V*—MORAL ERROR THEORY

by Hallvard Lillehammer

ABSTRACT   The paper explores the consequences of adopting a moral error theory targeted at the notion of reasonable convergence. I examine the prospects of two ways of combining acceptance of such a theory with continued acceptance of moral judgements in some form. On the first model, moral judgements are accepted as a pragmatically intelligible fiction. On the second model, moral judgements are made relative to a framework of assumptions with no claim to reasonable convergence on their behalf. I argue that the latter model shows greater promise for an error theorist whose commitment to moral thought is initially serious.

I

All reflective persons are familiar with the experience of moral disagreement with apparently reasonable others. Thus, many contemporary liberals confidently reject the restrictive norms of the pro-life movement, although they are often less confident about the explanation of their entitlement do so. The same applies to their opponents. According to the moral error theorist there is a deep explanation of this and related phenomena. Moral judgements in general are false or incoherent (I shall henceforth take this claim as a working definition of 'moral error theory'). Easy to refute, stalking-horse error theories keep appearing in the literature, mostly to be rejected on the way to more cheerful topics. The result is a tendency to underestimate the different forms that a moral error theory could take. Given that there is a form of error theory corresponding to every claim that moral judgements entail, not even the most optimistic moral realist could claim to have refuted every possible form of moral error theory.

Some forms of error theory are implausible. It is implausible to attribute to contemporary moral thought in the secular West a constitutive commitment to a divine lawmaker (this is not to deny

that some contemporary moral convictions depend on such a world view for their justification). Other forms of error theory are harder to dismiss. As Mackie and others have argued, there is a case for the claim that moral thought commits us to an irreducibly non-natural and mind-independent moral reality (Mackie, 1977). There is also a case for the further claim that the idea of such a reality is incoherent.[1] If both arguments are sound, it follows that moral thought is incoherent, and therefore that some relevant form of error theory is true.

There is a significant literature making the case for some form of moral error theory.[2] There is also a growing literature on the consequences of its adoption.[3] Thus, in response to Mackie, Blackburn points out the explanatory gap between Mackie's acceptance of an error theory and his discussion of utilitarianism later in the same book. I want to explore this explanatory gap (Mackie's brief remarks on the subject can be found in Mackie, 1977, Chapter 5). In doing so, I shall assume that a moral error theory is a non-obvious truth discoverable by either *a priori* or *a posteriori* reflection. My question is whether continued engagement in moral thought is a reflectively stable strategy for an error theorist whose initial commitment to moral thought is what I shall call *serious*. To have a serious commitment to moral thought is for one's moral judgements to express genuine convictions (in a cognitively neutral sense of that term), accompanied by a defeasible tendency to corresponding motivation in appropriate circumstances. I thus take serious commitment to exclude both a purely instrumental interest in moral thought, as well as complete indifference to its claims. By 'morality' I shall mean what Mackie calls morality in the 'narrow sense', namely a system of constraints on conduct 'whose central task is to protect the interests of persons other than the agent and which present themselves to the agent as checks on his natural inclinations or spontaneous tendencies to act' (Mackie, 1977, p. 106).

1. Mackie (1977); Garner (1990); Blackburn (1993).
2. Cf. Mackie (1946) and (1977); Newman (1981); Williams (1985); Garner (1990); Schiffer (1990); Joyce (2001).
3. Cf. Williams (1985); Sturgeon (1986) and (1994); Blackburn (1993); Burgess (1998); Lillehammer (1999); Joyce (2001).

II

For strategic reasons I shall not focus on error theories targeted at the notion of an irreducibly non-natural and mind-independent moral reality. First, it is unclear what concrete role the metaphysics of naturalism and mind-independence plays in moral thought. Second, the claim to incoherence entails that the non-existence of a moral reality is a necessary truth, this making it impossible for things to have been otherwise. It is not obvious what difference this necessary falsehood (as opposed to belief in it) could make to anything (cf. Sturgeon, 1986). Third, the incoherence of a mind-independent moral reality is a problem only for those who think moral thought aims to represent worldly facts. Non-cognitivists have historically evaded the error-theoretic challenge by denying this claim (cf. Blackburn, 1993). Yet there are other ways of doubting whether moral thought can be everything it seems, some of which raise questions for cognitivists and non-cognitivists alike.

According to Mackie's 'argument from relativity', the persistent absence of convergence among convictions is consistent with the existence of objective truth only if it derives from epistemological defects, such as faulty reasoning or ignorance of relevant evidence (Mackie's terms are 'speculative inferences' and 'inadequate evidence': Mackie, 1977, pp. 36–38). Mackie claims we have reason to believe that the persistent absence of moral convergence is not due to such epistemological defects. Thus, we have reason to not believe in objective moral truth. Mackie thinks we also have reason to believe that moral thought commits us to the availability of convergence in the absence of epistemological defects. We therefore have reason to believe that moral thought includes an erroneous claim to objectivity. (Mackie's further claim that variations in moral code are better explained by the hypothesis that they reflect naturalistically describable ways of life is not strictly necessary for this, purely negative, part of his argument to go through.)

Mackie is not alone in thinking that some claim to convergence is fundamental to moral thought. Among cognitivists, Smith proposes an analysis of moral reasons in terms of the convergence of the desires of fully rational agents in reflective equilibrium. In the absence of such convergence there are no

moral reasons so defined (Smith, 1994). Among non-cognitivists, Blackburn claims that we are naturally constrained to argue morally 'as though the truth is single' (Blackburn, 1985, Chapter 6). Blackburn's quasi-realist is faced with a quasi-error-theoretic challenge if it turns out that to argue as though the truth is single is wishful thinking. Poised somewhere in the middle, Wiggins puts forward the claim that in favourable circumstances 'There's nothing else to think' as the central cognitive ambition of moral thought; even though what we do in moralizing is to 'colonize' the natural world with our sentiments (Wiggins, 1990/91, pp. 65ff.). In the absence of such 'vindicatory' convergence, Wiggins thinks that moral thought is subject to 'indeterminacy or underdetermination' (*op. cit.*, p. 77). Some idea of convergence as fundamental to the ambitions of moral thought is explicit also in Lear, 1983; Williams, 1985; Wright, 1992; Sturgeon, 1994; Lewis, 1996; Korsgaard, 1996; Jackson, 1998; Scanlon, 1998; Pettit, 1999, and O'Neill, 2000. The obvious exceptions are three. The first is moral relativism (cf. Harman, 1975; Wong, 1984). The second is moral realism of the mind-independent variety (cf. Brink, 1989). The third is a doctrine of mysterious grace. I shall return briefly to the first two in Sections V and VI. I shall ignore the third for lack of space.

There are further motivations for exploring convergence-based error theories. First, the issue of convergence matters in substantial moral thought. We can understand someone's dissatisfaction with convergence produced by offers they can't refuse, their frustration at repeatedly hostile stand-offs, or their despair at the absence of reasonable compromise. Second, and depending on the notion of convergence in play, the availability of convergence could be a contingent matter. This makes it possible to imagine things being otherwise. Two groups of people could be committed to incompatible social norms, where adoption of the norms of the other group is not actually a reasonable option for either party (cf. Harman, 1975; Williams, 1985). Still, there could be a possible world where the circumstances of each group are sufficiently different to make the norms of the other group a reasonable option. Third, the argument from relativity is not obviously unsound. To make the argument more precise, I shall stipulate that the relevant claim to convergence is restricted to what I call *reasonable persons*. By a

reasonable person, I mean someone satisfying two conditions. The first is freedom from epistemological defects, such as ignorance of relevant evidence or faulty reasoning. The second is possession of epistemological virtues, such as consistency, coherence, and reliability. The associated *convergence-claim* is then that the correct application of moral judgements entails convergence on those judgements by reasonable persons. The convergence-claim is compatible with moral relativity resulting from the presence of indexical elements in the content of moral judgements. Tax evasion could be impermissible for us but permissible for them if our circumstances differ enough (e.g., we are well fed and they are starving). The convergence-claim is equally compatible with moral pluralism. Some moral judgements are disjunctive (e.g., 'You should move to Malaga or Madrid'). Rational uncertainty is also permitted, conditional on failure to meet the constraints on reasonableness. It does not follow from the convergence-claim that convergence entails truth. Not all convergence is reasonable. Finally, if P and not-P can both be reasonably rejected, neither P nor not-P is true. (While artificial, vague, contestable, and potentially subject to further indeterminacy, this interpretation of reasonable convergence is close enough to those present in the relevant literature to count as minimally charitable.)

I agree with Mackie that the persistence of moral disagreement is good evidence against the convergence-claim. Some moral disputes are stubbornly intractable even among apparently reasonable people, including the status of human and non-human life, natural habitats, or the claims to territory of different ethnic groups. Nor are such disputes obviously candidates for the pluralist or indexical solutions permitted by the convergence-claim ('Let them abort their foetuses, while we don't. OK?'). I also agree with those philosophers who think the convergence-claim is fundamental to some moral thought (although I do not claim that commitment to it is a necessary, or analytic, condition for basic competence with moral terms). The disposition to persist in arguing, to continue giving reasons, to refuse to accept that convergence is unavailable, to be disappointed when argument runs out, together constitute *prima facie* evidence for a conception of moral thought on which satisfaction of the convergence-claim constitutes part of what it is for moral

judgements to be true (cf. Smith, 1994; Jackson, 1998). I must therefore agree with Mackie that a convergence-based moral error-theory is less than obviously false, and that the consequences of its acceptance are a reasonable topic of inquiry.

III

The moral error theorist has three options. The first is to renounce moral thought in favour of other means to pursue valued social ends. Thus Lukes (quoting Hume): 'Increase "to a sufficient degree the benevolence of men or the bounty of nature" and you can "render justice useless"' (Lukes, 1985, p. 108). History suggests that this option is too optimistic. The second option is to continue with moral thought, albeit revised to avoid the convergence-claim. This response recommends itself to persons who value the social ends promoted by moral thought, but who are so committed to the norms of truth and truthfulness as to make their retention axiomatic, indefeasible, or to be relinquished only at great cost. I shall return to this option in Section V. The third option is to retain the convergence-claim, thereby aiming to secure the benefits of moral thought by the continued acceptance of an acknowledged falsehood. This response recommends itself to persons whose ends are promoted by moral thought, and whose commitment to the norms of truth and truthfulness allows their trade-off against ends better promoted by falsehood. Most people fall into this category about restricted areas of thought. Not only do normal people habitually lie, deceive and mislead in pursuit of their ends. Falsehood, myth, and idealisation are ubiquitously applied without any intention to deceive, as when people tell fairy stories, read fiction, psych themselves up to perform difficult tasks, or model the behaviour of theoretical entities in complex circumstances. The ubiquitous application of acknowledged falsehood in ordinary life might thus be thought to provide a model for moral thinking on error-theoretic terms.

A person could retain an inferential commitment to the convergence-claim in ordinary moral thought without literally endorsing it in moments of disengaged reflection. This idea has prominent precursors, e.g., in Rorty's idea of 'pragmatic irony' (Rorty, 1989), or in Williams's idea of 'proleptic invocation'

(Williams, 1995, pp. 35–45). A Blackburn-style quasi-realist could also argue himself into this position (Blackburn, 1985). More recently, Joyce has labelled the strategy of employing moral discourse in full knowledge of its falsehood *moral fictionalism* (Joyce, 2001; cf. Newman, 1981). I shall adopt this label in what follows, although I shall not address Joyce's proposal in detail (Joyce's error-theory is focused on Kantian-style categorical imperatives, the existence of which arguably entails, but is not entailed by, the convergence-claim). The moral fictionalist proposes that we regard moral thought as a convenient fiction, our engagement with which is pragmatically intelligible in light of the valued social ends which moral thought promotes. Thus, by pretending that the convergence-claim is true, the moral fictionalist could earn greater rhetorical weight for her judgements in cases of persistent disagreement. She could thereby encourage her interlocutor to assign deeper significance to considerations offered in their support than would otherwise be the case. More good things, including convergence, could then be forthcoming. To this extent, the fiction of reasonable convergence resembles the idea of an unrealisable regulative ideal, such as the capitalist dream of a perfectly free market in a world of constant, but variable, political interference (cf. Blackburn, 1993, Essay 1). By analogy, the state of reasonable convergence, although actually unrealisable, could be one to which moral argument can approximate. If so approximating is conducive to valued social ends, the fiction of reasonable convergence can function as a social bulwark for these ends.

In order for the fictionalist strategy to be effective, the fictionalist needs to immerse herself in moral thought to the extent of practically identifying herself with the literally false claims it embodies. To give further precision to this idea, I follow Joyce in adapting to the case of moral thought Walton's model of make-believe, originally formulated to account for fictional engagement with objects of aesthetic appreciation (Walton, 1990; Joyce, 2001). In acts of make-believe, persons entertain propositions they do not actually believe but rather pretend to believe, with whatever psychological or physical accompaniment this requires. To make-believe that some judgement is true involves more than just asserting it. Make-believes requires cognitive and emotional effort, as when someone makes-believe

that he sees gods or angels, that Hamlet is dying, or that distant church-bells can be heard in the score of a symphony. Make-believe also extends beyond the purely introspective, as when someone make-believes while walking down Trinity Street that she is Wittgenstein, that she is gracefully dancing Flamenco in the Seville feria, or that she is a monkey. The scope for cognitive, emotional, and physical engagement associated with acts of make-believe suggests that this notion is rich enough to serve as a fictionalist surrogate for the convergence-claim.

Moral fictionalism offers four advantages to the error theorist. First, it promises to continue serving the valued social ends which moral thought functions to promote. Second, it promises to serve the end of convergence (whether or not actual or reasonable convergence is forthcoming). Third, moral make-believe is *prima facie* analogous to non-controversially sound fictionalist strategies in other areas of thought. Fourth, fictionalism offers a unified approach to moral disagreement. Inevitably, persons are often ignorant about the range of judgements for which reasonable convergence is available, or the extent of reasonable convergence available for a given judgement. Fictionalism provides a rationale for treating all moral judgements in the same way.

IV

The advantages of fictionalism are counterbalanced by five complications, of which two put pressure on the analogy with other fictions, two put pressure on its pragmatic value, and one constitutes an evaluative conflict between fictionalist and pre-fictionalist moral thought.

First, moral thought is not a clearly delimited fiction like a novel, where the parameters of truth and falsity are defined to a high degree across a tightly specified domain. The limits of what moral thought forbids, recommends, or requires are essentially open to debate and negotiation, with even deeply entrenched moral claims open to scrutiny or rejection. The potential for change is obvious if persistent disagreement previously seen as avoidable comes to be regarded as bedrock. Where the coherence and unity of moral thought is challenged, so is the analogy with more conventional fictions.

Second, it is hard to isolate a significant set of non-vacuously defined non-moral ends (such as control, predictive power, or amusement) which can be taken for granted in the form of ends to be served by the moral fiction of reasonable convergence. It is natural to think there are few such ends of any determinacy because in moral thought the status of most ends is reasonably contestable, including the ends of social co-operation, benevolence, and reasonable convergence. While from the definition of 'morality' in Section I it follows that morality should protect the interests of others, this advice is no less reasonably contestable than the notion of 'interest' it employs. Where the ends to be served by moral thought are challenged by moral thinking, so is the analogy with more conventional fictions.

Third, pragmatic considerations favour different fictionalist strategies in different contexts. For example, the efficient implementation of fictionalism requires different degrees of transparency in different circumstances. In some cases, the strategy will succeed if everyone make-believes that reasonable convergence is available. Where multiple ends are shared and the stakes are low, knowledge that reasonable convergence is a fiction could weigh less with participants than the benefits offered by co-operation. For the sake of pretence, I will happily fetch the water, cook the paella, wash the dishes, or even kill the ants (much like children undertake make-believe obligations when playing house). In other cases, the fictionalist strategy will only succeed if someone is literally taken in by the fiction. Where fundamental ends conflict and the stakes are high, there are limits to what agents would reasonably agree to for the sake of make-believe convergence. Moral thought frequently prescribes costly sacrifices, such as the abandonment of basic personal projects or the involuntary termination of life. Such recommendations are ones the individuals concerned might be only too happy to reject as unreasonable. In some such cases, fictionalism could require that relevant persons believe the fiction, thereby regarding refusal as unreasonable (which is not to say that people are inevitably motivated to act reasonably). It follows that fictionalism is only efficient on the condition that in these common scenarios some (possibly most) people have a mistaken understanding of their moral predicament. In such scenarios, efficiency argues for the implementation of a non-transparent moral system along the

lines of the Government House utilitarianism made famous by Sidgwick and Williams, among others. Yet since the Government House model is not efficient in all scenarios, efficiency also argues that it be applied only selectively. Two questions thereby arise, each of which puts pressure on the pragmatic value of the fictionalist strategy. The first question is who should apply it. The second is how, if at all, the Government House model can be effectively applied only selectively.

Fourth, pragmatic considerations do not favour application of the fictionalist strategy in all cases. Fictionalism is only effective where acceptance of the convergence-claim has pragmatic value. Regardless of whether reasonable convergence is available or not, it is an open question how far it is useful either to think it is available or to aim for it. In a scenario where narrow failure to reach agreement between conflicting parties would result in unimaginable disaster, whereas the suspension of discussion would preserve a minimally tolerable *status quo*, the most sensible response could be to leave the negotiating table early. Much as the rhetoric of human rights can obscure the point that it would be a good idea if more countries had credible laws against torture, so the rhetoric of reasonable convergence can obscure the point that it would be a good idea for conflicting parties to cease hostilities.[4] Further pressure is put on the pragmatic value of fictionalism by the fact that efficiency argues for its selective application only, with associated problems of who is to selectively apply it and how. It follows that fictionalism cannot be defended as a universal strategy on the basis of its pragmatic value alone.

The fifth complication arises from an evaluative conflict between the pragmatic values of fictionalism and various truth-related norms to which morally serious persons are likely to be committed. First, any fictionalist inside Government House is committed to live a form of bad faith, internally divided between the 'anguish' of her insight on the one hand, and the 'seriousness' with which she is prepared to falsely prescribe personal sacrifices on the other (the terminology is from Sartre, 1977). This conflict undermines her authenticity, and thereby constitutes an evalua-

4. Cf. Geuss (2001), pp. 146ff, who suggests that talk of human rights is 'convenient, self-reinforcing fiction'.

tive cost. True, exceptions can be made to most norms of thought, including norms of truthfulness. Yet a consistent policy of producing deceitful demands in the service of reasonably contestable ends does not automatically recommend itself to morally serious persons. Second, while fictionalists outside Government House could avoid the problem of bad faith by forgetting that the convergence-claim is a fiction, this is a Quixotic trade-off. While sane commitment to the truth allows for its abandonment in cases of minor cost and major benefit, there are limits to how far commitment to the truth can be suspended while retaining one's grip on reality. The Quixotic fictionalist renounces her capacity to know what she is doing when engaged in moral thought, and thereby a central element of her moral agency. Such tragically deluded individuals will be no more attractive to morally serious persons than the less contrived person who falsely believed the convergence-claim all along. Third, at least some fictionalists could find themselves both outside and inside Government House at different times. Such persons are struck between bad faith and tragic delusion. *Qua* the former, their contribution to morality is sub-optimal when faced with demands requiring false belief for their implementation. *Qua* the latter, they are of no use inside Government House. While there could be a reflectively available route from the former to the latter, it is not obvious that one exists from the latter to the former. In consequence, there can be no comfort in a constant change of masks.

If the prospects for a universal convergence-based fictionalism are questionable, it does not follow that there are no persons for whom the price is worth paying. A loss of authenticity could be a small price to pay for a selfish person whose attitude to morality is purely instrumental, just as playing by the democratic rules is a small price to pay for totalitarians wishing to overthrow the liberal state. Yet the fact that such specimens are possible is cold comfort for morally serious persons. I have not shown that the same applies to other fictionalisms, such as fictionalism about intrinsically motivating and mind-independent facts (Mackie, 1977), transcendental freedom (Newman, 1981), or categorical imperatives (Joyce, 2001). Nevertheless, the analogy with more conventional fictions is at least equally strained for these forms of moral fictionalism. Furthermore, all forms of fictionalism are

faced with internal and comparative questions of their pragmatic value. Finally, the importance assigned to norms of truth and truthfulness by morally serious persons is evidence that other forms of fictionalism also conflict with the values of pre-fictionalist morality.

<div align="center">V</div>

Morally serious persons can recommend valued social ends to themselves and others without making the convergence-claim on their behalf. Yet intelligent moral thought is impossible without the presupposition of some (potentially defeasible) framework of norms and assumptions taken for granted at least for the sake of argument. Various forms of relativism have been proposed to provide local frameworks relative to which moral claims can be interpreted (or re-interpreted) as true or false (cf. Field, 1994, who suggests that prior to being relativised to a set of norms moral discourse is 'factually defective'). Relativist frameworks have ranged from the norms of the speaker to those of the audience, the object of appraisal, or some combination of these (Harman, 1975; Wong, 1984). Historically, relativists may have failed to justify their choice among frameworks (Sturgeon, 1994). Yet moral relativism does not entail that any one moral framework is privileged over others. It is enough for intelligent moral thought to take place that some arbitrary moral frame-work is operative, and that this is understood by those involved. A convergence-based moral error-theorist can thus truthfully make recognisably moral judgements provided these judgements are made relative to some moral framework or other. I shall henceforth call the strategy of so doing *constructive relativism*.

Constructive relativism is motivated by three assumptions. First, there is an indefinite number of possible moral frameworks. Second, no moral framework can be assumed to command reasonable convergence. Third, it cannot be assumed that for all persons there exists some reasonable moral framework. Con-structive relativism consists in relativising moral claims (regard-less of scope) to some moral framework, without making any claim about the reasonableness of that framework. On this strategy, the reasonableness of a moral framework is a question external to moral judgement. True moral judgements make

reasonable normative claims only on persons for whom the adoption of a relevant framework is a reasonable option. While for morally serious persons the requirements of reasonableness may coincide with some moral framework, for others they may not (Foot, 1979; Williams, 1995, pp. 35–45).

VI

Constructive relativism has four advantages. First, it reflects the contestability that undermines the analogy between morality and non-moral fictions. The constructive relativist agrees that the content of morality is potentially indeterminate, as is the content of the ends morality should function to promote.

Second, the constructive relativist is not committed to value moral thought primarily as a pragmatic tool. She is therefore not vulnerable to problems of efficiency generated by Government House scenarios and the like. A constructive realist could reasonably value truthful moral thought intrinsically, while recognizing that not everyone is reasonably bound to do the same.

Third, the constructive relativist can be morally serious. If she is, she can frame her moral thought against the background of ends widely regarded as axiomatic, indefeasible, or to be relinquished only at great cost, such as sentient interests, individual autonomy, or respect for life. She can think about the promotion of these ends against the background of constraints widely regarded as axiomatic, indefeasible, or to be relinquished only at great cost, such as consistency with truth, authenticity, or principles accepted by well-meaning persons aiming to reach agreement on universal moral principles. Furthermore, the constructive relativist can value systematic moral thought intrinsically, and recommend it as an illuminating way to make sense of the social world—all this while speaking truly.

Fourth, constructive relativism is realistically flexible with respect to the scope of convergence. On the one hand, constructive relativism respects the fact that claims to reasonable convergence have a moral and epistemological cost. This respect is consistent with working for convergence when it is forth-coming, likely, of use, or possible to admire while retaining one's

grip on reality. Retaining such a grip includes awareness that not all persons are interested in convergence, or interested in it for the same reasons. This has the advantage of guarding against anyone deceitfully appealing to convergence to serve her own contestable ends. On the other hand, while formulated for a world of irreducibly conflicting ends, constructive relativism does not entail that no ends are subject to reasonable convergence. If some are, these judgements can be generated from moral frameworks reasonable for any person to adopt. In a world where reasonable convergence is available, it need therefore make no practical difference whether people are constructive relativists or unrecon-structed moralists committed to the convergence-claim. (Although I shall not argue it here, constructive realism could conceivably be adopted as a reasonable epistemological strategy even for a moral realist of the mind-independent variety.)

VII

Constructive relativism generates a number of questions, four of which I have space to address here. One question concerns choice. How do persons decide which moral frameworks to adopt? The constructive relativist can answer that for any person some (possibly empty) subset of possible moral frameworks is reasonable to adopt. That there is a substantial and mutually recognized overlap between moral frameworks actually adopted in moral thought, from family meetings to philosophical debate, is a fact no more mysterious than that people can reasonably discuss 'the' colour of the sky. The constructive relativist can agree that some moral frameworks actually adopted are reason-ably adoptable, without making the convergence-claim on their behalf.

Another question concerns normative force. By which moral claims are persons bound? For the constructive relativist, this question is ambiguous. First, an indefinite number of moral claims can be applied to persons, each deriving from a possible moral framework. Given the falsity of the convergence-claim, there is no such thing as the uniquely morally right answer to a question. Second, persons are reasonably bound by claims generated by moral frameworks it is reasonable for them to adopt. The fact that a moral framework is applied to a person

does not entail that she is reasonably bound it. Nor does the fact that a moral framework is not actually applied to a person entail that it is not reasonable for her to comply with it.

A third question concerns purity. What prevents a morally serious error-theorist from adopting a mixed strategy, either making the convergence-claim where there is evidence for its truth, or make-believing that the convergence-claim is true where this promotes valued social ends? In response, the constructive relativist claims to offer a universal strategy for dealing with the fact that persons generally do not know the extent of available convergence, nor whether different kinds of make-believe will efficiently promote valued social ends. Morally serious persons could reasonably regard the consequent loss of efficiency as compensated for by the flexibility and realism offered by the constructive relativist strategy. It does not follow that no reasonable person would prefer a mixed strategy. A constructive relativist is no more committed to the convergence-claim for constructive relativism than she is for moral judgements themselves.

A fourth question concerns the permissiveness of the notion of a moral framework. Why not limit the range of permissible frameworks to actual, reasonable, productive, or interesting ones, thereby building a criterion of relevance into the notion of a moral claim? Two points can be made in response. First, constructive relativism respects the fact that it is possible to think intelligently about merely possible, unreasonable, unproductive, and uninteresting moral frameworks. Second, the constructive relativist does not deny that some moral claims are more relevant than others. What she denies is that less relevant moral claims are less than moral claims. Which moral claims a person should be interested in is a question of which moral frameworks it is reasonable her to be interested in (which may not all be frameworks it is reasonable for her to adopt). For the constructive relativist, that is not a moral question. Nor is it a question that can truthfully be assumed to have the same answer for everyone.

VIII

The case for a convergence-based moral error theory is a reasonable cause for concern for morally serious persons. Not only is there some less than negligible chance that such a theory is true in some form. The consequences of accepting it also include at least the partial abandonment of the convergence-claim. The discussion of fictionalism and constructive relativism suggests that it is mistaken to expect there to be, hidden somewhere in logical space, an error-theory which avoids all practical doubts about the universalistic ambitions of morality. The morally serious error theorist is therefore stuck in a predicament where her doubts about convergence will remain as long as she remembers what she really believes.[5]

Faculty of Philosophy
University of Cambridge
Sidgwick Avenue
Cambridge CB3 9DA
UK
hallvard.lillehammer@kings.cam.ac.uk

REFERENCES

Blackburn, S., 1985, *Spreading the Word* (Oxford: Oxford University Press).
Blackburn, S., 1993, *Essays in Quasi Realism* (Oxford: Oxford University Press).
Brink, D. O., 1998, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press).
Burgess, J. A., 'Error Theories and Values' *Australasian Journal of Philosophy*, 76: 534–552.
Field, H., 1994, 'Disquotational Truth and Factually Defective Discourse', *Philosophical Review*, 103: 405–452.
Foot, P., 'Morality and Art' in M. Burnyeat & T. Honderich (eds.), *Philosophy As It Is*, pp. 7–28 (Harmondsworth: Penguin).
Garner, R. T., 1990, 'On the Genuine Queerness of Moral Properties and Facts', *Australasian Journal of Philosophy*, 68: 137–146.
Geuss, R., 2001, *History and Illusion in Politics* (Cambridge: Cambridge University Press).
Harman, G., 1975, 'Moral Relativism Defended', *Philosophical Review*, 85: 3–22.
Jackson, F., 1998, *From Metaphysics to Ethics* (Oxford: Oxford University Press).
Joyce, R., 2001, *The Myth of Morality* (Cambridge: Cambridge University Press).
Korsgaard, C., 1996, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press).

Lear, J., 1983, 'Ethics, Mathematics, and Relativism', *Mind*, 92: 38–60.

Lewis, D., 1996, 'Desire as Belief II', *Mind*, 105: 303–314.

Lillehammer, H., 1999, 'Normative Antirealism', *Southern Journal of Philosophy*, 37: 201–225.

Lukes, S., 1985, 'Taking Morality Seriously', in T. Honderich (ed.), *Morality and Objectivity*, pp. 98–109 (London: Routledge).

Mackie, J. L., 1946, 'The Refutation of Morals', *Australasian Journal of Psychology and Philosophy*, 24: 77–90.

Mackie, J. L., 1977, *Ethics—Inventing Right and Wrong* (London: Penguin).

Newman, J., 1981, 'The Fictionalist Analysis of Some Moral Concepts', *Metaphilosophy*, 12: 47–56.

O'Neill, O., 2000, *Bounds of Justice* (Cambridge: Cambridge University Press).

Pettit, P., 1999, 'A Theory of Normal and Ideal Conditions', *Philosophical Studies*, 96: 21–44.

Rorty, R., 1989, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press).

Sartre, J. P., 1977, *Being and Nothingness* (London: Routledge).

Scanlon, T. M., 1998, *What We Owe to Each Other* (Cambridge: Harvard University Press).

Schiffer, S., 1990, 'Meaning and Value', *The Journal of Philosophy*, 87: 602–614.

Smith, M., 1994, *The Moral Problem* (Oxford: Basil Blackwell).

Sturgeon, N., 1986, 'What Difference Does it Make Whether Moral Realism is True?', *Southern Journal of Philosophy* 24: (Supplementary Volume) 115–41.

Sturgeon, N., 1994, 'Moral Disagreement and Moral Relativism', in E. F. Paul, F. D. Miller, and J. Paul (eds.), *Cultural Pluralism and Moral Knowledge*, pp. 80–115 (Cambridge: Cambridge University Press).

Walton, K., 1990, *Mimesis as Make-Believe* (Cambridge: Harvard University Press).

Wiggins, D., 1990/91, 'Moral Cognitivism, Moral Relativism, and Motivating Moral Beliefs', *Proceedings of the Aristotelian Society*, 91: 61–86.

Williams, B., 1985, *Ethics and the Limits of Philosophy* (London: Fontana).

Williams, B., 1995, *Making Sense of Humanity* (Cambridge: Cambridge University Press).

Wong, D., 1984, *Moral Relativity* (Berkeley: University of California Press).

Wright, C., 1992, *Truth and Objectivity* (Cambridge: Harvard University Press).