A Revisionist Defense of Metaethical Contextualism

Dylan Vollans

1. Introduction

Manuel Vargas has described and defended a revisionist account of free will that allows theorists to give different descriptive and prescriptive accounts: One theory about the common sense conception of free will, and another theory about how we ought to conceive of it.[1]. If a revisionist theory about a concept is successful, we may simultaneously hold (1) that our naive, common, or folk theory describes a concept that is not realizable, and (2) that the concept is nevertheless realizable, in virtue of the fact that there is more to understanding it than just understanding the associated folk theory. In particular, understanding the roles attributed to the concept may reveal that it is nested within a system of practices that is independently justified, and this may warrant a revision of our concept to fit that role, given the relationships between the concept, the associated role, and the normative indispensability of whatever in fact plays the role.

If revisionism about free will is successful – or perhaps even if it is not – revisionism about other concepts might be. Wherever we find a concept that is shown problematic upon analysis, we might be justified in revising the concept to eschew whatever the analysis tells us is problematic. Here, I consider the concepts that are widely disagreed about by metaethicists, who compete to best explain the concepts' apparent problems – or, in the case of error theorists, to prove the problems and move on. I will focus mostly on the concepts of *moral obligation*, but if another concept is regarded as more central, the discussion could be recast accordingly. I will argue that the folk theory of moral obligation describes a concept that is not realizable, but that considering other aspects of the language used to discuss moral obligation will render plausible and warranted a revisionist alternative to error theory. I examine three revisionist possibilities in order to show what sorts of conceptual revisions will be promising and which will not. Ultimately, I argue that revisionist concerns will direct us to a form of metaethical concextualism, according to which the truth-values

[1]Manuel Vargas, "The Revisionist's Guide to Responsibility" in *Philosophical Studies*, 125:3 (2005): 399-429.

of moral assertions depends on context, and on something like the *moral standards* of the interlocutors within the context. First, I offer a schematization of (at least some forms of) revisionism to prepare for the discussion that follows.

2. Revisionism: The Reexamination of Theoretical Priorities

Let me begin with an example borrowed from Vargas: That of the concept *water*. In the past, it was held that water – the clear, drinkable liquid filling lakes and streams, etc. – was one of four basic indivisible substances in the universe. However, over time, more data was acquired about this liquid, and it was discovered that it was neither basic nor indivisible. At this point, there were three inconsistent beliefs, at least one of which needed to be rejected for the sake of doxastic coherence: (1) That water was a basic indivisible substance, (2) that water was the clear, drinkable liquid filling lakes and streams, and (3) that that liquid was neither basic nor indivisible. As a matter of historical fact, it was (1) that went, in spite of the fact that it was supposed to express a conceptual truth about water. People did not go on to think, in light of (3), that water did not exist, and intuitively, it does not seem that they acquired a brand new concept, "*water**". And this particular belief modification seems warranted: Abandoning (2) would be much more doxastically costly than abandoning (1). People would have to reject all beliefs about ever having interacted with water, as opposed to merely rejecting beliefs about the metaphysical nature of the substance with which they interacted. So this case can easily be taken to provide a precedent for conceptual revision, showing us not only that revision has historically been warranted at least once, but that we have been tacitly accepting the revision all along. And this case is likely just one of many.

I will now generalize what we see in this case. As with the particular above case, revisionism as a theory of proper belief modification becomes relevant when there are conflicting beliefs of the right sorts. These sorts are what I will call *meta-conceptual*, *attributional*, and *world* beliefs, corresponding to beliefs like (1), (2), and (3) above.

So consider a concept C about which there is a folk theory whose characteristics we can come to know in part by studying widely shared intuitions about the necessary and sufficient conditions for C's realization (even if the intuitions are rarely or never explicitly recognized by the folk, but rather presupposed in certain contexts). This aspect of the folk theory can be modeled by a set of propositions that express these widely shared beliefs about the intrinsic features of C. These propositions are conditional in form (e.g. "x has free will iff x has alternative possibilities"), and are regarded as conceptual truisms. (It would probably be too strong to say that the folk theory takes these conceptual truisms to be *analytic*, and it is not necessary to

say so here.) These propositions characterize the first sort of beliefs listed above: The meta-conceptual beliefs. I will call the set of these propositions M.

In addition to meta-conceptual beliefs, there will be attributional truisms about C: These beliefs pertain not to the nature of C itself, but rather to the sorts of environments in which C is realized. For water, one of these is the belief that water is what fills the lakes, for example. Another attributional truism about water is just that the concept *water* is realized. Many of these beliefs concern paradigmatic cases. For free will, one such belief might be that normal humans in normal circumstances have free will. Attributional truisms like these give character to the world around C, and can perhaps help us learn about C by working from the outside: We can look for commonalities among paradigmatic cases that might account for our intuitions about C's being realized, such as when we conclude that water is H2O after seeing that this chemical composition is common to all paradigmatic samples of water. Call the set of propositions describing a concept's attributional truisms A.

M and A could be considered different aspects of the folk theory of C. The next category, *word* beliefs – corresponding to (3) above – is not folk theoretic. These are agreed-upon purported facts that are not at issue in debates about C. These include scientific discoveries, for example.Where C is *free will*, the discovery of universal determinism would fit in here, but only if described in a way that does not employ our concepts of freedom or free will. These facts must be described in neutral terms like those in (3), one's that can be agreed upon by any theorist of the relevant concept. Because these propositions are not at issue, they are given highest priority when theorizing about C: Folk intuitions – no matter how powerful and convincing – that contradict these considerations must be false, since we hold these scientific beliefs and conceptions constant when engaging in philosophical debates that employ them. This is what happened with the concept *water*, when the folk theory had to be changed in light of the beliefs characterized by (3). Call the set of these propositions describing these agreed-upon beliefs about the world W.

The folk descriptions of C contained in M are often taken to tell a privileged story about the actual nature of C. They are, after all, our initial intuitions *about* the concept itself. Even when folk intuitions are employed in philosophical theorizing, it is not always foregrounded that these are, at root, *folk* intuitions, but it is often taken for granted that these *descriptive, meta-conceptual* intuitions (the folk's or the philosophers') about concepts offer the only, or the best, microscope under which the concepts may be analyzed. If an analysis carried out starting with M leads to unacceptable conclusions, in light of A or W, the suspicion is not typically that our conception of C is inaccurate and misleading, but rather that we missed something in M, and that we need to take a more careful look at it – that we misunderstood

our own conception of the concept, not that our conception could have been *wrong.* Revisionism suggests that the correct course of action may indeed be to become skeptical about propositions in M, instead trusting the credence we have in A – propositions that outline the characteristics of our *employment* of C – to illuminate the concept.

So, as is clear, these issues arise when the propositions of M, A, and W cannot all be simultaneously satisfied. W is beyond critique, and the traditional route prioritizes M over A, under the assumption that studying the *actual* nature of concepts amounts to studying our conceptions *of* them. As a result, the traditional route would likely produce an error theory in such instances of conflict, which entails the falsity of A's contents, and then some – that no state of affairs which intuitively realizes C actually does.

The revisionist alternative is to rearrange our theoretic priorities, such that we deal with the contradictions among M, A, and W by looking for items in M – not A – to judge to be false. This move is not always warranted, but revisionism is supposed to offer something new by suggesting, in a distinctively systematic and self-aware way, that it sometimes is. There may be numerous ways of establishing warrant, but the one I will highlight is reflective equilibrium: Once we have established that M – our conception of C – may not actually accurately characterize C, reaching conclusions about C may be a matter of weighing the doxastic significance of elements of M against those of A, and resolving the contradictions in the least doxastically damaging way. Of course, this process will be informed by what we know about W as well. It may not be enough to warrant belief in free will compatibilism that we confidently intuit that some people have free will, but it may be enough once we consider the empirical facts about how belief in free will pervades our communities and structures our systems of praise and blame. That is, we might be more confident about everything in M than we are about anything in A, but relevant facts in W may predict unacceptably drastic belief and habit modifications as we continue to pursue reflective equilibrium after accepting the error theory.

In order to apply this schema to the problems of metaethics that lend credence to moral error theory, we need to first develop an understanding of folk metaethics and isolate its problematic aspects, to second identify and characterize the attributive intuitions of the folk, and to third look for the most satisfying way of resolving contradictions between the folk theory and our knowledge of the world, in order to establish warrant for the revision.This third task – establishing warrant, may be accomplished before we reach conclusions about which *particular* aspects to revise, but the manner in which we establish warrant will inform us about how best to revise. The primary goal, however, is simply to show that we should sooner alter our

conception of the concepts in question than abandon them altogether.

3. Folk Metaethics

The goal of this section will be merely to describe what some relevant commitments of folk metaethics might be. Afterward, I will describe how these commitments put pressure on us to adopt an error theory: No metaethical success theory can accommodate all the folk conceptual commitments, so if we are to honor everything the folk theory claims about the nature of moral concepts, we have to admit that the concept is unrealizable. In the terms of the previous section, since the content of M with regard to *moral obligation* includes inconsistent propositions, all the elements of A must be false.If this is correct, any good success theory will need to be revisionist with respect to the nature of moral concepts, which means that in constructing theories of these concepts, we are free to proceed with the methodological liberties distinctive of revisionism.

First, I need to make an assumption that I hope is plausible. I will be assuming that there is a central moral concept that is in play whenever moral judgments are made. In particular, I will be assuming that this concept is *moral obligation*, but the discussion could be recast for other concepts. The idea is that, according to the folk theory, when we discuss goodness, rightness, or duty, we are discussing what we ought to do – which actions we ought to perform, which attitudes we ought to have, which states of affairs we ought to promote, and so on. my goal is to capture the basic normative aspect of moral discourse under the broad umbrella of *obligation*.

So, with this assumption in place, understanding what folk metaethics has to say about moral obligation can be achieved in part by understanding what folk metaethics has to say about moral judgment: Folk theoretic commitments about moral judgment will entail folk theoretic commitments about moral obligation, given the presumed connection between the two. Joyce argues for some possible folk commitments, which I quote below[2]:

- Moral judgments (as public utterances) are often ways of expressing conative attitudes, such as approval, contempt, or, more generally, subscription to standards; moral judgments nevertheless also express beliefs; i.e., they are assertions.

- Moral judgments pertaining to action [are] deliberative considerations irrespective of the interests/ends of those to whom they are directed; thus they are not pieces of prudential advice.

[2]Richard Joyce, *The Evolution of Morality* (Cambridge, MA: The MIT Press, 2006), 70.

- Moral judgments [are] inescapable; there is no "opting out."
- Moral judgments [...] transcend human conventions.
- Moral judgments centrally govern interpersonal relations; they seem designed to combat rampant individualism in particular.
- Moral judgments imply notions of desert and justice (a system of "punishments and rewards").
- For creatures like us, the emotion of guilt (or "a moral conscience") is an important mechanism for regulating one's moral conduct.

Though perhaps any item on this list could be contested, it is not my place here to discuss the individual merits or demerits of each. I will take it for granted that this list provides a plausible and sufficiently accurate partial characterization of folk metaethics, that these claims are in some capacity among our naive metaethical intuitions. My proposal is, after all, conditional in nature: I hope to sketch an account of what we ought to do with the folk theory if its commitments do turn out to be problematic. A conscientious and full defense of this antecedent claim is far beyond the reach of this paper.

What I have listed above, then, provides some information about folk truisms of moral obligation. Moral obligation – the concept centrally involved in the making of moral judgments – is something that is not a form of prudential obligation, it is inescapable, it is non-conventional, and so on. These are the sorts of commitments that ordinary philosophic introspection might reveal to us as well, and therefore the sorts of commitments we would hope to build into our metaethical views. Certainly not every metaethicist has maintained all these views, but if Joyce is right about the folk theory, such metaethicists have already taken a revisionist move without necessarily being explicit about it.

On one line of thinking, we should not doubt that Joyce is right, since all that is under discussion at this point are our most naive intuitions. Metaethical theories that abandon one or more of these commitments require often quite elaborate arguments in order to maintain plausibility at all: Relativistic theories must explain away apparent categoricity, just as non-cognitivist theories must explain away apparent property-attribution.(Recall Section 2, where I claimed that when our intuitions clash, philosophers tend to try to discredit or explain away some intuitions, rather than accepting them all equally at first, and then systematically choosing ones to revise.) Once again, however, it is not my objective here to carefully describe the folk theory, especially because this task is likely an empirical one. I just want to

point out that Joyce's observations do justice to our most naive intuitions, as we can see in part by the fact that leaving these commitments out of one's metaethics creates *prima facie* problems that require independent arguments in order to be made convincing.

In the next section, I try to motivate the view that any success theory requires departure from the folk theory – that is, that vindicating moral discourse and practice requires metaethical revisionism, at least once we have accepted that the folk theory is approximately as Joyce describes.

4. Metaethical Conceptual Error Theory

First of all, while it is perhaps to some degree debatable – and in fact a somewhat complex issue – it seems fairly plausible that the folk intuitions about morality include the belief that, as Joyce points out in the first bullet point above, moral judgments are *assertions*. Normally, when someone says, for example, "Hitler was evil", they take themselves to be stating some matter of fact. She is trying to express something about *Hitler*, not about her feelings toward Hitler.

Nevertheless, moral judgments do have symptoms of non-assertoric language, such as the regular relationships between our moral judgments and our conative, motivational states: The person who says "Hitler was evil" is also very likely to have negative attitudes toward Hitler, and it is hard (if not impossible) to imagine someone sincerely and competently agreeing with this statement without sharing such attitudes. Observations like this, pertaining to moral language's function in expressing conative attitudes such as approval or disapproval – coupled with the difficulties that attend cognitivism about morality – have pressured many theorists to adopt forms of expressivism. But this would already be doing a little bit of revision: Such theorists seem to be required to say that the folk theory is wrong at least with respect to its assumptions about the assertoric status of moral claims. It is worth noting, though, that this is revision of a different kind: The folk intuitions about, say, whether moral beliefs are cognitive or not do not necessarily carry commitments about the nature of concepts, so this sort of revision is not necessarily *conceptual* revision, but it is revision nonetheless.

We might think that cognitivism can provide a fully non-revisionist alternative. After all, there is some hope that we could accommodate these motivational features of moral judgments within a cognitivist framework. Ordinary assertions are used to express approval or disapproval, or other similar attitudes, quite regularly. When I am trying to decide which computer to buy, my judgment that "This one runs awfully slow" may also express my negative attitudes about the computer, as long

as the interlocutors all share the assumption that a computer that runs slowly is, by my lights, a bad one.

There are, however, other contradictory aspects of the folk theory, such that no single cognitivist theory will be able to satisfy all the folk commitments. To see this, look at the second, third, and fourth bullet points from the preceding section. Together, these bullet points show us that moral judgments purport to provide *deliberative concerns*, that they are supposed to apply to everyone *independent of anyone's particular interests, concerns or motivations*, and that they *transcend convention.* Put in another way, moral judgments purport to describe facts that would provide non-conventional *external* reasons. (As I will use the term, a reason for some agent to ϕ is external – as opposed to internal – iff it obtains independent of particular motivational features (e.g. desires, values, or interests) of that agent.)

For all that has been said so far, it could still be the case that *moral obligation* is a realizable concept. It is only when we consider *reasons* more generally, and the relationships between reasons and obligation, that error theory becomes seriously alluring. Bernard Williams began a continuing controversy about external reasons by arguing that there are in fact *no* such reasons, and his argument seems to rest comfortably with folk intuitions: In order for some consideration to be a reason for someone, it must be relevant to them, given the things *they* care about.[3] When we ask which reason an agents acts for, and by this we mean to ask what the agent took to justify their actions, the reasons we are looking for are things that figure into agents' motivations for action. Moreover, when we ask why *we* should perform some action – that is, when we ask what *reason* we have to perform it – we only take some consideration to be a reason if it is motivationally significant to us. For example, if I find wedding cake to be overly sweet, it would make no sense to tell me of its sweetness in order to try to provide me with reason to it. Similarly, if I lack concern for convention, the conventional status of eating wedding cake at weddings will also fail to give me reason to eat. In order for something to give me a reason it has to be something I actually *care* about. This is just how we think about reasons.

Of course, when we think about things like morality, we are pressured into thinking that *some* reasons are different. Even though a murderer may feel no compunction for his actions, we still want to say that the *wrongness* of his murderous actions provided a reason for him not to perform them. But this intuition cannot be reconciled with our intuitions about reasons more generally. In normal circumstances, such as with wedding cake, not much is at stake in whether someone sees things the same we do: They just have different reasons, and that is okay. However, with

[3]Bernard Williams, "Internal and External Reasons" in Harrison, *Rational Action and Moral Luck*, 101-13.

morality, a lot is riding on whether or not others share our perceptions about what reasons there are. (This ties in with the fifth and sixth points in Joyce's list from the previous section.) I think it is this fact that causes us to maintain these inconsistent intuitions.

According to Joyce, these inconsistent aspects of the folk theory should cause us to adopt an error theory of moral discourse: Attributions of moral properties are, he says, systematically false. This is made clear in what I will refer to as *Joyce's Deduction*[4]:

1. If x ought to ϕ, then x ought to ϕ regardless of whether he cares to, regardless of whether ϕing satisfies any of his desires or furthers his interests.

2. If x morally ought to ϕ, then x has reason for ϕing.

3. Therefore, if x morally ought to ϕ, then x has a reason for ϕing regardless of whether ϕing serves his desires or furthers his interests.

4. But there is no sense to made of such reasons.

5. Therefore, x is never under a moral obligation.

At any rate, Joyce is right that no cognitivist success theory of morality will be able to keep all the relevant intuitions intact: If Williams is right about reasons, we must change how we think about morality, or else we must change how we think about reasons. There is no way to achieve a success theory without such changes. And if we want to avoid an error theory, this assures us that revisionism is the only way to remain cognitivists. (And, as we have seen, even non-cognitivism would constitute a form of revision, though much different in character.)

So, in the following sections, I try to show why revisionist cognitivism is preferable to error theory, and I map out a few revisionist possibilities in order to show what revisionism is likely to end up recommending for metaethics.

5. Motivating Conceptual Revision

If what has been said so far is approximately accurate, the conceptual truisms of folk metaethics force a descriptive error theory upon us. The folk theory contains commitments that are not simultaneously satisfiable, in such a way that the folk concept of moral obligation is incoherent given the folk conception of reasons. Now, the task is to motivate conceptual revision and to show that it may be a good response

[4]Richard Joyce, *The Myth of Morality* (Cambridge, U.K.: Cambridge University Press), 42.

to descriptive error theory. It will not be enough to just point out the practical utility of moral discourse and practice, since this can be salvaged by metaethical fictionalism, as Joyce has suggested[5], or by a revolutionary non-cognitivism. What I hope to show is that we should do more than keep around the verbal relics of abandoned concepts: We should rather look for philosophically sound and warranted ways of reforming our conceptions and adjusting their contours until they give us concepts we *can* talk about without error. It is the good, the right, and our duties that are important – not 'the good', 'the right', or 'our duties'.

First of all, there really are significant differences between morality as fiction and morality as real. When we take morality to be real, we can be both theoretically conscientious and morally judgmental when we witness apparent transgressions against our morals. No amount of prodding with questions like "But is it *really* wrong?" will tempt us to back down. When morality is taken to be fiction, even though things may most often carry on as they do now, theoretic conscientiousness alone sweeps away all moral pretensions.

In a similar way, the experience of watching a horror film can be worsened if one is continually reminded that one is only watching a movie . So the error theorist could reply that just as we get lost in a horror movie enough to maximally enjoy it without ever really losing track of reality, we could get lost in moralizing enough to fully reap the benefits without ever really losing track of reality. But there is something disturbing to me about this answer. Namely, while people view horror films in order to be thrilled and entertained, people carry out moral discourse and debate in order to get nearer to the truth, not just in order to reap instrumental benefits of the process. Being a horror-film fictionalist is fine, since the point of watching a horror film is to be entertained once one has taken their guard down enough to be absorbed in the fiction. Being a moral fictionalist, however, requires one to forget about reality enough to engage in an activity whose ostensible end is the discovery of truth – not entertainment or other forms of instrumental benefit. I don't mean this to be in itself a knock-down argument against metaethical fictionalism, but I do mean to highlight the strangeness of it: It calls for systematic cognitive dissonance and theoretic disingenuity for the sake of preserving practices that are instrumentally beneficial. This is costly, although not in a way that can easily be cached out in terms of the doxastic damage caused by the adoption of fictionalism, but still in such a way that we ought to, *ceteris paribus*, prefer alternatives that do not call for such dissonance.

The primary cost of metaethical error theory is much more general and much more obvious: It calls for the rejection of intuitions that are among our most powerful.

[5] *Ibid.*, 206-31.

To most, it seems unshakably obvious that Hitler was evil, for example. There are likel ymany who would sooner deny the existence of the external world than agree that torture-for-fun is okay or that genocide is permissible. (Idealism is much more popular than amoralism.) The history of theodicy has shown that many are more confident in their moral beliefs than they are in their religious beliefs: If there is a contradiction between the two, it is commonly thought that the moral beliefs stand on firmer ground than the religious beliefs. It's not that God exists and suffering is okay, but that either God does not exist or else his existence is somehow compatible with the wrongness of such suffering. These moral intuitions are extremely powerful, and the universal rejection of them should not be taken lightly. The error theorist seems to have taken them very lightly, since their view – by revisionist lights – commits them to saying that they are more sure about the subtleties of categoricity, obligations, reasons, and rationality than they are about Hitler's being evil. This would be a unique way to be, since most others would likely assent to any success theory that allowed them to be firm in their moral convictions: Error theory yields unacceptable entailments, and even if it is diagnostically correct, revisionism shows us that our prescriptions may, and sometimes should, differ from our diagnoses. We can look more closely at the premises, and if we can resolve the contradiction by eliminating a premise about which we are less confident than we are about Hitler's being evil, we should sooner abandon the premise than accept the conclusion.

To summarize: Given the contradiction in the folk theory that Joyce has highlighted, together with the comparably powerful intuitions that things sometimes *are* good or bad, right or wrong, we have more reason to assume that our intuitions about the nature of moral obligation or reasons are wrong than we do to assume that our moral judgments are systematically false. Even though Joyce's fictionalism would provide a way to leave discourse and practice largely untouched, it would be missing the point: The cost of the view is that it is an error theory, not that it would change our actual behavior in an unacceptable way. The doxastic damage of the theory is immediately obvious, since it calls for us to accept, among competing contradictory beliefs, the one with the lowest credence. Revisionism grants us means to avoid error theory, and since error theory *is* so doxastically damaging, conceptual revision is warranted.

6. Conceptual Revisions: Comparing Alternatives

I have so far tried to show how revisionism would be a promising alternative to error theory if moral obligation (as conceived of by folk intuitions) turns out to be a problematic concept. Now, I will address the question of how to move forward

once we admit that revisionism is viable. This will include examining a few different possibilities for revision and developing standards for adjudicating between them. I consider three broad categories of revision here, one for each premise in Joyce's Deduction, since revising the folk commitment that any one premise characterizes and summarizes would be sufficient for dodging the error theoretic conclusion. Ultimately, I will promote and expand one strand of revision that looks most promising, and will recommend on its basis a form of contextualism.

Before I get started with these possibilities, however, I should say something about methodology. Because I am addressing explicitly and self-consciously revisionist options, I will not spend much time focusing on familiar objections to the theories in question. For example, when considering the view that moral obligation is relative, I do not make my case by simply building arguments based on individual absolutist intuitions. Rather, I ask whether we should give priority to our competing relativist or absolutist intuitions, primarily in light of what consequences for our moral discourse and practice such a prioritization would have. Revising our conception of moral obligation may amount to more than inert metaethical revision, bear in mind. In many cases, the revisions can force upon us changes in our first-order moralizing, where our intuitions are often very powerful and important to us. Metaethical revisions thus have costs of two related kinds: Not just revision of our metaethical beliefs, but also revision of the ways we actually moralize. In terms of these costs, any of the revisions may be *prima facie* warranted, since each would allow us to avoid an error theory, but it is by asking which revision would require the most and weightiest change that we can decide between the competitors.

Often, I will be focusing on considerations that have historically supported expressivistic theories, since expressivism tends to give current, actual moral discourse and practice the most credit, and does not recommend changes in first-order moral beliefs, which would be a form of costly doxastic damage. These considerations will include phenomenological facts about what it is like to make moral judgments, motivational facts about what mental states typically correlate with moral judgments, facts about the function of moral language and the speech acts it is used to perform, as well as more general facts about the role of moral standards in structuring and coordinating social groups. These facts are important, and have given support to expressivism in various ways, but they are – in the end – truths that may be made false by the modification of our treatment of moral language in the post-revision future.

Revisions will be more or less warranted in large part due to how much or how little they disrupt these facts. If I am right, there is one form of revisionism that will disrupt very few – or at any rate fewer than competitors – and so will be, all things

considered, warranted and recommended. It is clear that moral language has certain uses and functions in our lives, allowing us to express ourselves in important ways – ways that carry consequences for how we get along together. A revision of moral concepts that would impoverish these expressive capabilities would drastically alter how we get along, and force us to reconsider much of what we think we know about the relationships between our feelings and our expectations for ourselves and others. A theory that preserves these expressive capabilities would employing problematic concepts is to that extent preferable, calling for much less doxastic restructuring. Morality is fundamentally practical, and our theory of it should acknowledge this.

6.1. Revising Obligation

The first revisionist possibility would require us to change how we think about moral obligation. As Joyce puts it, the folk theory assumes that "If x morally ought to ϕ, then x ought to ϕ regardless of whether he cares to, regardless of whether ϕing satisfies any of his desires or furthers his interests".[6] Put another way, our moral obligations obtain neither in virtue of our desires to do as the obligation requires (*de re* – relying on *de dicto* desires would be inescapably circular here), nor in virtue of facts about what is in our interest. We can say also that moral obligations do not obtain in virtue of what the obligated agents value, so as to not beg the question about the relationship between desires and values. Revising the folk theory in any of these ways would avoid the error theoretic conclusion, but all three revisions belong in the same family: They all suggest that agents' moral obligations are consequences of the same facts about them that would provide them with internal reasons for doing what morality requires (*de re*).

To make this point more clear, it is important to note that only certain types of desires, values, or interests are even contenders for being the basis of moral obligation. It would make no sense to say that an agent ought to give to charity because they desire to eat ice cream, or because they value literature, or because buying a new car is in their best interest. The connections between desiring, valuing, or interest and moral obligation must be of the right kind: The morally obligatory action must either satisfy some relevant desire of the agent's, be valued by the agent, or be in the agent's best interest. Any action that stands in any of these relationships with an agent will also give the agent internal reason to perform the action, and this is exactly why the revisions allow us to avoid error theory: They state that for any moral obligation, there is also a corresponding internal reason, so we avoid commitment to external ones. The moral reasons are themselves internal.

[6] *Ibid.*, 42.

The most obvious cost common to this family of views is that each flies in the face of moral phenomenology. When we make moral claims and come to conclusions about what morally should be done, it simply does not feel like we are making commentary about the desires, values, or interests of the agents in question. Indeed, we often make moral claims categorically, without any mention of particular agents at all. But more importantly, I think, are the implications for how we should go on moralizing in the post-revision future: Insofar as we want to continue making moral claims as we do – without relativizing to facts about agents – we must find a desire, value, or interest so basic to human psychology that we would be allowed to assume nearly everyone has it. Those who lack it would simply be outside of morality, escaping its obligations.

There are a few problems here. It is empirically dubious that such a psychological fact could be found, especially because not just any fact will do. It must also be one that actually has something to do with morality. Eating a well balanced diet may be in almost everyone's best interest, but eating a well balanced diet is not therefore a moral obligation. (A revisionist theory could in principle rely on such unlikely candidates, but the doxastic damage caused by the revision would increase in direct proportion to the proposed candidate's intuitive implausibility.) Among universal desires, we might consider the desire to be happy, for example, but different things make different individuals happy, and the things that make them happy most often have nothing to do with morality. It would be equally difficult to find universal values or interests. There are, however, some candidates that are more promising than others, but I think there are also general objections that are unavoidable.

In terms of desires and values, there is psychological evidence of nearly universal, though limited, evolutionarily explicable human altruism. We are social beings, and we most often do, as a matter of fact, have desires to help others, and we value a society where everything goes well and equitably. We value fairness, we desire to be kind, and we want to behave in ways that our social companions will not object to. Moreover, many apparent divergences in values between different groups may really be manifestations of more general desires in different environments.[7] With all this empirical support, it is indeed tempting to try out these desires and values as the basis of moral obligation. Unfortunately, even these values and desires would be insufficient. It is not only, as I mentioned, that those lacking the value or desire would escape the obligations, but that those who lack the desire are the ones morality ordinarily condemns the most. (In Section 3, it was noted that the folk views morality as designed to "combat rampant individualism".) We could no longer say that the purely selfish person would have done something wrong by torturing for pleasure, and

[7] James Rachels, *The Elements of Moral Philosophy* (New York, N.Y.: McGraw-Hill), 23-5.

we might find it more difficult to punish them: Do they deserve punishment, given that they are not a wrongdoer? We could only criticize those whose behavior was misaligned with their own desires or values, and morality would thus be identified with a branch of instrumental reasoning. Moral language would become practically useless – impotent to condemn the worst evils, and at best only used to tell others they acted in a way consistent or inconsistent with their desires or values, as though moral injunctions had as little significance as telling a friend who enjoys ice cream that he ought to buy some if he wants to be happy.

I consider interests separately because desires and values consist primarily in conative mental states of which an agent is aware, while agents may be ignorant of what is in their interest. This difference presents different theoretic possibilities, although what is in an agents' interest does depend to some extent on agents' desires and values. This means that, using interests as the foundation for moral obligation, we would not need to rely on contingent psychological facts in the same way, but could rather rely on more general facts about what is in individuals' interest, which might at first seem more promising. Whether or not one values having positive relationships with others, having friendships, or behaving in ways others would accept, we might be able to say that such things would nonetheless be in such persons' interest. Even the misanthrope who loathes the company of others would likely be better off if only he could manage to forge meaningful friendships, for example.

For all its merits, however, it seems unlikely that this account could make sense of many of our important first-order moral intuitions. If a snake oil salesman lacks concern for the townspeople, knowing he will flee in the night, it might not be in his interest to refrain. On the other hand, it might be in his interest to adopt a lifestyle of honesty and kindness, since this would promote healthy, meaningful, and even beneficial relationships with others.

There are again other exceptions. There are of course individuals who can maximize their well-being in intuitively immoral ways. If the snake oil salesman lacks guilt, and works with other rogues with whom he has lasting friendships, he might not gain anything by converting to an honest lifestyle. The misanthrope who finds genuine enjoyment in solitude may have no reason to cooperate with others. And once again we find that the revision in question would prevent us from making moral judgments about those who most deserve our criticism. We could only shrug our shoulders about the truly calloused and depraved, while the remorseful and regretful would be guaranteed our full censure.

There is one final worry for views of this kind that has been hinted at throughout: Each view renders moral facts as something empirically discoverable, as facts about psychology and social circumstance. Many philosophers do not find a problem with

views of this kind, but it is worth noting what the consequences would be for our moral discourse and practice if we went this route. Morality as a discipline would be passed from the philosophers to the scientists, and those who firmly disagreed with the scientists' moral conclusions would be ignored, guilty of not being up-to-date with the revised language. Indeed, they would lose the ability to properly express their dissent at all. When trying to make moral claims, they might not be making any claims at all about desires, values, or interests, but rather aiming to express – perhaps among other things – their approval or disapproval. Blackburn describes this as something like taking a place on a normative staircase, where different places on the staircase are characterized by the expression of different degrees and varieties of emotions and by the having of certain motivational predispositions.[8] Perhaps there are many positions on the staircase that one can be in while moralizing, but there are positions one cannot reach without moralizing in a way that this revision would disable us from using our moral language to do.

In summation, the problems with this sort of revision are, at least, the following: (1) It would have moral investigation amount to the empirical task of trying to find what would satisfy relevant desires, (2) There is insufficient similarity throughout human psychology to support anything but an unsatisfyingly relativistic morality, (3) Moral language could no longer properly be used to censure those who intuitively most deserve censuring – those who lack the relevant desires or values, and (4) Moral recommendations would become a sub-class of instrumental recommendations, and so individuals could never use moral language to pass higher-order judgments about the desires themselves, or, more generally, to express the sort of approval or disapproval moral language is currently uniquely positioned to express – that of people's actions, independent of facts about what instrumental reasons they may have.

6.2. Revising the Obligations-Reasons Relationship

The next possibility to be considered concerns the second premise of Joyce's Deduction, which captures another folk intuition about morality: Anyone who is under a moral obligation to ϕ has a corresponding reason to ϕ. Combined with the first premise, this implies that agents may have reasons for action that are independent of their interests or desires. Here, I will consider what we can do by accepting the first premise while rejecting the second. Perhaps moral obligation is inescapable, but moral obligation alone does not provide reasons for those who lack the relevant motivations.

[8] Simon Blackburn, *Ruling Passions: A Theory of Practical Reasoning*, (New York, N.Y.: Oxford University Press, 1998), 8-12.

Put more concretely, the view in question allows us to say that it would be wrong of the sensible knave to give a lying promise for the sole sake of personal gain, but it requires us to say that he nonetheless might not have a reason to refrain. This is exactly the sort of theory espoused by Railton in his "Moral Realism": Morality is understood as an independent "species of evaluation" that provides a categorization of actions into those that ought to and ought not to be done, but it only has reason-giving force for those who have (*de dicto*) desire to do as morality requires.[9] In this way, morality is rendered comparable to an institution or game that provides rules: A sly cheater in a game trivially ought not cheat according to the rules, but he may nonetheless lack reason to refrain, provided he is sly enough to get away with it.

This possibility has some definite advantages over the previous one. It allows us to continue moralizing about even the most callous and comfortably wicked agents, and we may properly call such agents wrongdoers. But it does have disadvantages as well. We lose the ability to cite moral facts as reasons, and consequently, morality itself seems to lose significance. Moral obligations remain inescapable, but only in a very weak way, since they themselves cannot provide reasons. We must admit that evil agents might lack reason to cease doing evil.

I will not put too much emphasis on this last point, though. While it is perhaps initially hard to swallow that evil agents lack such reasons, we do have other intuitions that back it up. We become so accustomed to citing moral facts as reasons because people generally are concerned about the moral status of what they do, but when we encounter someone who lacks this concern, it *does* seem pointless to continue saying "*Why* not ϕ? Because ϕing is just *wrong*". Once we realize that someone will not be convinced by what we take to be reasons, there is some pressure to assume that maybe the person does indeed lack that reason (i.e. that they lack the appropriate reason-giving motivations). When we believe that doing what is required to succeed in school, for example, is very important, we might think this gives us reason to study and work hard. But when trying to convince a lazy peer that he ought to start a paper early because it will boost his grade (and because this will ultimately help him succeed in school), we might cease to believe that the peer actually has any reason to start early if he repeatedly responds with apathy to the supporting reasons we cite (that it will boost the grade, and so on).

There are other, more important problems. This sort of revision would recommend a change of the way we use moral language, and as a result of this change, the status of morality would be *degraded*, and in some cases morality would become wholly irrelevant. It would render all-too-reasonable questions of the sort "Why should I care about morality?", which are intuitively bizarre at best. (Recall Sec-

[9] Peter Railton, "Moral Realism". *Philosophical Review*, 95, pp. 163-207.

tion 3: The folk theory takes moral facts to *just be* deliberative considerations, i.e., things that should be taken into account when deliberating.) An analogy would be defining 'delicious' in terms of a standard of evaluation for taste, about which we could make judgments of foods, in terms of 'better' or 'worse'. This would allow people to sensibly say, "Well, this food is certainly delicious, but I don't like how it tastes". My argument – and the argument I think expressivists would respond with – is that moral language shares this function of *approval*, and so should not be revised in such a manner. Let me start by first discussing what this new use of moral language would be like.

As mentioned, this sort of revision leads us to accept morality as a standard of evaluation that is independent of – although normally correlated with – what we have reason to do. The revision, however, does not tell us what this standard is, and it does not necessarily need to, at least not yet. Presumably, there will be many different arguments for and against the various competing substantive characterizations of the standard, but they need not detain us here. What is important is that there *will be* such a standard, whatever it ends up being. Relativism would not be a helpful alternative, because moral disagreement would no longer be a matter of disagreement about what to do, but rather just about what belongs to a certain class which we can reasonably not care about. If we were relativists about morality in this framework, the relativism would be no more significant than relativism about the meaning of 'chips' in different English-speaking countries. Moreover, since in this framework we are stipulating standards in some sense – by the revisionist nature of the framework – it is hard to see what the harm would be in stipulating a universal standard. Fine-grained standards for different areas would not, in any case, help to avoid the big problem here: Whatever the standard, people could reasonably ask whether or not to guide their conduct by it. Moore's open question argument arises most conspicuously here, which again just goes to show how this revision would force a fundamental change in the way moral language is used.

One common response to the open question argument is that it reveals to us that the primary function of moral language is to approve or disapprove, not to describe. This would account for why, for any object with any natural properties, we can still coherently ask whether or not it is good. Even if some act or circumstance is *good* according to the standards we have decided to identify with morality, one could still coherently disapprove of it. Of course, the open question looks different now, because the word 'good' itself is appropriated by the set of standards, but the open question argument shows us that it becomes thereby *necessary* to invent new language to do the job 'good' and its counterparts were designed to do: Express certain sorts of approval and disapproval. In this case, rather than *vindicating* moral language via

revision, we would reduce its role to one that is nearly irrelevant, and leave a vacuum to be filled by a *new* sort of language, one which would itself become the one that should interest metaethicists. We would take the name of something we cannot yet describe and stipulate a description to correspond with the name, but we would not thereby illuminate anything about what we set out to describe. A more successful revision will be one that does not alienate the revised moral concepts from the prior functions and uses of moral language.

For these reasons, I think this species of revision is also deeply flawed, and so I move on the final possibilities explored in this paper.

6.3. Revising Reasons

Lastly, I consider the possibility of revising the folk commitments about reasons, which requires finding a way of making sense of external reasons, those that apply to agents independent of their particular interests, desires, or concerns. It is this sort of revision that I think is most warranted. There are indeed costs to be incurred, but they are somewhat insignificant in comparison to those incurred by the foregoing rivals.

The largest and most immediately obvious cost in developing a theory of this kind is that it recommends a new conception of *rationality* itself. According to what has been said, the folk conceive of reasons as the sorts of things that obtain in virtue of certain facts about agents' motivations, but this form of revision would overturn this whole instrumental conception of rationality. Thus, the largest part of the burden here lies in softening this blow while stressing other benefits.

One way of softening the blow that might be considered would be to stipulate that the only sorts of reasons that defy this instrumental paradigm are moral reasons. This would minimize the doxastic damage in some sense, since we would only need to think about one sub-class of reasons differently. But the stipulation is very *ad hoc*. The problem with that, in revisionist terms, is that it defies some truly powerful intuitions about epistemic norms that favor the more parsimonious of competing theories. So, in our effort to minimize doxastic damage, we would in the end increase it. The softening of the blow must come in other forms.

In order to make progress toward this end, it will be useful to think more carefully about what the acceptance of external reasons would amount to. *What are* external reasons to be, and what would we say about them if we allowed them into our theory? We will want to know more about which purported external reasons are genuine and which are not, and about what allows us to determine this.

When it comes to internal reasons, the truth or falsity of reasons-ascriptions

depend on motivational facts of the agents for whom the reasons are said to obtain. Obviously, we cannot look in the same place for the truth-makers of external reasons-ascriptions. What we can do, though, is look for relevant differences between internalist and externalist language . Internalist language is notably used to issue recommendations or to advise. We can use it to make suggestions to our peers about things they should do, given their own interests and concerns. It provides a way to help them reason rationally given the normative premises they antecedently accept. Externalist language is quite often not used in this way at all. We use this kind of language to, for example, censure criminals, saying that murderers failed to do what they ought to have done. Our goal in using such language is not to criticize the murderer's instrumental rationality. More plausibly, our goal is to express something about ourselves to our fellow interlocutors, such as our disapproval of the murderer's actions. We can express our commitments to not murdering, our expectations that others act in like manner, and our preparedness to censure those who do not. This gives us quite a bit of direction in terms of characterizing external reasons: They are not, we may say, the sorts of things that have to do with agents' interests or concerns, but *rather* with commitments or expectations of the interlocutors – or, their concerns broadly construed.

What this suggests is that we can perhaps take facts about *speakers*, rather than about agents, to be the truth-makers for externalist reasons-ascriptions. This is clearly a form of relativism, but it is one of relativization to the *context of utterance*, rather than to the context of action. Accordingly, if we adopted this theory, we would not be able to judge absolute rightness or wrongness between two separate communities with incompatible concerns. We can, however, judge of the speakers in those contexts that they often say truthful things, and that they are employing language in a purposeful, effective, and ultimately *proper* fashion.

I will develop this sort of theory under the broad title of *metaethical contextualism* in the sections that follow. I think it provides the best and most comprehensive vindication of our moral discourse and practice, since the primary content of the revision concerns only meta-normative beliefs about the nature of reasons. But not only does this revision thereby vindicate the great bulk of our moral discourse and practice, but it coheres well with other facts about what we intend to accomplish with externalist language.

7. Metaethical Contextualism

I have considered two broad strategies revision might take and have taken some steps to diagnose their shortcomings. If my suggestion is correct – that the root of their

failure is not taking our first-order judgments seriously enough – then we can begin to work our way toward a more successful revision by heeding this suggestion. That is, when we think about how best to revise our moral theory, we should place a maximal amount of significance on our first-order judgments. We must grant that most people, most of the time, are using their moral language properly. That is, we should allow Moore's open question to remain open. Whenever we can avoid it, we must not categorize uses of moral language as incorrect: We should rather strive for a theory that is maximally vindicating with regard to these judgments. Along the way, we may be forced to shuffle around any number of our higher-order intuitions about moral concepts or moral language, but by maximizing first-order vindication and minimizing higher-order disruption, we should arrive at the most palatable revision available.

Of course, as we have already seen, there is great *prima facie* difficulty in giving so much credit to our first-order intuitions: Between individuals, and between groups, there is significant disagreement. Behind some of this disagreement, there may be shared, fundamental intuitions from which different people draw different conclusions, but we can't assume that this will be the case for all, or even most, of such moral disagreement. But given this disagreement, taking all these intuitions for granted would build contradictions into our theory before it even got started.

There are ways of vindicating these first-order intuitions without arriving at such contradictions, however: We will simply need to reconsider higher-order intuitions in order to get things going. In particular, we need to reconsider the assumption that moral truth is absolute, that it is the same everywhere. What we need instead is a way of binding moral assertions to the contexts in which they are made, so that no abuse of language is involved when people are tracking different properties with their moral terms. What we need is to adopt a form of relativism, specifically the sort contextualism according to which the truth-value of people's moral assertions is relativized to the context of utterance.

This sort of contextualism is just what results from Stephen Finlay's semantic analysis of moral language. According to his "unified theory" of reasons, "all normative reasons [...] are *end-relational*: a fact is a reason for ϕ-ing, relative to a system of ends E, iff it explains why ϕ-ing is conducive to E".[10] If we are assuming that, by and large, most uses of moral language are felicitous, and that they require for their felicity a contextually-given system of ends, all we need to vindicate our first-order judgments is to show that they are made relative to a system of ends. And this doesn't seem so difficult: When someone makes a claim about what others ought

[10]Stephen Finlay, "The Reasons that Matter". *Australasian Journal of Philosophy* 84 (1): 1-20, 2006.

to do, and are asked *why*, they can say why by making a claim about the action's conduciveness to some end.

It is crucial to this account that we do not discard any potential ends as being somehow illegitimate or non-justifying: If some context makes it the case that the end of murder is most relevant, claims about what ought to be done will be evaluative relative *to that end*, whether we like the end or not. We may try to effect changes in context to make a different relevant, or we can opt out of the conversation, but those who operate within the context make no linguistic or conceptual errors answering "*why* ϕ*?*" questions by pointing out the conduciveness of ϕ-ing to achieving murder.

This account of normative language sits well with plausible theories of modal language in general. In particular, the account given by Angelika Kratzer in her "What 'must' and 'can' must and can mean" describes these modal terms as being relative and taking two arguments: Roughly, one argument is what is being evaluated, and the other argument provides what that argument is being evaluated *in view of*.[11] It wouldn't be helpful to go into detail about this, but the point is that plausibly all modal language is evaluated in view of *something*, whether that something is *what is known (by anyone)*, *what is known (by a particular person or group)*, *what is metaphysically possible*, or something else. If modal language turns out to be like this generally, it shouldn't surprise us if deontic language is all evaluated in view of systems of ends. Indeed, if one wanted to accept Kratzer's general approach to modality but wanted to remain a reasons internalist – *contra* Finlay – they would have to contend, apparently *ad hoc*, that when it comes to certain kinds of normative language, only systems of ends that matter to the agent(s) under discussion can fill in that argument of the modal function.

However, within this acceptability lurks exploitability: One could use Finlay's framework to support non-relativistic theories of morality, in spite of the initial appeal of Finlay's approach being in its accepting even *prima facie* contradictory reasons-ascriptions as being equally good, provided each is equally accurate relative to the system of ends within its context. The complication here arises when we consider how systems of ends are to be specified, specifically when it comes to moral ends. When conversations are engaged in *moral* discourse, we should expect that the system of ends provided by the context is also in some way *moral*, but in what way?

There are a few ways of characterizing the moral context, not all of which get us what we want. We can say that the discourse is moral if the relevant end is the end of *being moral*, or that the discourse is moral if the ends in the relevant

[11] Angelika Kratzer, "What 'must' and 'can' must and can mean". *Linguistics and Philosophy* 1: 337-355, 1977.

system of ends are all individually moral ends. For the first possibility, we must determine what is conducive to the end of being moral, and for the second, we must determine which ends count as moral ends. On either route, there are difficulties, and in trying to say enough to characterize morality, there is the serious risk of, as mentioned, undermining the relativism that our revisionist priorities prompted us to pursue. My contention is that the first is preferable. It does better justice to how we intuitively characterize our moral discussions and disagreements.

One consideration that favors the first of the above two approaches is that people often do think of themselves as acting for the end of *being moral*, and it is natural to say, in moral contexts, that the reason to do something is that it would be *moral* to do it, i.e. conducive to the end of being moral. There are also some negative considerations that favor this view. According to the opposing view, we should be able to give criteria for an end's being a moral end, or at least be able to list the moral ends, but it is hard to see how this could be done in a way that begs no questions, or that does not discount various first-order intuitions that – for revisionist reasons – should not be discounted. If we took the moral ends to be provided by any particular, substantive moral theory, we would obviously be committing this error. But even if we favored no particular theory and instead listed the moral ends according to some descriptive features (e.g. their having to do with interpersonal relationships or social structures), people's beliefs about obligations and so on would have to either fit these criteria or would just be non-moral by definition. This is a big bullet to bite, and if we can avoid it, we should. And there is an even worse problem. If a given context of conversation is decidedly moral, and it is moral in virtue of the relevant system of ends containing only ends that belong to the class of moral ends, we should expect the interlocutors to be in agreement about these ends' being the relevant ones: As Finlay puts it, context is "roughly the totality of information mutually recognized by the participants in a conversation as mutually understood".[12] If these ends are part of the context, they should thus be mutually recognized as mutually understood, but in many conversations about what there is moral reason to do, *which* ends are moral ends is precisely what is *at issue*. So, it really does look like this is not the way in which we want to characterize such contexts: We want, instead, to adopt the first approach, and take the end in contexts of moral discourse to be the end of *being moral*.

An analogous issue, however, arises for this approach. What, after all, is it to be conducive to the end of being moral? Without an answer to this question, we have no way of determining truth values for reasons-ascriptions in such contexts. Again,

[12]Stephen Finlay, "The Reasons that Matter". *Australasian Journal of Philosophy* 84 (1): 1-20, 2006.

we could try answering this question by appealing to some substantive moral theory, but this would still be question-begging, discounting all first-order intuitions that are contrary to the theory we choose – and, again, the revisionist considerations that prompted us to pursue contextualism would be flouted by such a maneuver. The goal, once again, is to be as hospitable to these first-order intuitions as possible, since they are what are most significant in representing moral language as shared language. Disregarding some intuitions in order to easily preserve our personal favorites would do injustice to the linguistic competencies of those who do not share our intuitions.

Fortunately, there does seem to be a way to arrive at a satisfactory answer to the question of what is conducive to the end of being moral. In the sections that follow, I will begin to develop one such answer.

8. Moral Contexts

Before exploring this much further, there is one more question I must address: Given that – according to what I have argued – we cannot say which contexts are moral just by examining whether or not their systems of ends are subsets of the set of moral ends, as described by substantive or descriptive criteria, how are we to know when contexts provide the end of *being moral* as the most relevant? In other words, what are the hallmarks of moral contexts or of particularly moral reasons-ascriptions? It might be objected that there are no such hallmarks, and that it was my assumption that were such hallmarks that misdirected me in the previous section: We don't need to say *which* ends or systems of ends are *moral* (the objection could go), since all that matters is that certain ends are relevant in certain contexts. Sometimes conversations view the relevant ends as being moral, but sometimes not. All that matters is that they are prepared to accept reasons-ascriptions that accurately describe actions' conduciveness to that end, whether or not the end falls into such-and-such a category. When interlocutors agree that they are in a moral context but disagree about which ends are moral, we can just characterize this as a failure to accommodate each other's conversational goals, without ever needing to say *who's right* about which ends are really the moral ends.

But, contrary to this objection, there is a great deal that is important and distinctive about moral contexts. We can see this, I argue, by examining the way in which putatively moral reasons-ascriptions affect context in ways that other reasons-ascriptions do not. I admit in advance that there are other ways of explaining the phenomena, but I think recognizing moral discourse as a distinct sort of discourse, warranting an independent analysis, is not only the most intuitive, but also the cleanest way of explaining what we observe. To take a simple example, consider a context

in which the relevant end is *winning at chess*. In such a context, there is a reason to ϕ just in case ϕ-ing is conducive to winning at chess. Now, when a new interlocutor enters this context and says that someone ought to ϕ because ϕ-ing would be conducive to losing at chess, this claim is straight-forwardly false in the context, since the speaker attempts to make a claim relative to end that is inconsistent with the contextually provided end. But if the speaker had instead said that someone ought to ϕ because ϕ-ing would be the moral thing to do, it becomes an open question in the context whether or not on ought to ϕ, even if ϕ-ing is not conducive to the end of winning at chess. What happens, of course, is that the context *changes* in light of this new claim. Context-shifting is not itself unfamiliar, but my contention is that moral claims effect changes in context in a unique and particularly systematic way. Once putatively moral claims are made, context seems to *almost always* shift, and it often seems to shift independently of interlocutors' will to the contrary. Moreover, once shifted in this way, context resists reversion: The old ends cannot just be taken for granted again, but must be defended in moral terms.

So my proposal is that we can tell when a conversational context is moral context by examining the interlocutors' behavior with respect to new and old potential ends. If the relevant ends in the context override previous ends and cannot themselves be overridden by new ends, we know that the interlocutors are engaged in moral discourse, and that the end of the present context is the end of being moral.

There are, of course, cases of amoralists – agents who purported do not care about morality – that will cause difficulty for this analysis. Contexts these agents are a part of seem to behave differently, and thus could be given as counterexamples to what I have argued. I will consider this more fully later when addressing the semantics and pragmatics of moral disagreement, but there is one thing I can say for now on the matter. If it is true that moral reasons-ascriptions are overriding in the manner I have described, sense can be given to our intuitions that some some reasons are *not* end-relative, but rather obtain no matter what ends are presently relevant. We treat moral reasons as not only external (obtaining independent of agentive motivations), but as also categorical (obtaining independent of particular ends). If in moral contexts, the end is one that overrides all other ends, we can interpret talk of all-things-considered reasons as talk of moral reasons without losing much. I think talk of overriding, all-things-considered reasons *is* worth vindicating, and thinking of moral language as the language we can use to criticize other particular ends by shifting context to override them would allow us to do just this. If, in order to achieve such vindication, it turns out to be necessary to somehow disregard amoralists or to give up vindicating their beliefs about morality, this may just be a price worth paying. Nonetheless, I hope to show later on that there are actually a

few plausible ways of thinking about amoralists so that they should not cause too serious of problems for the theory I am developing.

9. Metaethics Doubly Contextualized

9.1. The Proposal

It would be useful at this point to briefly review what has been said. I argued that a revisionist framework would be interesting and worthwhile – if not necessary – to apply to metaethics, and that according to such an application, we should develop a theory that focuses primarily on the vindication of our first-order moral intuitions. I considered two broad revisionist possibilities each of which failed in its own ways. I diagnosed the general problem, roughly, as being one of treating moral language as too absolute. If we can, we want morality to apply to everyone with the force of providing reasons for action, but when we couple this with the diversity of people's motivational dispositions, it became clear that a form of speaker-relativism, or contextualism, would be our best bet for vindicating our language. However, in the process of fleshing out this contextualism, it was shown that there are difficult issues regarding *what* we are relativizing to. The only plausible candidate, I argued, was the end of *being moral*, but this unfortunately did not get us very far. The contextualism provided promise by showing how external reasons-ascriptions can make sense, but did not go far enough in characterizing moral contexts and describing what it is to be conducive to the end of being moral.

My suggestion at this point is that we can solve the remaining problems by contextualizing once again. We started thinking of which reasons-ascriptions were true by thinking about what ends interlocutors take to be relevant in contexts of utterance, and now we should be able to think of which *moral* reasons-ascriptions are true by thinking about what *moral commitments* interlocutors have. Once context fills in the relevant end as *being moral*, it goes on to fill in what being moral *amounts to* by examining another aspect of interlocutor concerns.

9.2. The Determination of the Moral Context

There are many candidates for what we should take to play this role in the contextual determination of the moral standards. We can begin to hone in on the best candidate by considering different contexts and looking for relevant commonalities and differences among them, as well for relevant commonalities and differences among the mental states of the interlocutors within those contexts. In particular, we should compare and contrast the contexts in which moral discourse goes as smoothly as

possible with the contexts in which something intuitively goes awry. Many everyday contexts fall into the former category, and contexts in which there is intractable disagreement in fundamental moral standards fall into the latter.

9.2.1. Shared Values

In diagnosing what causes moral discourse to go awry in such contexts, A.J. Ayer remarked that argument is possible on moral questions only if some system of values is presupposed.[13] This is a good start, and intuitive, but does require refinement. We may think of systems of values roughly as sets of beliefs about what sorts of objects bear which moral properties non-derivatively. If two consequentialists engage in a disagreement about whether some x is morally good or not, but each has a different belief about which properties of objects are responsible for their moral goodness, argument, in Ayer's sense, does seem impossible: Neither interlocutor can present an argument that would rationally convince the other of their moral conclusions, since there deductions would all hinge on their unshared basic moral premises. So, if argument requires shared systems of values, perhaps we could say that what determines the moral facts in conversational contexts is the shared system of values. This would explain why contexts that meet this requirement tend to produce successful, cooperative discourse, while those that do not tend to look dysfunctional, producing only intractable disagreement.

But it seems too strong to require convergence in whole *systems* of values. This restraint seems most plausible when we limit our discussion to consequentialists whose whole systems of values can be summarized by single premises. But we can picture deontologists engaging in moral argument successfully even when their is deep disagreement between their systems of values *except* with regard to the value that is in play in this particular conversation. Suppose that two interlocutors disagree morally about virtually everything *except* that murder is wrong. They may discuss an action and make arguments for the conclusion that the action is wrong by combining their shared moral belief about murder with premises according to which that action counts as murder. So, here, moral discourse proceeds quite successfully. We might thus conclude that it is not a shared *system* of values that is required, but rather just shared individual, conversationally relevant values.

Even if we limit the requirement to convergence upon the individually conversationally relevant values, it will still be too strong. As I have been discussing values, they can be characterized by conditional beliefs about which non-moral properties of objects entail their having which moral properties, but given this characterization,

[13] A.J. Ayer, *Language, Truth, and Logic* (New York, N.Y.: Dover Publications, 1952).

people can have entirely different values while still agreeing morally. That is, two people can believe that something has a certain moral property, but *for different reasons*. So, as it turns out, I don't think *values*, understood in this way, do the job quite right. Perhaps in contexts without any shared basic values, argument for particular first-order conclusions will not be possible, but this is not the only kind of argument or conversation that can take place when the moral end is presupposed. People can discuss quite simply what ought to be done morally, and as long as they share the relevant first-order evaluative beliefs, conversations of this kind can proceed happily. One person might believe murder is wrong for utilitarian reasons, and another person might share this belief, but for religious reasons. They may nonetheless argue about whether some action is wrong by arguing about whether or not it counts as murder. No shared *values* are required, at least not in the sense of *basic evaluative beliefs*.

9.2.2. Shared Norms

So we want neither shared systems of values or even shared individual values to be the things that determine what, in a particular context, is conducive to the end of being moral. But in rejecting these candidates, we have gotten much closer to seeing what a better candidate might be. What we want is something can be separated from a full system of values, and something that people can agree about even when they do so for different reasons. It is the thing that allows for people to reach meaningful agreement about, say, the wrongness of murder, even when they are in disagreement about why this wrongness obtains.

I think there is precedent in the metaethical literature for a mental state like this. According to Allan Gibbard's norm-expressivism, moral language is used to express mental states of norm-acceptance, where these mental states are understood in terms of their places "in a syndrome of tendencies toward action and avowal".[14] When one person values the maximization of utility and another values the word of God, their action- and avowal-tendencies toward things such as murder may be shared, *even though* their values are in an important sense different. So, my ultimate proposal is that what is responsible for the determination of conduciveness to the end of being moral in conversational contexts is the *shared acceptance of norms*. There is much more that can and should be said about such mental states, but here I can only hope that the reader finds it plausible that sense can be made of them. If the idea does turn out to be psychologically defensible, it is likely the best candidate for playing

[14] Allan Gibbard, *Wise Choices, Apt Feelings*, (Cambridge, MA: Harvard University Press, 1990), 75.

the role under discussion.

There is one last thing on this topic that must be mentioned. Quite often, more than one norm is relevant at a given time. Norms interact with each other in certain ways, and there may be higher-order norms accepted by thinkers that govern how they think about the relationships between other norms. When discussing what to do, it is often a *system* of norms that is in play, but the system need not include all the norms accepted by the interlocutors, but rather just the ones that are relevant to the discussion at hand. This relevance can perhaps be understood in terms of *employment*: If the interlocutors employ certain norms in their discourse, they become relevant to the conversation. Where these employed norms are shared, the conversation should run relatively smoothly. Where the employed norms are not all shared, intractable disagreement is likely to arise again. I will say more later on about this kind of disagreement.

9.3. Further Motivation for the Proposal

With this in place, we can begin to examine how this particular contextualist approach allows us to vindicate certain crucial aspects of moral discourse and practice in a way that neither the previously considered revisionist alternatives nor Finlay's unamended contextualism can.

Among the benefits of this approach, its most important is the salvaging and vindication of three important aspects of our moral discourse and practice: The externality of our language, the categoricity of our reasons-ascriptions, and the confidence commonly placed in our first-order intuitions. I will take these in turn, along the way stressing *why* it is so beneficial for a theory to treat these three aspects in this way.

9.3.1. Externality

There are two senses in which moral language is used externally: We intuitively take moral obligations to apply to agents independent of their particular motivations, and we intuitively take these obligations to provide reasons that apply in like manner. The first two revisionist paths considered failed to do justice to these intuitions. The first of these failed somewhat spectacularly by disabling us from morally criticizing instrumentally rational wicked agents. The second did slightly better by permitting moral criticism, but its deference to our intuitions about the internality of reasons generally rendered our moral criticism nearly vacuous. Morality was viewed as merely one system of evaluation among many, so the importance of doing good by this system was rendered entirely contingent on what happen to be our concerns. I argued

that these considerations should cause us to seek out a vindication of externality if possible. Finlay's contextualism offered this.

By noticing the different uses to which we can put normative language, we can divide it roughly into two categories: The language used to make recommendations, and the language used to express our commitments and expectations. (This division is imperfect and provides categories which are neither mutually exclusive nor comprehensive, but thinking of these two aspects of our language separately is nonetheless informative.) The language of recommendations unproblematic and internal: We discuss what agents have reason to do given their motivations. But what Finlay observed is that we can think of such language as itself being relativized to instrumental ends. We can salvage externality by relativizing the language of commitments and expectations to *other* ends. The mistake all along was supposing that we could only properly talk about reasons relative to agentive ends. Finlay's unified theory does away with this error. So, when we use external language, we can truthfully say that people have reasons to do things they have no instrumental reasons to do.

This defense of externality is preserved in my modified contextualism. When engaging in moral discussion, our claims about what reasons people have are made true or false by facts about the conduciveness of actions to the end of being moral, whether or not the agents we are discussing behave with this end in mind. Moreover, we can properly disapprove of the behavior of people in different moral communities, with different moral standards, in spite of the relativistic character of the theory: Which actions are counted as being conducive to the end of being moral in a given context are determined by the relevant and shared norms of that context, and are *not* determined by the accepted norms of the agents being talked *about*. Accordingly, this contextualism vindicates our intuitions about the irrelevance of agentive motivations to the questions of which obligations and reasons we have. When agents have moral obligations or reasons to act or not act in certain ways, they have these whether they like them or not.

9.3.2. Categoricity

The categoricity of folk normative language is exhibited by questions or claims about what we have reason to do"whatever our ends". This is similar to externality, but different in important ways. Most notably for the purposes here is that Finlay's theory salvages externality but thoroughly does away with categoricity.

Finlay won externality by going end-relative across the board for reasons-ascriptions, but with the consequence of disabling us from properly making any reasons-ascriptions that are explicitly not end-relative. The problem is that we *do* often say things that,

in order to ever be true, require there to be categorical reasons. This happens regularly when we are in contexts whose relevant ends are ones we do not endorse. For example, some interlocutors may be discussing how a politician could successfully deceived the public, given that he wanted to. It might be said, "He should have shredded the documents", and this would be true in the context. But another interlocutor can shift the context by interjecting that, "Just because it would have been in his interest doesn't mean he should have done it". The utterer of this claim could be shifting the context to a new, external end (such as *maximizing utility for all*), but this hypothesis would not do a great job of explaining why the context would resist new shifts, and why the speaker would be prepared to reinforce their claim with new ones, like "and that's just how it is", or by making judgments about which ends are *good* ones to have. We, at times, also say quite explicitly that things just ought to done. Asking to know "according to *which end*" the claim is to be evaluated would only puzzle the non-philosopher: They don't mean to speak relative to any particular end or ends.

I agree with Finlay that all reasons-ascriptions are end-relative, but what I have gone on to say about the particular end of being moral admits a form of vindication of categorical language. Because I have proposed a single, distinctive moral end, I have also been able to say things about distinctively moral discussion. Among other things, I have said that we enter into moral contexts in order to criticize previously relevant ends, and once moral discussion begins, the conversation resists shifting backward. I have contended that this is not merely due to the contingent fact that many of us care most about the end of being moral, but rather that this is just what we use moral language *for*: To engage in higher-order evaluations and to stake out our positions on what to do *even when antecedent contexts presuppose non-moral ends*. This opens a doorway for vindicating all the language that exhibited our folk categorical intuitions. What is required is a small revision of what exactly categoricity itself amounts to or requires.

I began with a conception of categoricity according to which a reason obtains categorically when it is a reason that can truthfully be ascribed *whatever* the contextual ends. But now the idea is to preserve what is most important about categoricity in order to make sense of the language that makes use of the concept. The language is used, as I have said, to express criticisms of potential ends: When we make a categorical claim about what ought to be done, we are tacitly criticizing ends adherence to which precludes performing the actions we prescribe. But this is the role filled by moral language. It is what is used to effect shifts in context that allow us to criticize other ends, and the shift sustains for the duration of the conversation. So, my proposal here is that we do not think about categoricity as requiring independence

from literally every end, but rather just non-moral ends.

9.3.3. First-order Intuitions

After all that has been said, the vindication of our first-order intuitions comes straight-forwardly. The moral context is determined by the shared acceptance of norms among us and our conversational counterparts, and the content of the moral context supplies external reasons available for our discussion whenever we enter into moral discussion. As a consequence, we are not only using language properly, but also saying true things when we express our community's common moral intuitions in a conversational context made up of members of that community. Because our intuitions can be truthfully expressed in this manner, they may be regarded as true by us.

Of course, conversational contexts will not always be so perfectly hospitable to us. Intuitions will often not be shared, or not shared with the perfect precision we would desire. But here the employment of new norms comes into play. During friendly moral debates, there are mutual demands for consistency among the interlocutors[15], and so by employing norms of consistency, we can help each other along by working together to coordinate our divergent lower-order accepted norms with our shared higher-order norms. So even when contextual indetermination at first threatens to undermine our ability to say true things when expressing our intuitions, it can actually move us to improve upon our intuitions *by our own lights.* Once such refinement has taken place, the appropriate moral norms will be shared, and we will be able to say true things, and it will make sense to us that expressions of our since-discarded moral beliefs could not have been truthful.

10. Moral Disagreement

Perhaps the biggest obstacle to the plausibility of a view of this kind is its treatment of moral disagreement. It must face objections to descriptivism, objections to relativism, as well as objections unique to contextualism. In doing so, it must give us an account of what exactly is going on in a variety of problematic contexts, including those where shared norms are absent and those where amoralists do not treat the contextual end of being moral, which the view suggests linguistic competence requires of them. And even though it is a revisionist view, there are limits to the flexibility revisionism grants it: When trying to motivate the view, I have relied on many observations of the present use of language, so if I have said that our moral

[15] *Ibid.*

language *does* work in a certain way, I must defend this hypothesis. If I cannot, the theory can still hold, but it will acquire more revisionist baggage, which will weaken the overall plausibility. So, in what comes next, I try to show how we can make sense of our intuitions about which sorts of linguistic behavior exhibit competence.

10.1. Contradiction in the Absence of Shared Norms

We already know a little bit about what happens when interlocutors do not have the same ends in mind when making reasons-ascriptions, but it is not clear that something similar will work when they agree that the *end* is the end of being moral but disagree morally on a fundamental level.

One of the big problems here is that it is much less appealing to say that in such a moral conversation, the disputants are simply talking past one another. In the case of disagreement about non-moral ends, it does seem that the disputants just have different things in mind, but in the moral case, we want to say that there *is* agreement about the topic of discussion: What to do *morally.* So we should not take the problem so lightly.

Fortunately, there are a number of possible solutions. I will briefly discuss one of them to stress the flexibility of contextualism as so far characterized, and will then linger on the second and, in my opinion, more promising option.

First, I will consider for moral disagreement what Finlay has said for disagreement about ends, or the refusal to conversationally accommodate by speaking relative to the relevant ends. When interlocutors are discussing which move is the best in a game of chess, the presupposed end may be something like *winning at chess.* A new interlocutor who doesn't finding winning at chess to be a worthwhile pursuit may then enter the context and begin saying things like, "So ϕing will make you lose, who cares? That's no reason". What Finlay suggests is that this claim is literally false in the context, but serves another important role: Expressing the speaker's attitudes toward the presupposed end.[16]

The analogy for moral disagreement would be to say that when shared norms are absent in a context, at least one of the speakers will end up saying things that are literally false, but they will thereby express their disavowal of the other speakers' moral standards. One problem with this approach is that it says little about some very common contexts in which there is no *dominant* moral standard: Without a dominant standard, how do we say *which* speakers' claims are literally false?

In spite of this problem, this explanation is somewhat informative. It draws our

[16]Stephen Finlay, "The Reasons that Matter". *Australasian Journal of Philosophy* 84 (1): 1-20, 2006.

attention to the non-descriptive goals of language use. Sometimes, we want to use our language to do much more than to say true things, and sometimes we do so even at the expense of uttering truth. So part of our motivation in making moral utterances is indeed to express our attitudes. I think there is no doubt about this. My remaining worry, however, is that we can say a bit more than *just* this about the disagreement. That is, we can say more than just that the whole of the linguistic activity involved in such problematic contexts is purely expressive. We can say more than just that the disagreements are ones of attitude. Specifically, we should aim to say that when people express their divergent and contradictory accepted norms in the same context, they are *contradicting* each other in a more substantial sense.

I think this can be done by shifting the burden in some way from semantics to metaphysics. We can begin by stressing the extent to which such interlocutors *do* agree: The both at least presuppose the same end, the end of being moral. The problem is that there is insufficient agreement among accepted norms for us to say what, within that context, constitutes conduciveness to this end. So, according to my theory, it would make sense to say that there is no answer to the question of who is right with respect to their disagreements. If this were described as a semantic problem, we might think that it just makes no sense to try to say something that bore unfulfilled semantic requirements. We would characterize moral claims in that context as suffering from presupposition failure, and thus being infelicitous. But this is somewhat counter-intuitive. The alternative is that we do not describe the issue as a semantic problem, but rather a metaphysical one: There is no presupposition that truth-making norms are mutually accepted, but rather just a presupposition that the end under discussion is the end of being moral. It just so happens that the metaphysical nature of morality is such that it is different in different contexts. In some contexts, we can say, it is indeterminate, and yields no definitive prescriptions. So the interlocutors' claims meet their semantic requirements, and are thus felicitous, but they are claims about something that is itself indeterminate. So, because the claims meet their semantic requirements, we can meaningfully say that they contradict without needing to say that there is a method by which we can determine which of the contradictory claims is correct.

10.2. Amoralists

There are a number of different kinds of amoralists that might each present unique problems to the theory being defended here. There are amoralists who engage in moral discourse willingly, but do not take its recommendations seriously, so that there is a lack of connection between their moral beliefs and their motivations. And

there are amoralists who simply think morality provides no reasons at all: They do not accommodate the shifting of the context to the end of being moral. I will discuss both of these amoralists in turn.

The main problem with the first kind of amoralist is that they might express agreement with the moralists with whom they interact, but they do not actually accept the same norms as their fellow interlocutors. So, intuitively, there is agreement, but what I have said so far suggests their is just contextual indeterminacy, given the absence of shared norms. In order to reconcile these hypotheses, we must either discredit the intuition that there is actually agreement, or we must say why there is not in fact indeterminacy.

We might discredit the intuition that there is actually agreement by asking ourselves whether what the amoralist expresses really fit the bill for moral beliefs. The amoralist's failure to abide by their own purported beliefs supports this line of argument. But I think this is bound to be unsatisfactory, and is not the best way of addressing this amoralist. First of all, it is in the best interest of a revisionist theory to accommodate whatever intuitions we can, and this intuition about agreement is one of them. Second of all, it is additionally counter-intuitive that moral facts would become indeterminate for a group of coordinated moralists just because a particular amoralist entered into their discussion, *especially* given that the amoralist at least superficially agrees with the moralists.

So I think it is better to show why there is not in fact indeterminacy. I propose an analogy with the fulfillment of presuppositions in conversations. Because terms like 'the' carry presuppositions (about existence and uniqueness in this case), expressions such as "the x ϕs" normally express, among other things, our beliefs about the fulfillment of those presuppositions. But sometimes, in order to communicate more effectively with each other, we simply accommodate each other's presuppositions, which amounts to treating certain things as true for the purposes of conversation. So, even though normally people only say things with presuppositions that they accept, they may at times say things with presuppositions they merely *accommodate* for the conversation.

We can treat the amoralist's relationship to the accepted norms as one like that of accommodation. Given the amoralist's motivational insensitivity to the purported moral facts, it would not make sense to say that they literally accept the norms, but we *can* say that they accommodate the use of language that, in order to be true, requires the acceptance of those norms. We can think of the shared acceptance of norms as the shared acceptance of norms *or* the accommodation of this acceptance. Thanks to this maneuver, we can say that there will remain definite answers to moral questions even in contexts where such amoralists are present.

The next kind of amoralist is somewhat trickier. The problem is that I claimed that, as a matter of fact, it is one of the features of moral language that it causes shifts in context away from any other ends and allows us to criticize those ends, but this amoralist will hear moral pleas and respond with "Who cares?" or "Just because it's wrong doesn't mean I shouldn't do it". I'm not sure we have clear intuitions about whether this amoralist's claims are true or false, but this form of contextualism also does not have any clear analyses readily available.

The first question to be addressed is whether or not this sort of amoralist's verbal behavior can actually cause moral contexts to shift to being non-moral, taking instrumental rather than moral ends. I think the answer is 'no'. Certainly the amoralist thinks that they can shift the context, but this does not make it so. And moreover, we, as moralists, should be fairly comfortable in regarding the amoralist's claims as straight-forwardly false: Morality *does* provide reasons, whether the amoralist thinks so or not.

But even once we have accepted this, we must deal with the even more difficult question of how the *moral* context is determined, given that the amoralist accepts no moral norms and refuses to accommodate others' acceptance for the purposes of the conversation. There are at least a few ways of addressing this question. None are perfect, but they should at least show that, while troublesome, amoralists present no insurmountable obstacles to this contextualism.

The first possibility is familiar: We can just say that since the relevant norms are not shared by the interlocutors, the moral context becomes indeterminate. We lose ability to say definitively who is right or wrong in the context. This solution *works*, but it is somewhat unappealing, since we might be more inclined to think that the amoralist is just wrong about what reasons there are. So we may instead say that the moral context remains determinate, but that the amoralist just gets things wrong.

But the example can of course be tweaked. We can picture amoralists outnumbering moralists in a given context, and here it is less appealing to just defer to the norm-acceptance of the minority. This move could be defended by contending that there is just something distinctive and important about the moralists: They are the ones who actually engage in moralizing, and since the end of being moral is the relevant one, their say is more important. Unfortunately, this defense seems somewhat *ad hoc*, and I think better can be done.

An alternative solution would tend more toward contextual indeterminacy, but would not force us to place arbitrary importance on moralists. We might like moralists more than amoralists, but that is no indication of their superior linguistic competence. According to this alternative solution, we can treat amoralists as *actually* accepting moral norms, even if they themselves do not believe it. Namely, by being

disposed to promote the pursuit of self-interest, even in moral context, they are perhaps expressing their acceptance of norms according to which self-interest *morally ought to be* pursued. The amoralist does not take their beliefs to be moral beliefs, but other aspects of their behavior indicate otherwise.

If this solution is adopted, we can say that what happens when an amoralist enters into others' moral discussions is that the amoralist is treated simply as another kind of moralist: The distinction evaporates. Because the amoralist's self-interested language in fact plays the role that I claimed was distinctive of moral language, it can itself be treated as moral language, even if it lacks other intuitive hallmarks of moral language. With the evaporation of the distinction comes the disappearance of any distinctive theoretical problems caused by this kind of amoralist. Contexts are determined in the ordinary way: By the shared acceptance of norms. Where amoralists dominate, the norms they share may determine truth. In reversed scenarios, we may regard their reasons-ascriptions as false. In mixed contexts, we may just say there is no fact of the matter about who is right.

11. Conclusion

For all that has been said, much is quite tentative, and there remains an extensive number of questions unanswered. However, I think there is great promise in the skeleton of a view that could leave so much of our first-order moralizing untouched. The true success of it is perhaps in its irrelevance to non-philosophers: If it is accurate, lay people can rest easy in their current discourse and practice. Nonetheless, the view cannot get off the ground without being backed by some form of revisionism. Even though it vindicates our first-order intuitions, it bears a heavy burden in explaining how our conceptions about reasons and other concepts could be so inaccurate, as well as why this does not matter. These issues are deep, wrought with contentiousness and uncertainty at many levels, and I have only shown what can be done if we allow that our conceptions can be erroneous. Still, this way of doing philosophy should itself be regarded with promise once we see how much it can accomplish, such as the wholesale vindication of going moral discourse and practice.

So, my hope is that I have both begun to motivate a particular form of metaethical contextualism and to present considerations in favor of the methodology that carved out the path to this theory.