MODEST QUASI-REALISM AND THE PROBLEM OF DEEP MORAL ERROR

TERRY HORGAN & MARK TIMMONS

Simon Blackburn's quasi-realist program seeks to accommodate the deeply embedded assumptions of common sense morality, one of which is the thought that one's current moral conviction on some topic might be mistaken. One may think: "I strongly believe that X is a good person, but I might be wrong" Blackburn has observed that this kind of moral-fallibility context is "the hardest context of all for an anti-realist to understand" (1973, p. 127). Of course, some errors in moral belief are due to mistaken (non-moral) factual beliefs, which present no particular challenge to the moral quasi-realist. One might have false (non-moral) beliefs about X's character, for instance. More worrisome is the thought that one may be mistaken at the level of one's current moral standards. Call mistakes of this sort "deep moral error." The challenge of making sense of first-person affirmations of the possibility of deep moral error represents a hard, perhaps the hardest, challenge to the quasi-realist.

In various writings including "Moral Realism," "Truth, Realism, and the Regulation of Theory," "Rule Following and Moral Realism," *Spreading the Word,* Ch. 6, *Ruling Passions*, appendix, and more recently, "Truth and A Priori Possibility: Egan's Charge against Quasi Realism," Blackburn has addressed the issue of deep moral error in ways that we ourselves (as irrealist moral expressivists) find congenial. Roughly, in thinking that one's current moral conviction might be mistaken, one envisions (as an open epistemic possibility) undergoing a change in moral sensibility that one would regard as a transition to an improved (and not merely a different) moral sensibility, and one expresses one's current moral approval of adopting such an altered sensibility under such hypothetical circumstances. Our principal goal here is to offer a more detailed elaboration of this basic expressivist treatment of moral error than has hitherto been offered by Blackburn or by other expressivists (ourselves included). This will advance the cause of quasi-realism and that of moral expressivism generally.

In subsequent sections of the paper we will briefly address two objections that might be raised against the quasi-realist program, objections that take a specific form vis-à-vis our own

proposed expressivist treatment of the problem of deep moral error. The first objection is that the proposal unavoidably renders moral agents unpardonably smug, because the position entails that one can obtain *a priori* knowledge of the following first-person claim: for any moral opinion M that I now possesses, if I would not give up M under any potential change in moral sensibility of the kind featured in the proposed expressivist account of moral error, then M is correct. We will argue in reply that such a claim of *a priori* first-person knowledge, construed expressivistically, is not objectionably smug at all; rather, it only reflects a potential asymmetrical privileging of one's own moral opinion, vis-à-vis the potential moral opinion of someone else with whom one might find oneself in disagreement about a given moral issue. We argue that such privileging is or can be unobjectionable not only from the perspective of expressivism but also from the perspective of moral realism.

The second objection is that certain moral claims that could potentially be made—and made correctly—within ordinary moral discourse cannot be construed within the proposed version of expressivist quasi-realism in a way that (i) plausibly captures the pre-theoretic content of such a claim, and (ii) allows for the possibility that the claim is correct. These are claims of the form "I might be wrong in my moral opinion M even if I would not give up M under any potential change in moral sensibility of the kind featured in the proposed expressivist account of moral error." We will argue in reply that although it is indeed true that such claims cannot be regarded as a legitimate part of ordinary moral thought or discourse—because that thought/discourse does not embody morally realist ontological commitments—nevertheless the claims make good sense, as intended, when they are regarded instead as metaphysical assertions within *metaethical* discourse. (They make good sense as metaethical claims even though, according to the irrealist metaphysics of expressivist quasi-realism, they are false.)

An important corollary of this reply to the second objection is that the quasi-realist approach to ordinary moral thought/discourse should not be extended to metaethical claims about the metaphysics of morals. By way of elaboration of this point, we will sketch how we think an expressivist quasi-realist can, and should, avoid such quietistic over-reaching vis-à-vis metaphysical debates within metaethics. Our anti-quietist stance represents what we call modest quasi-realism that is to be contrasted with a stronger version that does take a quasi-realist stance toward metaethical claims as well as first order moral claims. However, for the time being, our

exclusive focus in the following five sections will be on expressivist quasi-realism as a position regarding ordinary moral thought and discourse.

Here, then, is a brief preview of how we plan to proceed. We begin by setting forth a series of desiderata that we think should guide the expressivist quasi-realist's endeavor to make sense of self-ascriptions of moral fallibility concerning one's current moral commitments. Since an expressivist who wants to make sense of such judgments of fallibility is restricted by her irrealist moral metaphysics in explaining moral error and, in particular, deep moral error, it is also important to make clear the various resources available to the expressivist in developing an account of the sorts of judgments of moral fallibility which, as Blackburn says, are the hardest for an irrealist to understand. These methodological reflections will then set the stage for our proposal in sections 2 and 3 in which we make use of the resources to explain how our proposal satisfies the various desiderata. Having put forth and made a presumptive case for our view, we turn in sections 4 and 5 to consider the lately mentioned worries about smugness as well as a challenge from the moral realist who may complain that our view cannot accommodate certain judgments of moral error that the realist thinks can be sensibly affirmed within ordinary moral discourse. In section 6, we complete our account of moral fallibility by explaining how the resources available to our brand of modest quasi-realism provide the basis for making good sense of metaethical debates between moral realists and their irrealist opponents, thereby avoiding what we and many others take to be an unacceptable quietism about metaethical disputes that might otherwise seem to be a looming worry for the quasi-realist.[1]

1. Methodological Considerations

A. Desiderata

---

[1] A referee has asked, "To what extent is a paper like this aiming to quell doubts among the quasi-realist faithful, and to what extent is it about making converts?" We certainly aim to quell doubts among the *erstwhile* quasi-realist faithful—those who are otherwise inclined toward quasi-realism, pending an adequate expressivist treatment of the problem of deep moral error. As for making converts, we regard that as a matter of *large-scale* comparative cost-benefit evaluation of quasi-realism vis-à-vis its competitors. Although any quasi-realist treatment of the problem of deep moral error—including ours—might inevitably have to be too baroque to be highly plausible all by itself (independently of wider considerations favoring quasi-realism), we do think that our proposal in this paper has the potential to tip the scales—i.e., to give quasi-realism the dialectical tools it needs in order to be fare better, in terms of overall comparative theoretical costs and benefits, than do its meta-ethical competitors.

We propose the following eight desiderata (set forth as a set of directives) that we think ought to guide the expressivist quasi-realist in developing an account of deep moral error for first order moral thought and discourse. The first three give general directions for developing the kind of account of moral fallibility that is friendly to both expressivism and quasi-realism. The next three specify the range of cases involving fallibility the expressivist should aim to accommodate. The last two have to do with properly addressing the issues about smugness and metaethical quietism mentioned in the introduction.

D1.   Give an account of the state of mind expressed by utterances like "Abortion is morally permissible, but I might be wrong." It won't suffice to make vague and general remarks about understanding the whole practice, or form of life, in which such remarks occur; one needs to be told about the specific state of mind expressed. (It is true enough that expressivists should repudiate any demand to give an "analysis" of such a mental state that construes it as a belief with nondeflationary *truth conditions*; but there remains the task of explaining the nature of moral opinions, as states of mind that *differ from* beliefs with nondeflationary truth conditions.[2]) It also won't suffice to be told, for instance, that such remarks express, say, an openness to listen to, and weigh, counterarguments on the matter at hand; for, pre-theoretically and common-sensically, one means more than that in saying/judging that one might be wrong about a given moral opinion one holds.

D2.   Give an account of the pertinent state of mind that assiduously avoids error theory concerning it, i.e., assiduously avoids construing it as embodying some built-in mistake—e.g., as embodying a mistaken ontological commitment to moral properties and facts of the kind that the moral realist believes in and the moral irrealist does not. Expressivist quasi-realism denies that ordinary moral thought/discourse embodies erroneous ontological commitments. More generally, this position denies that such thought/discourse embodies any kind of systematic error.

---

[2] In section 6 below, we discuss various different potential uses of the truth predicate within ordinary moral thought/discourse and within metaethical thought/discourse. On some contextually appropriate uses within metaethics, we maintain, an irrealist about the metaphysics of morals should say that moral assertions are *neither true nor false*. Within ordinary, first-order, moral thought/discourse, however, normally the contextually appropriate use of the truth predicate is disquotational, and accords with Tarski's schema (T). This deflationary construal of the truth-ascriptions that occurs within ordinary moral thought/discourse is a familiar component of quasi-realism.

D3.   Give an account of the pertinent state of mind that firmly eschews treating it as a belief that imposes nondeflationary truth conditions on the world. Although the overall state of mind might *include* such a belief—say, a belief that it is possible that there is some form of cognitive evolution, meeting certain kinds of constraints, that would take one from one's present moral opinion to a different and incompatible moral opinion—nevertheless the total mental state cannot itself *be* a belief that imposes such nondeflationary truth conditions on the world. Rather, the total state must be one that fits smoothly with metaphysical irrealism about morals and with expressivism. Presumably this means that this mental state must include some kind of attitude other than a belief with nondeflationary truth conditions. The total mental state of mind can't, then, be something as simple as a belief whose content is "I might undergo some change in moral belief or moral sensibility that I would regard as an improvement," or "I might undergo some change in moral belief or moral sensibility that I would regard as an improvement, and this altered belief or sensibility-aspect would not be subject to becoming dislodged by yet further changes in moral belief or moral sensibility that I would regard as improvements." Beliefs of that kind have realist (albeit subjectivist) truth conditions. Expressivism firmly eschews both objectivist and subjectivist versions of moral realism, and should do so for the morality-involving states of mind expressible by "I might be wrong." (Obviously, this desideratum represents a way of satisfying the previous anti-moral error theory desideratum.)

D4.   Make expressivist sense of the distinction between (a) a change in moral sensibility or in moral opinion that is an improvement, and (b) a change in moral sensibility or moral belief that is a change for the worse.

D5.   Make expressivist sense not just of moral improvement involving gradual, piecemeal, alteration of one's moral sensibility, but also of moral improvement involving radical, discontinuous, moral-conversion experiences.

D6.   Make expressivist sense not just of moral improvement involving one's first-order moral opinions and standards, but also of moral improvement involving changes in one's second-order opinions and standards concerning the kinds of changes that could count as improvements in one's first order opinions and standards. First-order standards might include, for instance, simple exceptionless moral principles of the sort *Committing adultery is always wrong*, as well as more complex principles of the sort *It is wrong to break one's promises unless....; Lying is wrong unless....*, where one may or may not have articulated what one considers to be a complete set of

overriding (or undermining) considerations that would fill out the "unless" clause and thus serve to permit (if not require), for example, breaking a promise or telling a lie. Among second-order standards for moral improvement are those articulative of the sorts of "virtues" mentioned by Blackburn: coherence, maturity, imagination, sympathy, fairness, having sufficient relevant information, and culture. (See Blackburn 1998, pp. 310, 313-4, 318.) One may wish to add to this list. For instance, depending on how one conceives of the virtues of maturity and having sufficient information, one may want to add to the list the importance of having a certain range of moral experiences that may serve as "formative" in both shaping and improving one's moral outlook.

D7. Give a credible account of how one can avoid being smug in one's own moral opinion, in circumstances in which one adheres to that opinion despite believing in a deep moral disagreement between oneself and someone else. A special worry seems to arise here because of expressivism's metaphysical irrealism about morals: since (according to expressivism) there's no metaphysical fact of the matter to have a nondeflationary true belief or a nondeflationary false belief about, one can't privilege one's own moral opinion by thinking (say) that this opinion is more likely to be in conformity with how the world is (with the ontologically robust facts of the matter). Why then isn't it just inexcusably smug to privilege one's own moral opinion over the other person's?

D8. Meet all the preceding desiderata 1-7 in a way that does not undermine the intelligibility and significance of the metaethical dispute between moral realists and moral irrealists about the metaphysics of morals. Likewise for the metaethical dispute between expressivist quasi-realists and non-expressivists—where the latter include not only moral realists, but also error theorists like Mackie. If one deploys the deflationary tactics of quasi-realism in too flat-footed a way, without due appreciation for the full variety of intelligible, contextually legitimate, uses of notions like *fact*, *attribute*, *objectivity*, and (nota bene!) *truth*, then one runs the risk of backing oneself into quietism about metaphysical and semantical debates within metaethics. (Beware of becoming beHorwiched!)

B. Resources[3]

   We turn now to the various resources (eight in all) that are available to the expressivist quasi-realist for developing an account of deep moral error which, when utilized in the way we propose, yields an account that satisfies the above desiderata.

   Conditional attitudes. For whatever attitude or attitudes an expressivist account might appeal to, there also can be corresponding *conditional* attitudes. We ourselves have a version that embraces certain kinds of psychological states that we call ought-commitments (which we treat as a species of belief). We also here posit what we call good-commitments. On our view, then, both ought-claims and good-claims are expressions of certain attitudes, which generically can be called ought-attitudes and good-attitudes. So, for example, one can have, say, an ought-attitude toward keeping one's promise to meet his or her spouse for lunch at noon. But one can also have a conditional not-ought attitude toward doing that, given that a colleague walks into one's office at 11:45 to discuss a crisis situation with a student.

   Epistemic possibility. We propose to make use of epistemic possibility, which we take as swinging free of metaphysical possibility. Relative to one's available evidence and information, various claims can express open epistemic possibilities even if they are not all metaphysical possibilities—and even if one knows that they are not all metaphysical possibilities. Concerning those two guys visible at a distance from behind—Sam, the guy on the left, and Dave, the guy on the right—I know that one is the other's father, but I don't know which. It's epistemically possible that Sam is Dave's father and it's epistemically possible that Dave is Sam's father, even though (as I well know) one of these is metaphysically necessary and the other is metaphysically impossible. The kind of modality that figures in "I might be wrong," as applied to some claim, is epistemic possibility—but needn't involve any genuine metaphysical possibility. (The importance of this resource, understood in the manner just mentioned, will become apparent later in the chapter).

   Structural complexity. We propose to construe the pertinent states of mind that express fallibility as structurally complex. Of course, there needs to be an expressivist element, and the overall state of mind shouldn't be a belief with robust, nondeflationary, truth conditions.

---

[3] Some readers might prefer to quickly skim this subsection, then move on to Section 3 (where each of the resources we are about to describe gets deployed in our proposal), and later look again at the present subsection.

However, a belief with robust truth conditions could be a *component* of the overall state of mind, as long as there is also a suitable expressivist component too. (For instance, the overall state of mind could comprise both (i) a conditional ought-attitude (or conditional good-attitude) involving condition C, plus (ii) a belief that it is epistemically possible that condition C obtains.)

Stability. Certain moral attitudes or aspects of one's moral sensibility, either ones one actually possesses or ones one might come to possess through some kind of process of attitude revision that meets certain adequacy conditions, might be such that there isn't any further process meeting such conditions under which one would *lose* those attitudes or that sensibility. Such attitudes or sensibility-aspects, actual or potential, are stable ones.

Idealization. There are various kinds of potential evolution in one's moral beliefs, or in aspects of one's moral sensibility, that might be forms of evolution that could occur only under counterfactual conditions in which certain features of oneself—e.g., certain unrecognized, psychologically deep-seated, prejudices—are altered or absent. Some kinds of idealization along these lines might be legitimate to appeal to in providing an account of moral error. Here it is important to appreciate that appealing to idealization in providing an expressivist-friendly account of moral error does not commit one to non-deflationary truth-makers for thoughts or judgments of the form "Ought-p, but I might be wrong." There are possible uses of idealization in ethical theorizing according to which the sort of thought just mentioned would be made true by the following sort of descriptive state of affairs: *Under such and so ideal conditions for making moral judgments—conditions of full relevant information, etc., I would judge not ought-p*. However, this way of using idealization is to be avoided if one is to remain faithful to expressivism.

Competence/performance: morally neutral variety. One kind of competence/performance distinction that could figure in pertinent forms of idealization involves a morally neutral form of competence in cognitive processing—including in processing by which one comes to change in one's moral sensibility and/or one's moral opinion—that counts as competence independently of one's moral sensibility itself. For instance, one might be guilty of hasty generalization in one's inferences from particular cases to moral conclusions about certain types of behavior. One might be susceptible to special pleading, or again one might be guilty of affirming the consequent in reasoning one's way to some moral conclusion. The list of such morally neutral errors competence/performance errors could be extended.

Competence/performance: morally loaded variety. However a second kind of competence/performance distinction that also could figure in pertinent forms of idealization involves a morally loaded form of competence in cognitive processing—including a process by which one comes to change one's moral sensibility and/or one's moral opinion—whose status as competence depends in part upon the moral standards one possesses at the time the pertinent cognitive processing commences. We call this "process-initial, morally charged, cognitive competence" (PIMC cognitive competence). This sort of competence/performance distinction will play an important role in our account of moral error and will be explained by illustration in the following section.

Moral phenomenology. Finally, another relevant resource that can figure in an expressivist account of moral error concerns aspects of moral phenomenology. For example, there is a phenomenology that accompanies the fact that sometimes one finds oneself still holding a given moral opinion despite knowing that someone else holds an opposing moral opinion, and yet one still regards one's reasons as warranting that opinion despite knowing that the other person does not so regard those reasons. One might call this the phenomenology of firm moral conviction.

With appropriate desiderata and resources now in view that govern the project of making good expressivist quasi-realist sense of moral fallibility, we turn to our positive proposal.

2. The Proposal

One notion that will figure centrally in our account is this: a *factually accurate, cognitively competent, seeming-improvement trajectory* (for short, an *I-trajectory*). This is a process of cognitive transition in which (i) one undergoes some change in one's moral opinion and/or one's moral standards, (ii) any non-moral beliefs that figure in this process are factually accurate, (iii) the process is an exercise of one's morally neutral cognitive competence and/or one's PIMC cognitive competence, and (iv) one experiences this alteration as an improvement (a change for the better). In addition, we will make use of what we will refer to as Improvement-stability (I-stability). A moral opinion, or a given aspect of one's moral sensibility, is I-stable just in case there is no I-trajectory that would take one from that opinion (or sensibility-aspect) to an incompatible opinion (or sensibility-aspect).

Making use of these two related notions, our proposal, for "Ought p, but I might be wrong" is that this locution expresses a complex state of mind also expressible as follows:

Moral Error-Possibility:

(1) It is epistemically possible that some I-trajectory would take me from my current ought-p state of mind to a not-ought p state of mind that would be I-stable, and

(2) not-ought p, **given that** some I-trajectory would take me from my current ought-p state of mind to a not-ought p state of mind that would be I-stable.

The same holds, mutatis mutandis, for other kinds of moral judgments and "I might be wrong" claims concerning them.[4]

Before elaborating our proposal, it is important to take notice of the following points. First, in general, the pertinent kind of trajectory needn't involve just altering the single moral attitude; the new overall state of mind could include a variety of changes in attitude, including dramatic changes in one's overall moral sensibility. Second, note well that clause (1) does *not* say that there is some *metaphysically* possible trajectory that meets the specified conditions (that is also an open epistemic possibility for me). There might very well not be any such metaphysically possible trajectory. The key thing is that there being such a trajectory is an *open epistemic possibility*, for me.[5]

---

[4] Here and throughout, we use the locution "I might be wrong" rather than the locution "I could be wrong." To our ears, the word 'might' is slightly more evocative of epistemic possibility than is 'could'—and epistemic possibility is the kind of modality at issue. This may be a matter of regional idiolect: we two were colleagues for many years at the University of Memphis, and a common locution in the U. S. mid-south is the "might could" construction—as in "I might could reschedule my appointment for next week." Roughly, this means that it is epistemically possible [might] that I would be able to [could] reschedule my appointment. (We realize that this might/could locution might be something that could not be used within all y'all's own regional idiolects.)

[5] To regard something as an open epistemic possibility need not be to regard it as a *live* epistemic possibility—i.e., need not be to take it seriously in one's practical decision making or in one's belief formation. One can confidently believe that one is not an envatted brain—indeed, one can *know* this, and can confidently *claim* to know it—while yet regarding as epistemically open the possibility that one is an envatted brain. (For instance, one might regard this as an open epistemic possibility when engaging in philosophical reflection about epistemological skepticism). Similarly, one can confidently believe that torture is always morally wrong, and one can confidently claim to know this, while yet regarding as epistemically open the possibility that one is mistaken. (For instance, one might regard

Third, this approach meets the first three desiderata from above. It characterizes in a reasonably specific manner the state of mind expressed by locutions like those of the form "Ought p, but I might be wrong," (D1). The proposal avoids attributing any sort of ontological error to first order moral thought and discourse (D2). Furthermore, it construes the pertinent state of mind in an expressivism-friendly way, because although the overall state of mind does include a factual belief as one component—viz., the epistemically modal belief expressed in clause (1)— this fits smoothly with expressivism because the *overall* state of mind also includes a conditional ought-attitude—the one expressed in clause (2)—whose condition is the proposition within the scope of the epistemic-possibility operator in clause (1).[6] Thus D3 is satisfied.

The other above-mentioned desiderata still need addressing which we will take up in the following sections in which we further elaborate our proposal. But before turning to that, there is the following to underscore about the proposal. In saying that the state of mind expressible via "Ought p, but I might be wrong" is also expressible via *Moral Error-Possibility*, we do not mean to imply that the content of *Moral Error-Possibility* is consciously and explicitly present in the mind of someone who says or thinks "Ought p, but I might be wrong." In ordinary moral thought and discourse, one's appreciation of the nature of the pertinent state of mind is apt to be largely implicit, and such implicit appreciation can suffice for this state of mind to play a content-

---

this as an open epistemic possibility when discussing the issue in good faith with an intelligent person who believes that torture is sometimes morally permissible). Furthermore, sometimes one can be in a state of mind expressible as "Ought p, and I *couldn't* be wrong"—e.g., "I ought to refrain from torturing puppies for fun, and I couldn't be wrong about this." The negative modal claim here is that it is *not* an open epistemic possibility that some I-trajectory would take me from my current ought-p state of mind to a not-ought p state of mind that would be I-stable. As a sidebar to the above remarks, let us add that in our view, modal notions like epistemic impossibility, open epistemic possibility, and live epistemic possibility are highly context-sensitive; they are governed by implicit parameters whose settings vary with one's specific intellectual or practical or communicative purposes.

[6] Although we ourselves are inclined to regard the state of mind articulated by clause (1) as a factual belief, we realize that some people might be inclined instead toward the view that epistemic-possibility claims (and the states of mind they articulate) should be construed in an irrealist, expressivist, quasi-realist way. Blackburn himself might prefer to go the latter way, since he favors such a construal of *probability* claims. As far as we can tell, either construal is consistent with our proposal in the text.

appropriate role in one's moral thinking. That is our take on the "paradox of analysis" in general, and on the specific version of it that arises concerning our proposal.

3. The Proposal Elaborated

We now will describe various kinds of I-trajectories that could be epistemically possible for an agent who has a certain moral opinion (say, an opinion of the form "Ought p"), and why and how these various kinds of trajectory allow our proposal to meet desiderata 4-6 that concern the range of cases involving moral fallibility that an expressivist ought to be able to accommodate. By and large, the direction of discussion will be from less radical/dramatic to more radical/dramatic versions of change in one's moral opinion and/or one's moral standards/sensibility.

Perhaps the simplest pertinent kinds of I-trajectories involve nothing more than changes in one's factual beliefs. False factual beliefs one actually has, that actually figure in one's moral opinion, are replaced (in the given I-trajectory) by true beliefs; and/or the given I-trajectory includes true factual beliefs that one does not currently have. For instance, at some particular time, one might have false non-moral beliefs about fetal development, believing perhaps that a one month old fetus is capable of having pain and pleasure experiences which, together with one's (I-stable) commitment to a sentience-based view of moral standing, has led one to mistaken moral judgments about certain cases of abortion.

Then there are I-trajectories in which certain performance errors that one has made, and/or that one is prone to, are not present. For example, maybe one is in fact susceptible to arguments by special pleading, and maybe a given moral opinion one holds has resulted from such susceptibility; the pertinent I-trajectory expunges that performance error, and so it's a trajectory in which one's resulting moral opinion is actually in keeping with one's own current moral standards. Consider, for instance, a donor who believes that her college-age daughter ought to be admitted to her alma mater because of her relationship with the school, even though the daughter does not satisfy minimum entrance requirements. Upon reflection, the donor later realizes that she is guilty of special pleading, given her antecedent commitment to a principle of procedural fairness governing such processes.

We now turn to forms of potential I-trajectory that speak to desiderata 4-6 which address the issue of deep moral error. For example, one's moral standards themselves could evolve, in

the following way. At each stage of the trajectory, most of one's first-order moral standards remain intact, and all of one's higher-order standards concerning morally appropriate forms of change also remain intact; and yet some specific portion of one's moral standards gets altered, in a way that conforms with one's original higher-order standards for appropriate change. One might, say, have a higher-order standard to the effect that certain kinds of novel, enlightening, formative, experience might appropriately change some of one's first-order moral standards; and one might then have such an experience. Suppose, for instance, that one goes from thinking that the kind of universal health care provision just passed into law in the U.S. is a morally wrong mandate to the view that it is morally acceptable, as a result of interacting with people of one's own social status who, owing to economic bad luck, will be left without any insurance coverage during a time when they are in desperate need of care. This is an instance of gradual, piecemeal, moral evolution, ala Neurath's boat. (Blackburn often mentions Neurath's boat, in connection with "I might be wrong.")

Concerning the various cases just described, how does our proposal make sense of genuine improvement vs. regression? Well, a trajectory that leads to a different moral opinion (or a different moral-sensibility aspect) can be a regression if it fails to meet some of the conditions laid down in clause (1) of *Moral Error-Possibility*. For instance, it could fail to be factually accurate, where this failure contributes to the change of moral opinion/sensibility (our abortion example). Or it could involve a performance error—a lapse in one's morally neutral cognitive competence and/or one's PIMC cognitive competence—where this error contributes to the change in moral opinion/sensibility (our donor and health care examples respectively).

Also, even if the trajectory doesn't involve any flaws of the kind just mentioned—and actually counts as an improvement in one's overall mode of moral reasoning—it could still be a case of regression insofar as its specific outcome is concerned. For, that outcome might not be *I-stable*, and perhaps there would be some potential subsequent trajectory that would take one back to one's original moral opinion/sensibility and would also have all the features specified in condition (1) of *Moral Error-Possibility* (including I-stability).

What about radical, discontinuous, changes in one's total moral outlook? Could some of these have all the features specified in condition (1) of *Moral Error-Possibility*? Yes. The key thought is this: even a radical and discontinuous change could count as an exercise of one's I-cognitive competence, because it might well be compatible with both (i) those of one's first-

order moral opinions/standards that remain unchanged after the trajectory, and (ii) those of one's higher-order standards for appropriate change that remain unchanged before and after the trajectory.

There are various potential variations on this theme. For instance, one's first-order standards could discontinuously change a *whole lot*, where this change is largely "anchored" in its conformity to one's prior higher-order standards for appropriate change. Alternatively, one's higher-order standards could discontinuously change a whole lot, where this change is largely anchored in its conformity to one's prior first-order moral standards. Or, one's higher-order and first-order standards could jointly and discontinuously undergo dramatic change, where this change is nonetheless anchored in its conformity with those aspects of the first-order and second-order standards that do *not* change. As an example of a dramatic change that involves standards of both sorts, consider the case of Oskar Schindler. On one natural construal of how Schindler is depicted in the film *Schindler's List*, he undergoes a change in moral sensibility that is both dramatic and also discontinuous (in the sense that he does not reason his way to a new set of moral concerns and associated standards). Presumably, before his conversion experience, Schindler did not feel any moral obligation to help the Jews being persecuted by Nazi forces; he was merely an opportunist who realized that he would be able to use Jews on production lines in his factory at little or no cost. Suppose (and here we speculate) he also held morally loaded second-order standards according to which allowing one's moral views to be affected by sympathetic concern is a kind of moral weakness to be controlled and perhaps overcome. As this individual is depicted in the film*,* Schindler (played by Liam Neeson) has a powerful experience that dramatically transforms his moral views about his responsibility toward the Jews who were being persecuted. Suppose (again, we speculate) that along with his first-order moral views, some of his second-order standards undergo transformation as well, including those having to do with the significance of sympathetic concern. His transformative experience (as we are depicting it) resonated with some deep-seated features of his overall moral character. For him, there was an I-trajectory that resulted in a quite dramatic and discontinuous change in moral view that was the result of an exercise of his process-initial, morally charged, cognitive competence (PIMC cognitive competence) that he experienced as an improvement in his moral views. Having spent much of his fortune "purchasing" the lives of many Jews, at the end of the film Schindler feels deep remorse about the fact that he could have saved more Jews given his resources.

Do these variations on the theme of "improvement via radical and discontinuous change" conform to the image of Neurath's boat? Well, yes and no. No, because the pertinent kinds of trajectory are not ones in which almost all of the boat remains intact at any single stage, with only one or a few planks getting replaced. Yes, though, because these are trajectories in which one stays afloat on *some part* of the boat (maybe a fairly small part) while rapidly replacing all the rest. Cases in which radical and discontinuous change may count as improvement according to *Moral Error Possibility* are to be contrasted with cases such in which change results from, say, "knocks on the head or processes of conversion [that] might get a subject anywhere. But not by processes that he would regard, antecedently, as about to improve his position" (Blackburn 2009, p. 202).

In our view, then, desiderata 4-6 from section 2, concerning the range of types of case of moral fallibility that an expressivist should be expected to accommodate are adequately met by our proposed account of states of mind like "Ought p, but I might be wrong," as lately elaborated.

4. The Smugness Objection

We now turn to desideratum 7 from section 2. It is sometimes alleged that an expressivist ends up committed to a metaethical position that underwrites an inexcusable form of smugness concerning one's own moral opinions and/or moral sensibility, as compared with conflicting opinions or conflicting moral sensibilities of other persons, actual or potential. This kind of objection is especially apt to surface in connection with an expressivist attempt to make sense of judgments of the form "I might be wrong in my moral opinion M." The thought is basically this: The expressivist account of "I might be wrong" ends up immunizing one's own moral perspective in a certain respect, without doing the same for the moral perspective of any actual or potential person with whom one might have a moral disagreement.

In relation to our own proposal, the objection could be put as follows. On one hand, under our recommended treatment of "I might be wrong in my moral opinion M," the following thought/assertion ends up being one that I know *a priori* to be correct:

> *A Priori Truth*: For any actual or potential moral opinion M of mine such that no I-trajectory would take me from M to a state of mind M* such that (i) M* is incompatible with M, and (ii) M* is I-stable, M is correct.

15

On the other hand, I have no such *a priori* guarantee concerning the opinions of any other person, actual or potential, insofar as that person's opinions differ from my own. This asymmetry, the objection says, is just unpardonably smug. I have no legitimate basis, the objection says, for asymmetrically privileging myself that way, vis-à-vis actual and potential others. (For development of this line of criticism, see Egan 2007.)

In response, we acknowledge this asymmetry, but we deny that there is really any objectionable smugness here. In support of that denial, we have several points to make.

First, nowadays there is active debate in epistemology about the so-called problem of rational disagreement. Is it rationally legitimate to privilege one's own belief that P over someone else's belief that not P, in circumstances where one regards the other person as an epistemic peer and one believes that both oneself and the other have reflectively considered the same pertinent evidence (including any considerations that have been set forth either by the other person or by oneself)? Some say no, despite the awkward fact that this conclusion seems to mandate that one suspend judgment on a vast range of issues about which one actually holds specific beliefs—e.g., numerous political issues and numerous philosophical issues. Others say yes: if one has reflectively considered the evidence that seems pertinent (including the considerations that one's peer has set forth), and one finds oneself still believing that P and still regarding P as well warranted by the overall pertinent evidence, then one is rationally justified in retaining one's belief that P. And one is justified in concluding that, as regards P anyway, one's peer has an epistemic sensibility that is somehow skewed.

Consider for a moment one of the most well-known reflections on this sort of peer disagreement in the realm of morals and politics, namely, Putnam's discussion of his disagreement with Robert Nozick over the morality of public education. In responding to such disagreement, Putnam finds himself maintaining his socially liberal position despite Nozick's reasoned opposition. Putnam's phenomenology is one of firm moral conviction--one of the resources we set forth in section 2. Associated with Putnam's phenomenology is his diagnosis of the source of the disagreement as owing to a defect in Nozick's overall moral sensibility. Here Putnam reports that an aspect of his overall phenomenological response to Nozick is of one of "respectful contempt," which, as Putnam says, "is an ambivalent attitude [but] an honest one: respect for the intellectual virtues in the other; contempt for the intellectual or emotional

weaknesses (according to one's own lights, of course, for one always starts with them)" (1981, 166).

We think Putnam accurately describes a phenomenological response to such cases that we expect readers will recognize from their own experience, and which provides good (though admittedly defeasible) reason to side with those who think that the appropriate response to cases of deep disagreement (under the right conditions of proper reflection) is to retain one's conviction. Of course, addressing the current epistemological debate about peer disagreement is well beyond the scope of this paper. But our point here, based on the relevant phenomenology, is that we do not believe that it is inherently smug to stick to one's belief in the face of deep-seated disagreement with an intelligent interlocutor, as long as one has duly considered all pertinent considerations. This is no less so for moral opinions than for other kinds of opinion. (Moral realists can, and we believe should, side with us on this matter.)

Second, given what we have said about the relevant phenomenology associated with disagreement, there is no obvious reason why this general take on the issue of peer disagreement needs to depend on the antecedent assumption that there is a robust fact of the matter about the issue at hand. Expressivists, being irrealists about the metaphysics of morality, deny that there is a robust fact of the matter in moral disputes. Rather, from the perspective of expressivist quasi-realism, when one forms the judgment that the other person's *epistemic* sensibility is skewed with regard to the matter at hand, one is making a *morally engaged* epistemic appraisal rather than judging that the other person's opinion fails to coincide with robust facts of the matter. (A morally engaged epistemic judgment is a psychological state that is a fusion of moral appraisal and epistemic appraisal. More in section 6 below on morally engaged vs. morally detached judgments.)

Third, one has no *a priori* guarantee, concerning any *specific* moral opinion one has or any *specific* aspect of one's moral sensibility, that it is I-stable. Prototypical smugness, we take it, arises with respect to certain specific moral opinions or sensibility-aspects that one stubbornly insists on, come what may. So *A Priori Truth* involves no prototypical smugness.

Fourth, it is important to bear in mind the extent of idealization, vis-à-vis one's actual current self, that figures in the notion of an I-trajectory. It may well be that some moral opinion M of mine is both *plain-old-stable* and yet wrong—plain-old-stable in the sense that the actual me would never give it up because of certain deep-seated limitations in my actual cognitive

competence. Admittedly, it would be problematically smug (although still not prototypically smug) for me to regard the following as an *a priori* truth:

> *Putative A Priori Truth*: For any actual or potential moral opinion M of mine such that no trajectory would take me-with-my-cognitive-competence-limitations from M to a state of mind M* such that (i) M* is incompatible with M, and (ii) M* is stable for me-with-my-cognitive-competence-limitations, M is correct.

But on our account, it is not an a priori truth, and need not be true at all.[7]

Given these four observations, we ourselves do not see any genuine smugness that attaches to our recommended expressivist construal of "I might be wrong." On this matter, we submit, the ball is in the moral realist's court.

5. The Non-Accommodation Objection

Here is something that a moral realist presumably thinks could be sensibly affirmed within ordinary moral thought and discourse:

> *Ineliminable-Error Possibility***:** Ought p, and yet I might be wrong even if it is **not** the case that some I-trajectory would take me from my current ought-p state of mind to a not-ought p state of mind that would be I-stable.

As should be plain, a thought expressible this way is not one that our proposed, expressivism-friendly, account of "I might be wrong" can sympathetically accommodate. Expressivists are moral irrealists, who deny that moral judgments purport to describe metaphysically robust, nondeflationary, in-the-world, facts. Given that claim, and given our proposed account, when *Ineliminable-Error Possibility* is treated as situated within ordinary moral discourse, the content of its second clause is something like this:

> *Silly State of Mind*:
>
> (1) It is epistemically possible that both (i) some I-trajectory would take me from my current ought-p state of mind to a not-ought p state of mind that would be I-stable,

---

[7] Michael Huemer (2005, 41) objected to an earlier attempt in Timmons 1999, 167-70 to provide an expressivist-friendly account of moral error, by arguing that the view in question could not obviously accommodate such thoughts as *I might be mistaken in some of my current moral opinions in ways that I will never discover*. So, notice that distinguishing plain-old-stable from the sort of I-stability associated with our notion of an I-trajectory allows the view we are here proposing to accommodate such thoughts.

and (ii) it is **not** the case that some I-trajectory would take me from my current

ought-p state of mind to a not-ought p state of mind that would be I-stable, and

(2) Not-ought p, **given that** (i) some I-trajectory would take me from my current ought-p

state of mind to a not-ought p state of mind that would be I-stable, and (ii) it is

not the case that some I-trajectory would take me from my current ought-p state

of mind to a not-ought p state of mind that would be I-stable.

But this would be quite a silly state of mind to get into (hence the label): a state of mind in which one affirms the epistemic possibility of a blatant contradiction of the form (A & not-A), and in which one is conditionally not-ought p committed, conditional on this blatant contradiction. Surely nothing like that is what the moral realist thinks is the nature of mental state articulable as *Ineliminable-Error Possibility*.

These latest observations could be turned into an objection to our proposed treatment of "I might be wrong." The objection is this:

In fact, *Ineliminable-Error Possibility* articulates a perfectly sensible state of mind that someone might be in—and, moreover, a perfectly sensible thought that would reside within ordinary, first-order, moral thinking. An adequate construal of thoughts expressible as "I might be wrong about my moral opinion M" should accommodate this fact. Since our own construal fails to do so, it is inadequate. This is the *non-accommodation objection*.

We reply as follows. Although *Ineliminable-Error Possibility* does indeed articulate a perfectly sensible state of mind, that state of mind does not fall within the purview of ordinary, first-order, moral thought and discourse. (Construed in the first-order way, *Ineliminable-Error Possibility* does indeed articulate the Silly State of Mind.) Rather, although the first conjunct—"Ought p"—articulates a first-order moral opinion, the second conjunct, when construed as intended, articulates a *metaethical and metaphysical* opinion. After all, the second conjunct is intended to carry ontological commitment to moral properties and facts—putative properties and facts that the moral realist believes in, and that we expressivists do not believe in.

Now, if moral realists were right in their claim that ordinary moral thoughts/utterances carry moral-realist ontological commitments, then they would be right to claim that *Ineliminable-Error Possibility* articulates a sensible thought within ordinary moral thought/discourse. But we ourselves, and expressivists generally, deny that ordinary moral

thoughts/utterances carry such ontological commitments. So it is entirely appropriate, we say, for expressivists to reply this way to the objection. Although the second clause of *Ineliminable-Error Possibility*, as intended, articulates a perfectly sensible and intelligible thought, that thought is a metaphysical claim that really belongs within metaethics, rather than a first-order moral belief.

Of course, one could have a mistaken belief *about* this thought: one could mistakenly believe that it falls within ordinary moral thinking, because one mistakenly believes that ordinary moral thoughts embody moral-realist ontological commitments. Moral realists have this mistaken belief. But in saying this, we are not embracing an error theory about first-order moral thoughts/assertions, and thus we are not violating desideratum 3 from Section 1. The error is a metaethical one, concerning the nature of those first-order thoughts/assertions,

Moreover, we ourselves, as moral irrealists, would claim that the second clause of *Ineliminable-Error Possibility*, construed as intended, is false—and false in a dramatic way. For, the claim carries a false presupposition—viz., that the right ontology includes moral properties and moral facts. Nonetheless, the metaethical claim embodied in this clause is not silly or irrational, because moral realism is not a silly or irrational view at all. But we irrealists say it is false.

Here it is especially important to notice that one can perfectly well be in the first-order state of mind expressible as *Moral Error-Possibility* even if, *in fact*, no I-trajectory would take one from one's current ought-p state of mind to a not-ought p state of mind that would be I-stable. But that's because one regards it as *epistemically possible* that there is such a trajectory, and one is expressing a conditional not-ought attitude toward p that is conditional on that very epistemic possibility. But there is no sensible expressivism-compatible first-order state of mind articulable by *Ineliminable-Error Possibility*, because that would be a state of mind deploying *as a conditional supposition* that there is no I-trajectory that would take one from one's current ought-p state of mind to a not-ought p state of mind that would be I-stable.

Our reply to the non-accommodation objection presupposes the intelligibility of debates within metaethics between moral realists and moral irrealists—debates about the metaphysics of morals, and about the nature of first-order moral thoughts and first-order moral utterances. This brings into focus the one remaining agenda item from section 2—satisfying desideratum 8, namely, explaining why and how an expressivist who embraces quasi-realism about first-order

moral thought/discourse can avoid being driven to the conclusion that the same quasi-realist features are present within metaethical discourse too. Such a conclusion would threaten to undermine the very intelligibility of the debate between moral realists and moral irrealists, thereby leading to anti-metaphysical quietism. We turn now to that final item on the agenda.

6. Avoiding Metaethical Quietism

Quasi-realism as we understand it is, first and foremost, an approach to ordinary moral thought, and to ordinary moral discourse as articulative of ordinary moral thought, However, one can ask whether the quasi-realist approach to ordinary moral/thought discourse should be extended to metaethical thought/discourse too—and particularly to those parts of metaethics that normally are construed as being about the metaphysics of morals. One can also ask whether a quasi-realist about ordinary moral thought/discourse would somehow be unable to *avoid* thus extending the quasi-realist approach into metaethical discourse itself.

In our introduction, we briefly mentioned a distinction between what we call strong and modest quasi-realism. *Strong quasi-realism* is a form of expressivism, and of metaphysical irrealism about morals, that allows one to non-erroneously think or say, within ordinary moral thought/discourse, essentially *anything* that a moral realist believes can be thought or said within such thought/discourse. By contrast, *modest quasi-realism* is a form of expressivism, and metaphysical irrealism about morals, that (a) purports to allow one to non-erroneously think/say essentially anything that can be thought or said within ordinary moral thought and discourse, while giving the pertinent thought/utterance a deflationary, irrealist, expressivist construal, but (b) does not allow one to non-erroneously think/say everything that a moral realist *believes* can be sensibly and correctly thought/said within such thought/discourse.

As we are about to explain, we maintain that expressivists should provide an account of ordinary moral thought and discourse that implements modest quasi-realism, but not one that implements strong quasi-realism. Modest quasi-realism, we say, is quasi-realism enough.

A key idea in our approach to the issue now in focus is that a whole range of notions that figure importantly in much philosophical discussion, including in metaethics, are subject to contextually variable uses. This is a theme we have urged in much of our own work, and is also

urged in pertinent work by others.[8] Among these are *fact, property* (the noun, as in the noun *attribute*), *objectivity*, *proposition*, and *truth*.

There are uses of all these notions that are largely disquotational. In moral thought/discourse, for instance, if one holds the opinion "Abortion is morally permissible" then, under this disquotational usage, one can also hold the opinions "It is a fact that abortion is morally permissible," "Abortion instantiates the property *moral permissibility*," "It is objectively the case that abortion is morally permissible," and "The proposition that abortion is morally permissible is true." Similarly, notions like *belief* and *knowledge* can be deployed in a way that fits with this usage of notions like *propositions* and *truth*. One can hold the opinions "I believe that abortion is morally permissible," and "I know that abortion is morally permissible." The quasi-realist seeks to accommodate all such thoughts/assertions, and to construe them in such a way that they do not carry moral-realist ontological commitments. This is a point that Blackburn has emphasized over the years, and it has become a standard feature in contemporary expressivist views.

Now, the threat of metaethical quietism looms if one thinks, or is inclined to think, that such disquotational uses are the *only* intelligible, contextually appropriate, uses to which the pertinent notions can ever be put. For, how then will one articulate any genuine debates between moral realists and moral irrealists about the metaphysics of morals, or about the semantics of moral notions and moral locutions? But, we urge, such single-minded thinking is far too flat-footed, and fails to appreciate other contextually appropriate uses of these notions. On our recommended approach, adumbrated at some length in our other writings, the pertinent notions all are governed by implicit contextual parameters, and all can be used in other ways than the ways lately mentioned. The lately mentioned uses are the default uses in *ordinary* moral thought/discourse—the proper focus of quasi-realism. But not, we maintain, in much meta-ethical discourse.

One key distinction here is between morally engaged vs. morally detached uses of such notions as belief, proposition, property, fact, and truth. Morally engaged uses operate disquotationally, and thus operate semantically in an expressivistic manner. For example, when one thinks/says "The proposition that abortion is morally permissible is true," one is articulating

---

[8] See for example, Timmons 1999, Horgan and Timmons 2002, 2006a, 2006b sec. 5, and Wright 1992.

a state of mind that is a *fusion* of moral and semantic appraisal. This thought/utterance is to be construed, then, in much the same expressivistic manner as is the thought/utterance "Abortion is morally wrong."

But there is also a *morally detached* usage of such notions that is operative when one deploys the pertinent notions in a way that not does embody one's own first-order moral state of mind. In metaethics, often one does this to make metaphysical claims or non-fused semantic claims. The metaphysical irrealist, using such notions the latter way, can intelligibly say, e.g., "There are no moral properties or moral facts." The moral realist, using notions the same way, can intelligibly say, "Yes they are." This is a genuine dispute, given the contextually appropriate, morally detached, deployment of notions like *property* and *fact*.

Similar points apply to the notion of truth, which has certain uses that connect it platitudinously with other notions in the family (e.g., *fact*, *property*, *proposition*.) The metaphysical irrealist, using the notion of truth in a morally detached way, can intelligibly say, e.g., "Moral assertions are neither true nor false." The moral realist, using the notion of truth the same way, can say "Yes they are." Here, the notion of truth is deployed in such a manner that truth is a matter of how the world is: some possible ways the world might be are precluded, if the claim is (in the pertinent sense) true.

It is important to note that the contextual differences for all these uses involve fine-grained contextual semantic variation (via different implicit settings of contextual parameters), all compatible with coarse-grained commonality of meaning. To say that such a view posits massive *ambiguity*, then, would be far too crude. (See Horgan and Timmons 2002.)

So, we say that expressivists can, and should, embrace and deploy the kinds of conceptual tools and distinctions we've been sketching in this section. Thereby they can, as they should, firmly eschew being metaethical quietists. Quasi-realism rightly applies to *ordinary* moral thought/discourse, but not to metaphysical and semantic debates within *metaethical* thought/discourse. This is how we propose to satisfy desideratum 8 from our list. Modest quasi-realists can make good sense of the metaethical dispute between moral realism and other anti-realist views including our own brand of expressivism.

7. Conclusion

As stated at the outset, we are very sympathetic to Blackburn's quasi-realist project in metaethics. We do not know whether Blackburn would be willing to embrace the specific proposal regarding judgments of moral fallibility—*Moral Error Possibility*—that we have proposed, or whether he would agree with us that our modest version of quasi-realism is superior to its high octane cousin. In any case, our main objective in this chapter had been to advance the cause of expressivist quasi-realism by making what we hope is some progress on the issue of deep moral error, which we agree with Blackburn is among the hardest (if not the hardest) of contexts for expressivist quasi-realists to explain.

REFERENCES

Blackburn, Simon. 1973. "Moral Realism." In John Casey, ed., *Morality and Moral Reasoning* (London, Methuen). Reprinted in Blackburn 1993. Page references are to the 1993 reprint.

_____. 1980. "Truth, Realism, and the Regulation of Knowledge." *Midwest Studies in Philosophy,* vol. V. Reprinted in Blackburn 1993.

_____. 1981. "Rule Following and Moral Realism." In Holtzman, S. and Leich, S., eds., *Wittgenstein: To Follow a Rule* (London: Routledge).

_____. 1974. *Spreading the Word* (Oxford: Oxford University Press).

_____. 1998. *Ruling Passions* (Oxford: Oxford University Press).

_____. 2009. "Truth and A Priori Possibility: Egan's Charge against Quasi Realism." *Australasian Journal of Philosophy* 87: 201-213.

Egan, Andy. 2007. "Quasi-Realism and Fundamental Moral Error." *Australasian Journal of Philosophy* 85: 205-219.

Horgan, Terry and Timmons, Mark. 2002. "Conceptual Relativity and Metaphysical Realism." *Philosophical Issues* 12: 74-96.

_____. 2006a. "Cognitivist Expressivism." In Horgan, T. and Timmons, M., eds. *Metaethics after Moore* (Oxford: Oxford University Press).

_____. 2006b. "Expressivism, Yes! Relativism, No!" In Shafer-Landau, R, ed. *Oxford Studies in Metaethics*, vol. 1: 73-98. (Oxford: Oxford University Press).

Huemer, Michael. 2005. *Ethical Intuitionism*. (New York: Palgrave Macmillan).

Putnam, Hilary. 1981. *Reason, Truth and History* (Cambridge: Cambridge University Press).

Timmons, Mark. 1999. *Morality without Foundations* (Oxford: Oxford University Press).

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