referee for the connection.

10 I thank an anonymous referee for this objection.

her neighbor's heating on already without informing her, or warmer weather than she expects might be forecast, or she may for no good reason at all expect it to snow in August. These cases, I hold, are differentiated only by the degree of epistemic warrant our agent holds for the pertinent belief; since this is a non-moral element, they are morally indistinguishable. So her akratic failure to act on any such belief is *morally*, and not merely practically, reprehensible, even though her neighbor is not harmed by it. This view of our agent's actions is, in essence, the central commitment of this paper; while I do not provide a positive argument for it, and so will perhaps not change the minds of readers with different intuitions about the case, I will show that holding it is unproblematic.

We can formulate the principle more precisely. Moral agents are assumed, *qua* moral agents, to have genuine, binding obligations.¹¹ On most views, those obligations are in some manner imposed by some set of facts about the world.¹² But according to AC, one sub-class of the facts which thus impose obligations is a particular sort of psychological fact about the obligated person, namely that they believe themselves so obligated. The truthmakers of moral claims, in other words, include not just 'external' states of affairs, but at least some beliefs about those states. At the risk of excessive formality:

AUTHORITY OF CONSCIENCE (AC): For any individual *a*, the set of facts $\{f_1, f_2, f_3 \dots f_n\}$ which, if they obtained, would be severally sufficient to oblige her to perform some action *x* include *f* Ψ , the fact that *a* believes there to obtain a number of $\{f_1, f_2, f_3 \dots f_n\}$ sufficient to oblige her to perform *x*.

To be clear: AC does not entail that *a*'s believing herself to have obligations imposes those obligations *per se*. Rather, it states that those obligations may be imposed by her belief *that the same obligations are imposed by the relevant facts*.

¹¹ Of course, not everyone will be comfortable with talk of 'obligations.' Though I use that locution as a placeholder, there is nothing intrinsically deontological about my argument; analogous formulations should go through for the building-blocks of any normative ethical theory. Similarly, and for purposes of simplicity and brevity only, I treat 'obligation' as a simple, binary deontic operator; one is either obliged to do something or one is not, and obligation does not come in degrees of force. That assumption is as a substantive issue controversial, however, and while I do not believe that any of the arguments here are dependent upon it, proper consideration of that question is beyond the scope of this paper.

¹² Here and throughout, I shall talk as if what makes moral claims true are 'facts' or 'states of affairs,' and disregard well-known controversies about the 'fact/value distinction.' I do so purely for convenience; the argument I pursue here is independent of that dispute.

'The relevant facts,' in other words, properly include her having certain beliefs about those same facts as well as about the obligations those facts impose. So external states can impose moral obligations, and so can our considered, but fallible, judgements about those external states and the obligations they impose. With respect to their ability to generate moral obligations, there is no significant distinction to be drawn between external and psychologically internal states.

The principle seems intuitively obvious to me, to the extent that I have difficulty formulating a positive argument for it beyond the standard objections to *akrasia*. But it is nonetheless controversial; I have been informed both that the thesis of this paper is so obviously true as not to be worth stating, and that it is trivially false. What I will do, therefore, in lieu of such an argument, is to observe briefly that something of the sort may be seen to accord with the significant majority of mainstream positions in normative ethics,¹³ before addressing some of the chief concerns about and objections to such a principle. In other words, I will content myself here with showing that the principle is not obviously erroneous, in the hope that its intuitive appeal can complete the work of persuasion.

The first obvious worry is that such a principle may give rise to contradictory obligations, so that an agent is obliged to perform some action by her false beliefs about the state of the world, and simultaneously obliged to refrain from the same action by the actual state of the world. And indeed, it is true that it may: but it is not clear that this presents a serious problem. Certainly, as Jonathan Dancy has shown, any such contradictions are in purely formal terms

13 *E.g.* for Aristotle's criticisms of *akrasia*, see book VII of the *Nicomachean Ethics*. For Kant's account of conscience, see Immanuel Kant *The Metaphysics of Morals*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1991), pp. 160ff and Emre Kazim, *Kant on Conscience: A Unified Approach to Moral Self-Consciousness* (Leiden: Brill, 2017); for recent work on conscience and Kantian 'frailty,' see J. David Velleman, 'The Voice of Conscience,' *Proceedings of the Aristotelian Society* 99 (1999), pp. 57–76 and Seiriol Morgan, 'The Missing Formal Proof of Humanity's Radical Evil in Kant's Religion,' *The Philosophical Review* 114 (2005), pp. 63–114, 67–8. For Mill's view, see John Stuart Mill, 'Utilitarianism,' in J.S. Mill, *Collected Works vol. 10* (Toronto: University of Toronto Press, 1969), p. 228, and for versions of consequentialism taking seriously the normative authority of 'our best reflective judgement,' see Brian McElwee, 'The Rights and Wrongs of Consequentialism,' *Philosophical Studies* 151 (2010), pp. 393–412 and Frank Jackson, 'Decision-theoretic Consequentialism and the Nearest and Dearest Objection,' *Ethics* 101 (1991), pp. 461–82. For general historical overviews, see John Skorupski, 'Conscience,' in J. Skorupski (ed.), *The Routledge Companion to Ethics* (London: Routledge, 2010), pp. 550–561 and Alberto Giubilini, 'Conscience,' in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (2016) <<https://plato.stanford.edu/archives/win2016/entries/conscience/>>.

easily resolvable as long as we take the obligation to hold not over the action alone, but rather over the relation between the action and the beliefs which motivate it.¹⁴ That is, while $O(p)$ and $O(\neg p)$ may contradict each other, $O(x \rightarrow p)$ and $O(y \rightarrow \neg p)$ do not, even if x and y are both true. I discuss this point in more detail in section §3 below.

Moreover, as Aurel Kolnai points out, the attempt to avoid such contradictions must lead to one or other of two converse errors: *formalism*, which holds ‘that the agent’s acting or not in conformity with his conscience is the only test of his behaving rightly or wrongfully’; and *intellectualism*, which disregards subjective ‘good faith’ entirely, claiming that ‘right conduct depends on true moral knowledge ... alone, and all wrongful conduct expresses a moral error.’¹⁵ Both positions look erroneous, but there is no need to choose between them by making any special effort to avoid contradictory obligations; we can avoid both Scylla and Charybdis by recognizing that one’s conscience – one’s sincere, reflective belief about what it is right to do in a given circumstance – may simply be *wrong*, even though that error will make no difference to its phenomenologically binding normative force. In other words, it is quite enough as far as avoiding such contradictory obligations is concerned that we make our normal epistemic efforts to avoid the sort of false views about the world that give rise to them; adherence to the dictates of conscience – however mistaken those dictates may be – ‘expresses and presupposes the agent’s general response, assent and submission to the valid intrinsic principles of morality as we know them,’ and to cease following them would be to cease acting from moral motivations at all.¹⁶ The formalist route, then, denies the very obvious fact of human fallibility, while the intellectualist option ignores the significance of our moral agency. The middle route to avoiding contradictory obligations, by contrast, is simply to strive for true beliefs and act on the beliefs we have, eschewing any special further procedure to insulate against error.

A second ground for objection to the principle is that it permits something known as ‘bootstrapping,’ whereby merely thinking that something is the case – in this instance, that I have an obligation to do something – makes it in fact the case. ‘This,’ objects Niko Kolodny, ‘is absurd.’¹⁷ Indeed, he repeats the charge of absurdity on several occasions, without ever saying what, exactly, is supposed to be absurd about bootstrapping in the moral case.¹⁸ Perhaps its

14 Jonathan Dancy, ‘The Logical Conscience,’ *Analysis* 37 (1977), pp. 81–84.

15 Aurel Kolnai, ‘Erroneous Conscience,’ *Proceedings of the Aristotelian Society* 58 (1958), pp. 171–198, 172.

16 Kolnai, ‘Erroneous Conscience,’ p. 173.

17 Niko Kolodny, ‘Why be Rational?’, *Mind* 114 (2005), pp. 509–563, 512.

18 Kolodny, ‘Why be Rational?’, Niko Kolodny, ‘State or Process Requirements?’, *Mind* 116 (2007), pp. 371–385; Niko Kolodny, ‘Reply to Bridges,’ *Mind* 118 (2009), 369–376.

absurdity seems intuitively obvious to him; this would be perfectly reasonable, and a similar intuition will be shared by many. I could simply respond here by denying that I share the intuition – after all, one plausible explanation of the institution of promising is that we can ‘bootstrap’ obligations into being just by *saying* that we have them – and hold up the intuitive appeal of AC in response. But there is a more persuasive, if more technically involved, line of argument available.

John Broome, who agrees that bootstrapping would be problematic,¹⁹ denies that it is entailed by enkratic (that is, anti-akratic) normative requirements like AC, distinguishing the notion of a ‘normative requirement’ from the more familiar normative concepts of ‘ought’ and ‘have a reason to.’ Like ‘ought,’ but unlike the reason-having relation, the appropriate enkratic relation between our beliefs and actions – in his terminology – is ‘strict’ rather than ‘slack,’ meaning that ‘[i]f you believe you ought to see to it that *r*, but you do not see to it, you are definitely failing in one respect. You are definitely not entirely as you ought to be.’²⁰ But if bootstrapping is disallowed, the appropriate relation cannot just be the *ought*-relation, argues Broome; believing that something is the case does not make it so. So while there is an *ought*-component to the relation, Broome reasons, it is not ‘detachable’; normative requirements such as AC must be ‘wide’ rather than ‘narrow’ in scope.²¹

What Broome means by this is that normative requirements have the structure ‘ $O(p \rightarrow q)$ ’ rather than ‘ $p \rightarrow O(q)$ ’: not ‘if you think ... then you ought to *q*,’ but ‘you ought, if you think ... to *q*.’²² Thus, the normative requirement does not allow a normatively loaded consequent (*i.e.* ‘ $O(q)$ ’) to be detached from its antecedent by *modus ponens*, thereby preventing the bootstrapping move. That this structure is ‘widely ignored,’ Broome attributes to

an idiom of English. When a conditional proposition contains a modality, we tend to attach the modality to the consequent, even if that is logically not the correct place for it. We say, for instance, ‘If it is raining, it must be thawing.’ We do not mean, ‘If it is raining, necessarily it is thawing,’ but,

19 See, *e.g.* John Broome, ‘Normative Requirements,’ *Ratio* XII (1999), pp. 398–419, 404–5; John Broome, ‘Wide or Narrow Scope?’, *Mind* 116 (2007), pp. 359–370.

20 Broome, ‘Normative Requirements,’ p. 404.

21 Broome, ‘Normative Requirements,’ p. 410.

22 Broome, ‘Normative Requirements,’ p. 402. Strictly, Broome says that this form is merely a *consequence* of normative requirements, since it contains only a material conditional and is thus prone to counterexamples, but we may safely disregard these for the sake of clarity. Cf. also Lennart Åqvist, ‘Good Samaritans, Contrary-to-duty Imperatives, and Epistemic Obligations,’ *Noûs* 1 (1967), pp. 361–79.

'Necessarily, if it is raining, it is thawing.' ... Similarly, the proposition that believing p normatively requires you to believe q would be idiomatically expressed by: 'If you believe p , you should believe q .' ... These expressions are good idiomatic English, but they misrepresent the logic of what is said.²³

Jonathan Dancy, indeed, identified the same 'ambiguity ... cloaked in natural language' as motivating the earlier worry about contradictory obligations,²⁴ and Broome's approach may be seen as an elaboration of his.²⁵ Fulfilling normative requirements, Broome goes on to demonstrate formally, 'is unaffected by the choice between narrow and wide scope. ... Either way, you are rational at exactly the same worlds.'²⁶

This analysis ties felicitously to the issue of erroneous conscience in another way. The fundamental worry there was that AC would generate obligations to act on false beliefs, thereby conflicting with the obligations generated by the non-doxastic or 'external' facts of the matter. But our new structure for normative requirements – ' $O(p \rightarrow q)$ ' – is equivalent to an obligation over a disjunction, ' $O(\neg p \vee q)$.' In other words, as Julia Markovits writes, there are 'two ways to satisfy such requirements: we can change our actions, or we can change our beliefs about what we ought to do.'²⁷ So given AC, a belief will generate an obligation to *either* act on the belief *or* abandon or revise it, the latter option obviously being the normatively appropriate one when the belief in question is false or irrational. However, in rare cases where the option to change the belief is unavailable – where the antecedent is necessary – a form of bootstrapping *is* permissible. That is to say: $O(q)$ follows validly from $O(p \rightarrow q) \cap \Box(p)$, although not from $O(p \rightarrow q) \cap p$.²⁸ Note that the agent must be *unable* to revise

23 Broome, 'Normative Requirements,' p. 412.

24 Dancy, 'The Logical Conscience,' p. 84.

25 Broome, 'Normative Requirements,' p. 404, n. 4.

26 Broome, 'Wide or Narrow Scope?,' p. 363, pp. 369–70.

27 Julia Markovits, 'Acting for the Right Reasons,' *The Philosophical Review* 119 (2010), pp. 210–242, 216.

28 This can be demonstrated simply in tableau form for any standard Deontic logic on the lines of $K\eta$ or related systems, which either contain a standard necessity operator in addition to O , or in which $O(p) \rightarrow p$, (cf. Graham Priest, *An Introduction to Non-classical Logic* (Cambridge: Cambridge University Press, 2001), p. 51):

$O(p \rightarrow q)$	$O(p \rightarrow q)$
$\Box(p)$	p, \circ
$\neg O(q)$	$\neg O(q)$
$P(\neg q)$	$P(\neg q)$
$\neg q, 1$	$\neg q, 1$

her mistaken beliefs; it will not suffice to license this bootstrapping that she merely fails or is unwilling to do so.

Of course, there are different orders of necessity. Must it be *logically*, *metaphysically*, or *physically* impossible for an agent to revise her beliefs before detachment is permitted? This would be temptingly conservative in its implications, for only in rare circumstances would an agent be unable to avoid believing some statement X , have an obligation to either disbelieve X or perform q , and thereby be obliged to perform q . But there are many cases where I cannot avoid something even though it is not in any of these technical senses *necessary*. If I play chess against Magnus Carlsen, for instance, I cannot realistically avoid losing. Suppose that I have an obligation to either defeat Carlsen at chess, or admit that he is a better chess-player than myself. From these premises, one can reasonably infer that I am obliged to declare him the better chess-player; so much weaker construals of $\Box(p)$ will license the detachment of imperative consequents. And now AC looks far too permissive, for unless we have much stronger voluntary control over our beliefs than most people think to be the case, almost any mistaken beliefs I may possess will at the given moment requiring action be impossible to change, and so generate potentially heinous obligations. A Nazi prison-guard might over time come to realize the moral error of his ideology; but given an order to execute prisoners *immediately*, he has no realistic chance to discharge his AC-obligation by the preferable route of belief-change. AC implies that if he thinks in this scenario that he should obey his orders, then indeed he should. And this, surely, makes AC untenable.²⁹

Two possible responses seem open to defenders of AC: I will not opine here as to which I find preferable, since the aim of the paper is simply to defend the general principle as not entailing unwelcome consequences. The first, more conservative strategy is to block the weakening of necessity suggested above. The detachment is formally valid if and only if the antecedent is *necessary* in

$(p \rightarrow q), 1$	$(p \rightarrow q), 1$
/ \	/ \
$\neg p, 1 \quad q, 1$	$\neg p, 1 \quad q, 1$
$p, 1 \quad x$	$p, 0 \quad x$
x	...

So detaching will be problematic only when we can give a Kripke-semantic interpretation along the lines of: $W = \{w_0, w_1\}; w_0Rw_1; \nu w_0(p) = 1, \nu w_1(p) = 0, \nu w_1(q) = 0$. For a very thorough recent discussion of detachment in deontic logics, see Sven Lauer and Cleo Condoravdi, 'Preference-conditioned Necessities: Detachment and Practical Reasoning,' *Pacific Philosophical Quarterly* 95 (2014), pp. 584–621.

29 I thank an anonymous referee for this important objection.

whatever sense, and not merely overwhelmingly likely. I *could* beat Carlsen in the right circumstances, however unlikely or atypical; after all, as an 11-year-old I recorded a draw with a senior national champion when she had to abandon a match several minutes after it began. So cases where an agent faces great practical difficulty in changing their beliefs are merely contingent cases of p , not cases of $\Box(p)$ which enable detaching. And the intuitive plausibility of the Carlsen detachment simply reflects the imprecision of ordinary language, as noted by Broome and Dancy, which formalization aims to avoid.

But a proponent of AC could also accept this weakened necessity; the consequences of doing so may be a feature or a bug, depending on the reader's taste. Suppose that we accept as valid the inference from $O(p \rightarrow q) \cap (\text{cannot})(\neg p)$ to $O(q)$. In this case, I think we should say that an agent's obligation is *relative* to the opportunity she has to change her beliefs. That is to say, 'cannot' here seems to function in a way that is relative to various factors – duration, difficulty, *etc.* – so it seems that anything we infer from its presence might be treated as relative to them in the same way. For instance, our Nazi prison guard could change his views, but not immediately. Why not, then, treat him as possessing an obligation to follow the orders he is given that is strictly relative to the short timeframe available for decision-making, but as having no such obligation relative to the longer period over which we have a reasonable moral expectation that he will alter his beliefs, let alone in absolute terms?³⁰ 'Moral relativism' is a term that makes many uneasy, but in this case seems innocuous; after all, if Kant is right that 'ought implies can,' then our more conventionally-generated moral obligations are already relative to opportunity in essentially the same way.

To summarize, then, the great majority of relevant obligations are wide-scope ones, as described by Dancy and Broome: obligations which range over an action and the beliefs which motivate it. These do not generate genuine contradictions, and nor do they permit the sort of bootstrapping that Kolodny deems absurd. In certain rare cases, however, where it is impossible – however we construe that – to alter our beliefs, it is possible to detach the obligatory action from the motivating belief, generating the sort of narrow-scope obligation which can produce genuine contradictions and license bootstrapping. But in those cases, I argue, such results are not problematic; rather than creating absurdity, allowing bootstrapping intuitively gives the *right* results in the cases

³⁰ In this case, the longer-period obligation to alter his beliefs will remain operative even after the shorter-period obligation to follow orders has been discharged. I thank an anonymous referee for pressing me to clarify this point.

just discussed, while I will argue in §3 below that the contradictions so generated reflect a real and familiar feature of our moral lives.

A caveat is in order. Broome objects to the common practice of using the term ‘subjective ought’ in place of ‘normative requirement’ on two grounds:

[f]irst, it conceals the logical structure of the situation, because it does not make the ‘ought’ govern a conditional. Second, it implicitly makes the ought relative to the wrong thing. It makes it relative to the subject, whereas it should be relative to a fact: the fact that imposes the normative requirement.³¹

The terminological point is well-taken; nonetheless, for the sake of clarity and familiar usage, I shall in what follows phrase my argument in terms of obligations, and ask that the reader take Broome’s reservations as implicit. Nor, for that matter, will I distinguish explicitly between ‘subjective’ and ‘objective’ obligations; given that an agent is acting and required to act on the best information *available to her*, I shall hold, it makes minimal difference whether her felt obligations are ‘actually’ objective or subjective.³² This claim has received support in the literature; while there is a difference, Martin McGuire writes, ‘between what *is* right and what a person *believes* to be right, ... in the case of statements made in the first person [such as the dictates of (AC)], this distinction between objective and subjective vanishes, for the person making the statement.’³³ And ‘[a]lthough thinking of objectivity ... orients us toward “the external world”’, Peter Railton points out, ‘it could hardly demand the banishment of subjectivity. On the contrary, objectivity in representation, belief, or assertion requires the real presence of a representer, believer, or asserter’ who can be ‘more or less objective to the extent that they possess epistemic and semantic capacities that nonaccidentally result in representations that approximate features of the world around them.’³⁴

One remaining worry that has been raised regarding AC is that it may generate *self-referential* obligations or – more precisely – obligations grounded on self-referential beliefs; unless we restrict the set of facts in respect of which I can judge myself to be obligated to perform an action to external facts, there

31 Broome, ‘Normative Requirements,’ p. 415.

32 cf. Caspar Hare, ‘Obligation and Regret When There is No Fact of the Matter About What Would have Happened If You Had Not Done What You Did,’ *Noûs* 45 (2011), pp. 190–206.

33 Martin C. McGuire, ‘On Conscience,’ *The Journal of Philosophy* 60 (1963), pp. 253–263, 261.

34 Peter Railton, ‘Aesthetic Value, Moral Value, and the Ambitions of Naturalism,’ in P. Railton, *Facts, Values, and Norms: Essays towards a Morality of Consequence* (Cambridge: Cambridge University Press, 2003), pp. 85–130, 89–90.

is no reason to think that I cannot sincerely judge that I have an obligation to perform an action in respect of the fact that I sincerely judge that I have that obligation.³⁵ To make this worry specific: AC is meant to generate, for instance, a genuine moral responsibility to give my neighbor some of my firewood, if I can, from my (fallible) beliefs that he cannot cut his own, that he will need some during an imminent ice-storm, and that morally I should therefore do so. But it looks far too powerful if it is allowed to generate similar genuine moral responsibilities just from any belief, *apropos* of nothing in particular, that I do have such a responsibility. In the terms of our formal statement of AC: f^Y – the fact that a believes there to obtain a number of the obligation-generating facts $\{f_1, f_2, f_3 \dots f_n\}$ sufficient to oblige her to perform x – should intuitively not be fulfilled by her belief that f^Y itself obtains, even though f^Y is stipulated to be a member of the set $\{f_1, f_2, f_3 \dots f_n\}$. But how are we to justify this intuitive restriction?

Perhaps, we might concede, there is no reason that I cannot generate an obligation for myself simply by judging in this way; still, as a matter of phenomenological fact we *do not* typically do so. So one response to this worry would be to bite the bullet, and concede that such judgements would be problematic if they ever occurred which, fortunately, they do not. A second possible response would be to observe that such a judgement would be, in any number of other respects, *defective*; and to recall that the obligations generated by AC differ from the felt obligations in that they may be satisfied either by acting upon our judgements *or* by abandoning them, as would seem prudent in such a case.

A more thoroughly satisfactory response, however, would be to explain why self-referential obligation-generating beliefs would be impossible: why the beliefs which generate obligations – though they constitute internal facts – must themselves *concern* external facts. The concept of conscience ‘properly regarded delivers no substantive moral dictates at all,’ as Peter Fuss points out; ‘what it “dictates” is a purely formal relation between what we know or believe and what we are to do, namely, consistency between the former and the latter.’³⁶ That is to say, I take it, that the obligations generated by AC are obligations to act on the substantive moral claims to which we already subscribe concerning the state of the world and our obligations to act in it; it should not be understood as a source of substantive new moral claims in and of itself. So a self-referential judgement of the sort which motivates this worry – an f^Y which was fulfilled only by a ’s belief that f^Y – would be an inherently *empty*

35 I thank an anonymous referee for pressing this objection.

36 Peter Fuss, ‘Conscience,’ *Ethics* 74 (1964), pp. 111–120, 117.

one, never managing to relate actions to any substantive belief or knowledge concerning the external world (any other member of the set $\{f_1, f_2, f_3 \dots f_n\}$), but rather left to chase its own tail in something rather resembling a Bradleyan regress.³⁷

So the enkratic intuition which AC expresses looks both coherent and defensible; although it will not persuade everyone, there are no obviously compelling arguments for its falsehood. There remains the possibility, however, that it contradicts some other position which we wish to hold; most obviously, it appears to be incompatible with the realist picture of moral truth as independent of human evaluative attitudes, since it appeals to just those attitudes to generate the truth of the moral claims with which it is concerned. In section §3, I will show that this appearance is misleading; the principle is a metaethically innocuous one, and can be asserted without contradiction by those committed to any workable version of moral realism.

3 Conscience and Realism

Is the Authority of Conscience really metaethically innocuous? Or must we forswear moral realism if we are to hold it? I will not argue for or against realism here; my concern is simply to show that the assertion of AC does not entail an anti-realist metaethics. AC states that beliefs impose moral obligations; that my believing myself to be morally obliged to perform some action does itself oblige me to perform the action. But moral realism, by any number of formulations, requires that such moral facts be independent of my beliefs about them.³⁸ So on the face of things, one cannot be a realist and hold AC. Is this truly the case?

To answer this, it will be helpful to have a particular conception of realism in mind. There are, of course, numerous more-or-less overlapping versions of realism available; though I will also give explicit consideration to the positions advanced by Russ Shafer-Landau and David Enoch, I will here concentrate on the influential formulation given by Jane Heal, which displays most visibly those features I believe common to any form of realism, of which my defence of AC will take advantage. Heal's account of what she terms *minimal realism* involves two 'generally agreed' components. The first of these she terms the *principle of non-contradiction*; two incompatible moves in any realist language-game cannot both be fully acceptable. In any ethical dispute, then,

37 F. H. Bradley, *Appearance and Reality*, 2nd edn (Oxford: Oxford University Press, 1897), ch. 2.

38 William FitzPatrick, 'Recent Work on Ethical Realism,' *Analysis* 69 (2009), pp. 746–760.

‘the moral realist thinks that there is such a thing as a shared grasp of a standard for correct use of the moral term and that one or other of the disputants will be, even if unwittingly, violating a rule he or she aims to follow.’³⁹

The second component is *minimal doxastic independence*,⁴⁰ the claim that ‘the mere existence and nature of my thought does not constitute the existence of what it is a thought about.’⁴¹ This view, she writes, ‘insists only that if one is a realist one admits that “What do you think?” and “Are things like that?” are different questions.’⁴² The two conditions, moreover, though presented as separate, are linked. Any discourse which fails to satisfy *minimal doxastic independence* must also violate the *principle of non-contradiction*, since two disputants might hold contrary beliefs which, if those beliefs are *ipso facto* correct, will give equal justification to two incompatible statements. The question, then, is whether AC violates these principles; I shall argue that, despite appearances, it does not. The discussion will further demonstrate that stricter formulations of these principles than Heal’s, which AC would violate, lead to unacceptable consequences.

Let us begin with the second component. It is my contention here that while AC may violate some hypothetical condition of *strict* doxastic independence, whereby no causal influence at all is permitted between belief and world, it violates neither Heal’s more permissive condition nor the general spirit of the mind-independence constraint. That is, obligations may, according to AC, be imposed by some of our beliefs about them; but they do not in any substantive sense *depend* upon our having those beliefs. Rather, those beliefs are only one of a number of ways obligations can be imposed, and we may have the relevant obligation while remaining blissfully unaware of our burden.

Recall our definition:

AUTHORITY OF CONSCIENCE (AC): For any individual *a*, the set of facts $\{f_1, f_2, f_3 \dots f_n\}$ which, if they obtained, would be severally sufficient to oblige her to perform some action *x* include f_Ψ , the fact that *a* believes there to obtain a number of $\{f_1, f_2, f_3 \dots f_n\}$ sufficient to oblige her to perform *x*.

39 Jane Heal, *Fact and Meaning: Quine and Wittgenstein on Philosophy of Language* (Oxford: Blackwell, 1989), p. 16.

40 This characterization is not Heal’s own, but due to Alexander Miller, *Philosophy of Language* (London: UCL Press, 2000), p. 169.

41 Heal, *Fact and Meaning*, p. 16. Cf. David Enoch, *Taking Morality Seriously: A Defence of Robust Realism* (Oxford: Oxford University Press, 2011), p. 4; his ‘robust’ moral realism requires that ‘normative truths do not *constitutively* depend on our responses or attitudes or desires.’

42 Heal, *Fact and Meaning*, p. 17.

Obligations may be imposed either by external states or by our judgements regarding those states: specifically our judgements *that* those states impose obligations on us. So those judgements are a sufficient condition for the existence of obligations, but they are not a necessary one; indeed, even their sufficiency is effectively parasitic upon the sufficiency of the external states they fallibly take to obtain. They do not, in Heal's phrase, 'constitute' the existence of the obligations they are about but are merely one potential explanation, as $f\Psi$ is merely one member of the set $\{f_1, f_2, f_3 \dots f_n\}$. Since these beliefs are only one among a number of possible groundings for moral obligations, the obligations are not simply equivalent – or do not simply reduce – to the beliefs; the question 'are things like that?' is always a much wider one than 'what do you think?,' even though what one thinks is a part of how things thereby are.

One way of phrasing this claim, which should make clear both its attractiveness to those of a naturalist tendency, and its application to varieties of realism other than Heal's, is that there is nothing relevantly distinctive about mental or psychological facts. Our personal beliefs do not, according to the realist, affect the 'external' world: but any strain of realism which denied that they are certainly *part* of it would look implausible. Thus, if there are any moral 'truthmakers,' any facts at all which are compatible with realism and can ground moral obligations, then obligation-grounding facts about our mental states are on all fours with those.

A legitimate worry arises at this point: for the realist, our moral judgements must not be 'in corrigible' in Norman Malcolm's sense;⁴³ there must be a *question* whether we have judged erroneously. For AC, this appears problematic; there can be no question of my being in error about the extent of my obligations if judging that I have an obligation to perform x creates, *ipso facto*, an obligation to perform x .

To formulate a response, we need to think about the nature of the beliefs in question. Although I have spoken as though our bare obligations to action are justified or made true by our beliefs that we have them, neither our moral psychology nor the logic of obligation-discourse quite works like that. Rather than thinking of myself as having a bare, unadorned obligation, as the formal statement of AC makes clear, I think of those obligations as being imposed on me *by* some further facts or states of affairs. I do not feel the obligation to rescue a drowning child as *sui generis*, as unrelated to anything else: I judge the obligation to be imposed on me by everything else I know and believe about the situation, to be *made true* by whatever circumstances I take to be salient.

43 Norman Malcolm, 'Direct Perception,' in N. Malcolm, *Knowledge and Certainty: Essays and Lectures* (Englewood Cliffs NJ: Prentice-Hall, 1964), pp. 73–95, 85.

So when I judge that I am obliged to perform some action x , I judge that I am thus obliged *because of* some set of facts $\{f_1, f_2, f_3 \dots f_n\}$. And should I come to think myself mistaken about enough of those facts, I may no longer think myself obliged to perform x (that is, f^{Ψ} will cease to be true), and the conscience-grounded obligation too will lapse. So while it might be that the 'bare' thought that I have an obligation is incorrigible, my moral psychology is never in this way bare; and nor, as we have seen, is a bare obligation formally 'detachable' from its doxastic antecedent. My thinking that I have an obligation may not admit of error insofar as the thought creates an appropriate obligation; but my grounds for thinking that I have an obligation are certainly capable of error, and certainly open to correction.

This leads us to Heal's first condition for realism, the *principle of non-contradiction*. Again, AC may appear *prima facie* to violate it; wherever two agents hold contradictory beliefs about their obligations, contradictory obligations will be generated. We should keep in mind, however, that the dictates of conscience apply only to the bearer herself; my belief that we have an obligation to alleviate world hunger generates an obligation only for me.⁴⁴ I may believe that you, too, have such an obligation, but the truth of *that* belief does not follow from my possessing it. So there is no more a contradiction between your conscience-generated obligations and my contrary ones than there is between A 's externally-generated obligation as a police officer to bring a suspect to justice, and B 's as the suspect's brother to help him escape.⁴⁵

But it is not just between different persons that AC may seem to generate contradictory imperatives of the sort that moral realists are uncomfortable with. Consider again the *fallibility* of conscience: the fact that I may be mistaken about my obligations, or the facts that generate them. Suppose that I arrive, armed, at the scene of a shooting, and see a man holding a weapon. Although I take him to be the gunman, he has in fact disarmed the true culprit moments before my arrival.⁴⁶ Now, the relevant external facts $\{f_1, f_2, f_3 \dots f_n\}$ oblige me not to shoot anybody, since there is no imminent danger. But according to AC, the psychological fact f^{Ψ} – that I (wrongly) *believe* myself obliged by the facts to shoot the man holding the gun in order to prevent further loss of life – generates a contradictory obligation. Does this violate Heal's first condition?

44 Gilbert Ryle, 'Conscience and Moral Convictions,' *Analysis* 7 (1940), pp. 31–39, 31.

45 I assume for purposes of argument here that siblings do have such duties to each other. At any rate, if they do not, I take it that that is not because a relevant *contradiction* exists.

46 The example is inspired by the almost-tragic case of Joseph Zamudio; see William Saletan, 'Friendly Firearms: Gabrielle Giffords and the Peril of Guns; How an Armed Hero Nearly Shot the Wrong Man,' *Slate Magazine* (January 11, 2011) <http://www.slate.com/articles/health_and_science/human_nature/2011/01/friendly_firearms.html>.

I believe not. There are two ways to interpret Heal's reference to 'incompatible moves'; these moves may be just the obligations, or they may be the inferences from particular states of affairs to the obligations they impose. Likewise, beyond Heal's position, we may distinguish between two sorts of moral realist position generally, depending on how they construe this commitment. The first option is too strong; it rules out AC, but takes too much of our routine moral discourse with it. The second option, which I deem to be correct, presents no problem for AC.

Is there a contradiction in saying that *a* has an obligation to perform *p* and a simultaneous obligation not to perform *p*? This gives us, I think, too little information. It would certainly seem problematic to hold that *the same set of considerations* imposed both obligations on *a*. But there would be no obvious contradiction in holding that one set of considerations imposes one obligation on *a*, while a second set imposes another. So there is no difficulty in saying that while $\{f_1, f_2, f_3 \dots f_n\}$ may impose one obligation on me, f_4 imposes a contrary obligation.⁴⁷

This consideration is not limited to conscience-generated cases. Consider the example above, where *A* is obliged by his role as a police officer to arrest a suspect, and *B* is obliged as a brother to help the same suspect abscond. In Bruce Springsteen's song 'Highway Patrolman,' *A* and *B* are the same individual, the eponymous Joe Roberts who has 'always done an honest job, as honest as I could,' yet simultaneously holds that 'when it's your brother, sometimes you look the other way.'⁴⁸ So the apparently contradictory obligations generated by AC are not essentially different from common-or-garden tragic dilemmas; and any condition for a version of moral realism strong enough to rule out such dilemmas looks implausibly radical.⁴⁹

AC can generate a species of tragic dilemma, then, where 'external' facts or states of affairs impose one obligation, and our mistaken beliefs about them impose another. The flip-side of this claim is that when our beliefs are *correct*,

47 Again, this is essentially the result established by Jonathan Dancy in 'The Logical Conscience,' discussed in §2 above. Construing obligations thus dyadically – as imposed *given* some further claims – should be uncontroversial; for a succinct justification, see Lewis, 'Semantic Analyses for Dyadic Deontic Logic.' Lewis' opening sentences are particularly perspicuous: 'It ought not to be that you are robbed. *A fortiori*, it ought not to be that you are robbed and then helped. But you ought to be helped, given that you have been robbed' (p.5). Cf. also the earlier discussion of Broome, 'Normative Requirements.'

48 Bruce Springsteen, 'Highway Patrolman,' *Nebraska* (Columbia, 1982).

49 Peter Railton, 'Pluralism, Dilemma, and the Expression of Moral Conflict,' in P. Railton, *Facts, Values, and Norms: Essays towards a Morality of Consequence* (Cambridge: Cambridge University Press, 2003), pp. 249–290 makes a realist case for taking moral dilemmas

our obligations are in a sense over-determined; we are obliged both by the facts of the matter and by our belief that we are so obliged, where either would be sufficient. Some people find this implausible. Michael Ruse and E.O. Wilson, for instance, in the course of a skeptical argument against moral realism, insist that 'redundancy is the last predicate that an objective morality can possess.'⁵⁰ But redundancy of this stripe is scarcely remarkable; who will dispute, libertarians aside, that I should pay my taxes in full *both* because it is honest to do so and because it is fair?⁵¹

I have focused on Heal's account of realism because, as I say, it shows most perspicuously how AC can be made amenable to the realist; the same case should nevertheless go through for any realist view that is not implausibly strong, since any version of the principle of non-contradiction or its analog, which does not forbid 'tragic dilemmas,' must permit clashing obligations to

seriously. Some theorists do deny that there are tragic conflicts of obligations; see, *e.g.* the debates in Christopher W. Gowans (ed.), *Moral Dilemmas* (Oxford: Oxford University Press, 1987). On such a view, a moral dilemma could only occur if there were all-things-considered obligations both to *p* and to not-*p*; but excluding Burridan's Ass situations, either the role-obligations of sibling or those of police officer must prevail. So moral realists could deny that an agent can have contradictory obligations. But if this view is correct, AC does not in any case generate such contradictions. This is because an AC-generated obligation which conflicts with an independent all-things-considered obligation can only be a *prima facie* or *pro tanto* obligation, because it can at most reflect a proper subset of all the morally relevant considerations; if my belief about how things are diverges from how in fact things are, and so generates a different obligation, it follows that there must be something relevant that I am failing to consider. Whether or not, then, AC can generate all-things-considered obligations (the authoritative nature of conscience, one might argue, is only one of many morally relevant considerations), it cannot generate one that runs contrary to another, independent all-things-considered obligation. I thank two anonymous readers for pointing out this objection, one of those two for supplying the solution, and a third for pressing me for further clarity on this point. *Cf.* also the argument made by John Horty, 'Deontic Modals: Why Abandon the Classical Semantics?', *Pacific Philosophical Quarterly* 95 (2014), pp. 424–460, 429–30.

50 Michael Ruse and E.O. Wilson, 'Moral Philosophy as Applied Science: A Darwinian Approach to the Foundations of Ethics,' *Philosophy* 61 (1986), pp. 173–192, 187; *Cf.* David Papineau, 'Normativity and Judgement,' in D. Papineau, *The Roots of Reason: Philosophical Essays on Rationality, Evolution, and Probability* (Oxford: Oxford University Press, 2003), pp. 6–38, 30.

51 One might worry here that whereas honesty and fairness give us *separate* reasons to comply, in AC there are two obligations where the *same* facts are doing the work twice, once directly and once indirectly, insofar as they feature in the agent's beliefs, so that the redundancy might lead an agent to double-count her reasons for action and decide badly as a result. But recall that to the agent herself, there is no perceptible difference between her objective and subjective reasons, or those reasons she has because of how things are and those she has because of how she thinks things are (McGuire, 'On Conscience,' p. 261); she cannot double-count her reasons because she is unable to see them as distinct in the first place.

be generated by different considerations. However, the reader might well find something fishy about taking one's target version of moral realism from a 1989 book on Quine and Wittgenstein! To dispel any such concerns, I will briefly consider some more recent – and specifically moral – realist positions.

One of the most influential such accounts is due to Russ Shafer-Landau (2003).⁵² He begins by defining moral realism as 'the theory that moral judgements enjoy a special sort of objectivity: such judgements, when true, are so independently of what any human being, anywhere, in any circumstance whatever, thinks of them.'⁵³ At first glance, this might seem to cause problems: if human beings' beliefs can generate moral obligations for those human beings, then those obligations are not independent of what human beings think of them. But note that Shafer-Landau's point here is specifically about moral *judgements*. AC does not deny that our moral judgements can be false, and are not made true by our making them. Rather, it states that *even false judgements* can generate genuine moral obligations. Shafer-Landau's formulation of moral realism would preclude what we have termed 'self-referential' obligations, i.e. obligations generated only by my beliefs that my beliefs – and not any external facts – generated them. Judgements about obligations of this sort would, indeed, come out true just because of what human beings thought about them. But fortunately, we have already disclaimed the propriety of such obligations; this worry simply gives us another reason to.

Let us be sure, though, that we have not simply cherry-picked a conveniently loose formulation of Shafer-Landau's position. Elsewhere in the same text, he fleshes out the realist commitment to 'stance-independence' thus: 'Realists believe that there are moral truths that obtain independently of any preferred perspective, in the sense that *the moral standards that fix the moral facts are not made true by virtue of their ratification from within any given actual or hypothetical perspective*.'⁵⁴ The first thing to note here is the existential nature of the first clause: to whatever extent it licenses stance-dependent moral claims, AC in no way implies that at least some moral claims are not stance-independent. I say more about this point below. The italicized part of the sentence can be read, however, as implying greater generality; the moral standards fixing the moral facts are not *as such* made true by virtue of any process of ratification. Even so, I think, there is nothing problematic for AC here. The moral standards which generate obligations in the normal way are unaffected by AC; the question, then, is whether AC itself – *qua* fact-fixing moral standard – is made

52 Russ Shafer-Landau, *Moral Realism: A Defence* (Oxford: Oxford University Press, 2003).

53 Shafer-Landau, *Moral Realism*, p. 2.

54 Shafer-Landau, *Moral Realism*, p. 15. Emphasis in original.

true by virtue of its ratification within some actual or hypothetical perspective. But note that an agent with a mere AC-obligation, as opposed to an obligation established by external facts in the normal way, *will not believe* that this is the sort of obligation they possess, unless self-referential obligations are permitted; such agents are not in a position to ratify mere AC-obligations as such, only to falsely regard them as ratified stance-independently by the state of the world. So whatever validity AC possesses as a moral standard cannot depend on anybody's ratification of it.

There is further evidence that Shafer-Landau might be implicitly amenable to, if not quite openly supportive of, AC. Two positions he adopts in the course of outlining his version of realism are ethical pluralism⁵⁵ and moral reliabilism.⁵⁶ The first of these involves the thought 'that moral properties are multiply realizable,'⁵⁷ that 'a plurality of descriptive properties [may] realize a given moral one,'⁵⁸ so that good moral considerations may be adduced on both sides of a debate or dilemma, and there may be no 'single best standpoint from which to discover moral truth,'⁵⁹ thereby giving rise to genuine moral dilemmas. As a realist, too, he is committed to fallibilism: 'even ideal moralists may, on occasion, fail to get things right.'⁶⁰ So Shafer-Landau's reliabilism consists in the view that moral beliefs 'can be epistemically warranted, provided they emerge from a reliable belief-forming process,'⁶¹ but that possessing such warrant 'doesn't entail truth. When agents applying reliable processes emerge with conflicting beliefs, each of the beliefs is presumptively justified.'⁶² And agents may not be blameworthy for their possession of false beliefs even where these are formed by an unreliable process⁶³ if they are not in a position to distinguish reliable from unreliable processes.⁶⁴ All this is not yet quite to endorse AC and say that agents possess *positive* moral obligations to act on any false moral beliefs that they are justified in holding. But if not, it is only a little way short, and none of the missing elements are obviously inimical to Shafer-Landau's realism.

55 Shafer-Landau, *Moral Realism*, p. 7, p. 92.

56 Shafer-Landau, *Moral Realism*, p. 272ff.

57 Shafer-Landau, *Moral Realism*, p. 92.

58 Shafer-Landau, *Moral Realism*, p. 92.

59 Shafer-Landau, *Moral Realism*, p. 224.

60 Shafer-Landau, *Moral Realism*, p. 223.

61 Shafer-Landau, *Moral Realism*, p. 272.

62 Shafer-Landau, *Moral Realism*, p. 276.

63 Shafer-Landau, *Moral Realism*, p. 286–7.

64 Shafer-Landau, *Moral Realism*, p. 274.

Another variant of moral realism much-discussed in recent years has been the ‘Robust Realism’ developed by David Enoch,⁶⁵ which he summarizes as ‘the thesis that there are objective irreducibly normative truths’; ‘it states that some normative judgments are objectively non-vacuously true,’ and that ‘normative truths do not constitutively depend on our responses or attitudes or desires.’⁶⁶ Since this is intended as ‘an existential rather than a universal thesis,’⁶⁷ AC’s similarly existential claim that *some* genuine moral obligations are generated just by our beliefs about such obligations poses no obvious challenge to this characterization; AC holds, moreover, that the majority of truths about moral obligations are not thus, in Enoch’s phrase, ‘response-dependent,’ and hence that such response-dependence is not *constitutive* of moral truth as such. But let us look more closely at the details of Enoch’s account.

Enoch provides two chief arguments for moral realism, the second of which – an argument from explanatory and deliberative indispensability – is not relevant here. The first – an argument from the moral implications of metaethical objectivity – aims to block response-dependence, and so might be thought to cause problems for AC. It proceeds from a normative ethical intuition he names IMPARTIALITY, which requires us, in interpersonal conflicts, to openly ‘step back from our mere preferences, or feelings, or attitudes’ and ‘to the extent the conflict is due to those,’ seek ‘an impartial, egalitarian solution.’⁶⁸ Accepting this, if I would prefer that we spend the afternoon watching a movie, and you would rather we played tennis, we should try to resolve the disagreement by a neutral method, like flipping a coin or taking turns to choose; it would be – morally – wrong for me just to insist on getting my own way.

This is not the case in factual disputes, at least in those cases where the truth is important enough to outweigh any concerns about hurt feelings, future relationships, *etc.*; if we are trying to defuse a bomb and disagree whether to cut the blue or red wire, then if I am confident that the blue one should be cut I have reason to act on that basis rather than to seek a neutral compromise, even

65 Enoch, *Taking Morality Seriously*. Usefully, ‘Robust Realism’ is closely related, in Enoch’s estimate, to no less than 17 other realist theories (p. 7), including those of Derek Parfit, *On What Matters* (Oxford: Oxford University Press, 2011), T.M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998), Torbjörn Tännsjö, *Moral Realism* (Savage, NJ: Rowman and Littlefield, 1990), Torbjörn Tännsjö, *From Reasons to Norms: On the Basic Question in Ethics* (Dordrecht: Springer, 2010) and Terence Cuneo, *The Normative Web: An Argument for Moral Realism* (Oxford: Oxford University Press, 2007). If AC can be shown to be compatible with Enoch’s account, then, we can have at least provisional confidence that it is compatible with a broad swath of realist thinking.

66 Enoch, *Taking Morality Seriously*, p. 4.

67 Enoch, *Taking Morality Seriously*, p. 2–3; cf. p. 88.

68 Enoch, *Taking Morality Seriously*, p. 19.

if you are similarly confident. Standing my ground is the right thing to do here. If our disagreement is moral – if you think that it is fine to cause serious and gratuitous pain to dogs, and I think that it is my duty to stop you – it would be morally indefensible for me to agree to settle the matter by coin-flipping: I should again stand my ground. And this, to abridge Enoch's argument only a little, shows that morality cannot simply be a matter of our preferences, feelings, or attitudes, but is instead concerned with matters of fact.

This argument against the response-dependence of morality does not, I think, impugn the particular sort of response-dependence that AC involves. Discussing the bomb-defusing case, Enoch grants that a symmetrical or neutral solution would be appropriate if my reason for cutting the blue wire were simply *that I believe* that doing so will defuse the bomb. 'But my reason for cutting the blue wire,' he insists, 'is not the indexical *that I believe that cutting it will neutralize the bomb*; rather, my reason is *that cutting it will neutralize the bomb* (as I believe).'⁶⁹ So standing my ground is appropriate. Moral beliefs and motivations follow the same pattern, Enoch holds; if someone opposes gay marriage because he believes gay relationships to be unnatural, 'the consideration in light of which he acted was not *that I believe that gay relationships are unnatural*. The consideration in light of which he acted was *that gay relationships are unnatural*' even though, in fact, they are not.⁷⁰ This accords with what AC tells us; it is typically *the way I believe the facts to be* that generates the conscience-derived obligation I feel, and not merely the fact that I believe them so to obligate me. Once more, AC would be problematic for the moral realist if it allowed self-referential obligations; but once more, it does not.

4 Conclusion

For such a commonplace moral intuition, and a central feature of lay moral thinking, it is striking how little the *Authority of Conscience* has been investigated by contemporary moral philosophers. But examining it in detail and with philosophical rigor, it turns out, has benefits. We can now dismiss at least one motivation for the reticence of moral philosophy to discuss conscience-generated obligation; an authoritative (but fallible) faculty of conscience is not bought at the expense of a realist metaethics. Several more *prima facie* objections to the principle, which might have led philosophers to suppose it confused, can now also be disregarded. I have not provided a positive argument

69 Enoch, *Taking Morality Seriously*, p. 22.

70 Enoch, *Taking Morality Seriously*, p. 222.

for the *Authority of Conscience* here; but it is an attractive intuition, with few if any remaining barriers to embracing it.

Acknowledgements

I am extremely grateful to have had the opportunity to present this material to audiences at Trinity College Dublin and the University of Sussex, whose comments were enormously helpful. In addition to all who attended those talks, and to the anonymous referees who have commented on the paper, I wish to thank Aisling Crean, Hannah Galvin, Danny Kelly, Michael Ruse, and Stefan Storrie for stimulating and improving discussions of the material. I am also indebted to Ireland's Department for Social Protection for their support during the period of unemployment when I began this research, without which I would not have been able either to complete it or to persevere and subsequently secure my current position.

Biographical Note

Shane N. Glackin is a Senior Lecturer in Philosophy at the University of Exeter. His areas of interest include normative and meta-ethics, the philosophy of biology and medicine, and the philosophy of language, and he has recently published in the journals *Analysis*, *The Southern Journal of Philosophy*, and *The Philosophical Quarterly*.