Building bridges with words: an inferential account of ethical univocity

Mark Douglas Warren

Published online: 09 Jan 2018.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Building bridges with words: an inferential account of ethical univocity

Mark Douglas Warren

Department of Philosophy and Religious Studies, Daemen College, Amherst, NY, USA

ABSTRACT
Explaining genuine moral disagreement is a challenge for metaethical theories. For expressivists, this challenge comes from the plausibility of agents making seemingly univocal claims while expressing incongruent conative attitudes. I argue that metaethical inferentialism – a deflationary cousin to expressivism, which locates meaning in the inferential import of our moral assertions rather than the attitudes they express – offers a unique solution to this problem. Because inferentialism doesn't locate the source of moral disagreements in a clash between attitudes, but instead in conflicts between the inferential import of ethical assertions, the traditional problem for expressivism can be avoided. After considering two forms of inferentialism that lead to revenge versions of the problem, I conclude by recommending that we understand the semantics of moral disagreements pragmatically: the source of univocity does not come from moral or semantic facts waiting to be described, but instead from the needs that ethical and semantic discourses answer – a solution to the problems of what we are to do and how we are to talk about it.

1. Introduction

Moral disagreement is, we think, ubiquitous. We find it in the seminar room, at the dinner table, and pretty much everywhere on our social media feeds. But metaethicists have trouble explaining these disagreements. When fundamentally disparate sorts of considerations drive our moral judgments, how is it that we are to understand moral disagreements *as* moral disagreements, instead of just instances of people talking past one another? How can we account for the sense that such disagreements are genuine? What's needed is an account of the meaning of our moral language that establishes a 'semantic common ground'

(Merli 2007b, 26) for interlocutors – if we don't mean the same thing with the words we're using, we can't have a genuine disagreement. Metaethicists need an account of the *univocity* of moral language.[1]

Consider the problem as it appears for moral expressivism. Expressivists maintain that moral utterances express non-cognitive states. Our moral judgments are more like desires than beliefs, they argue, so expressivists can understand univocity in terms of the potential harmony or disharmony between the attitudes being expressed by disputants. But because they model disagreement in terms of a kind of conative clash, the viability of their account of univocity hinges on the identification of the particular kind of conative attitudes all moral claims function to express. This is problematic; we can imagine having genuine disagreements with 'atypical' agents who make moral claims without being in the appropriate mental state.

In this paper, I set out to argue for a metaethical position that maintains the principal appeal of expressivism – what I will call its deflationary advantages – but does so without running into this problem. What's needed is a deflationary approach that does not account for the meaning of ethical claims as a function of the attitudes they express. Such an approach is available: metaethical inferentialism.

I proceed as follows: In Section 2, I explain the appeal of deflationary approaches, both in metaethics and elsewhere. In Section 3, I explain how the possibility of conative variation makes univocity problematic for expressivists. In Section 4, I offer inferentialism as a deflationary alternative to expressivism, and argue that because this position doesn't explain the meaning of moral claims in terms of the mental states they express, it offers us a tempting way to avoid problems associated with expressive accounts of univocity. In Section 5, I consider two ways of explicating the link between use and meaning: *dispositional* and *regulatory*. I argue that both face a problem of inferential variation that echoes the expressivists' difficulty with conative variation. In Section 6, I argue that these problems stem from an inflationary assumption about meaning talk itself. I suggest a deflationary alternative: *normativism* about meaning. This approach is uniquely well equipped to deal with the dangers that variation poses for giving a univocal treatment of moral communication.

2. Background: the appeal of deflationism

A traditional realist approach to metaethics has a seemingly straightforward account of genuine disagreement: when interlocutors have a moral disagreement, that's because there's some moral property that they're both referring to.

Adam: Eating meat is wrong.

Amy: No, it's not wrong at all.

Adam thinks the property of wrongness attaches to the act of eating meat; Amy doesn't. So long as we can give an account of how our moral predicates 'lock on' to the same properties, we can explain why this is a genuine disagreement. Now, it turns out that it's actually a bit tricky to account for this mutual locking on – explaining how diverse speakers are connected to the same moral properties even when they have fundamental disagreements about their extensions.[2] I'll ignore these issues here; for my purposes, it will suffice to note that traditional moral realism faces a more basic problem, one that stems from its essentially *inflationary* approach.

An inflationary metaethics is any one in which moral properties or facts play an ineliminable role in the explanation of moral discourse. The idea is that our moral expressions aim to represent moral facts, and that these facts act as *truthmakers* for our moral claims or judgments. Accordingly, this truthmaker would play a crucial role in the inflationist's account of Adam and Amy's disagreement – each of them is making a claim which will turn out to be true only if it accurately represents the moral nature of eating meat. But positing moral facts in this way brings with it a host of well-known metaphysical and epistemic worries. The challenge for the inflationist lies in figuring out just what sort of things moral facts are supposed to be, and placing them within a scientifically respectable worldview.[3] Whether she can fulfill this metaphysical desideratum or not, the inflationist also has to contend with epistemic and practical questions: How is it that we hairless, language-using apes come to learn about these moral truthmakers, and how do these moral facts come to influence how we decide to behave?

A tempting response to these worries is to go deflationary. Instead of asking, 'What is the nature of moral properties?', the deflationist asks, 'What are we *doing* when we make moral claims like this?' If we can answer the latter question, we might find that a satisfying explanation of moral thought and practice that doesn't hinge on the answer to the former question. In contemporary metaethics, the dominant answer to this question comes from expressivists, who hold that what we're doing, fundamentally, is expressing conative attitudes – roughly speaking, we are booing or hooraying the eating of meat. The expressivist foregoes an explanation of the nature of moral properties or facts, and instead gives an account of the practical significance of moral assertions.

I call expressivism 'deflationary' to highlight its connection to deflationism in other domains. There are many of these – in ethics, but also in epistemology, modality, and theories of truth. The hallmark of such approaches is the turn away from inflationary questions about *truthmakers* – about the facts or properties that are ontologically required for an assertion or thought to be true. Deflationists instead step back to ask the *pragmatic* question: What purpose is served by thinking and talking about these things – about morality, knowledge, necessity, truth? So, a deflationary approach to epistemology might look past questions about the necessary and sufficient conditions for knowledge itself,

and instead focus on how knowledge claims help us keep track of who is reliable and why.[4] Modal deflationists don't attempt to understand modal claims in terms of truthmakers like possible worlds, but instead favor explanations of how our talk about necessity and possibility helps to make explicit the conditions under which particular concepts can be applied.[5] Deflationists about truth deny that there is some interesting property that all true statements must share; they instead explain our use of the truth predicate as a 'disquotational device', one that frictionlessly moves us from the assertion p to the assertion 'p is true'.[6] In each of these deflationary accounts, the basic approach is to prefer a *functional* account of how we use certain types of expressions over a *representational* account that demands an investigation into the truthmakers for those expressions. The hope is that a proper understanding of this function will give a satisfyingly natural account of the discourse in question and the part it plays in our lives, but will also deflate the epistemic and metaphysical worries that bedevil inflationary approaches.

On an expressive account, moral claims get their meaning by expressing evaluative attitudes, and these attitudes are more akin to desires than beliefs. Expressivists argue that because moral language functions primarily to express mental states – and not to report on a domain of moral facts – we can give a naturalistically respectable explanation of moral discourse that doesn't appeal to any problematic moral truthmakers. This is the appeal of deflationism, in metaethics and elsewhere: by focusing on the practical matters that drive a discourse, we won't get mired in questions about what sorts of properties or facts the claims in such a discourse are supposed to reference.

3. Genuine disagreement for expressivists

The expressive account of moral disagreement can be read off of two central theses:

> (*Mentalism*):   the meaning of moral claims comes from the mental states that these claims express

> (*Non-cognitivism*) :   unlike beliefs, these states do not primarily function to describe some moral aspect of the world, but instead, like desires, they play some sort of non-cognitive or conative role.

To keep its deflationary credentials, an expressivist account of disagreement cannot start from the truth-making assumption that there is some property an act has (or lacks) that disputants disagree about. Instead, an expressivist explanation of disagreement must work through an account of the attitudes moral claims function to express. Expressivists understand disagreement in terms of the expression of non-cognitive mental states that have a kind of incompatible practical significance. Interlocutors disagree when they are expressing conative states that systematically lead to behavior that puts them at odds with one another.

We might, for example, follow Stevenson (1937) in modeling ethical disagreements on clashes between our interests, instead of our beliefs:

> Disagreement in interest occurs when *A* has a favorable interest in *X*, when *B* has an unfavorable one in it, and when neither is content to let the other's interest remain unchanged … *A*. "Let's go to a cinema to-night." *B*. "I don't want to do that. Let's go to the symphony." *A* continues to insist on the cinema, *B* on the symphony. This is disagreement in a perfectly conventional sense. They can't agree on where they want to go, and each is trying to redirect the other's interest. (Stevenson 1937, 27)

Moral claims conflict in the same way – because they express attitudes that clash in their practical implications for our behavior. We read their content as contradictory as a result of this clash. We might think of these mental states as if they were jigsaw puzzle pieces from the same box; they 'fit' with one another because they systematically share practical implications for how we behave.

Given that the notion of mental states being expressed does so much work for expressivists, we must ask: Exactly what sort of conative state are we supposed to be expressing when we make moral claims? It won't do to simply assert that moral claims express interests, as Stevenson seems to suggest; when I say, 'The Star Wars prequels were awful', I'm expressing a kind of interest (in never watching them again, say), but for all that I'm not making a moral claim. It also won't do to just contend that moral assertions express *moral* interests. This may be true, but it illuminates nothing about what sets moral interests apart from other kinds of interest, like prudential or aesthetic interests.

It turns out that this challenge – the Moral Attitude Problem, or MAP (Miller 2003, 43–51) – is a major obstacle for expressive accounts of morality. David Merli (2007b) points out that the expressivist response must meet two criteria: it must give an account of the attitude in question that is specific enough to differentiate it from other kinds of normative claims (aesthetic or prudential, say), but also be general enough that it doesn't render moral disagreements that we intuitively take to be genuine as instances of equivocation. We see that Stevenson's suggestion fails the former criterion; expressivists have to be sure that the account they give is specific enough that it differentiates moral attitudes from non-moral ones. But they must also meet the latter criterion; their account of the moral attitude must be general enough to make sense of the robust diversity of participants in genuine moral disagreements – we don't want an account of mental states that rule out disputants that our commonsense intuitions would rule in as participants to a real disagreement.

The danger here is that the more exact the expressivist's specification is, the more open she becomes to counterexamples. For example, Allan Gibbard analyzes moral attitudes as endorsements or rejections of the rationality of feeling guilt or resentment towards an act (1990, 42). So, in our example above, Adam endorses the rationality of feeling this way towards eating meat, and Amy rejects it. Plausibly, if he's in favor of that emotional reaction and she's against it, they're expressing mental states that could lead to incompatible behavior in a

way that's in line with the general expressive account of genuine disagreement. But what if guilt isn't at the heart of Amy's judgment? What if what she's really expressing with her claim is a rejection of the rationality of feeling disgust at those who eat meat? What if she's from a culture where judgments about the appropriateness of shame are fundamental to morality? Or maybe she's one of those dreaded amoralists – someone whose moral judgments are motivationally inert because they don't express any particular conative attitude.

The worry generalizes. For *any* proposed attitude, it seems we can imagine engaging in what appears to be a genuine ethical dispute with someone who doesn't in fact have that attitude. This may be because our interlocutor is an amoralist, or because she is an iconoclast whose moral claims don't reflect common sensibilities, or because she is suffering from *akrasia* or cynicism and so doesn't have access to the relevant dispositions, or it can be because she comes from a culture whose members characteristically express significantly different attitudes with their moral assertions. This is the problem of 'conative variation' for expressivists. Above, I suggested that expressivists can think of disagreement in terms of the mental states being expressed by interlocutors having the appropriate practical 'fit', like puzzle pieces from the same box. The problem of conative variation forces us to consider situations in which interlocutors are drawing their tiles from different boxes, and so lay down pieces that don't have the right kind of fit. The implication is that such engagements are equivocal: because our interlocutor is expressing a different sort of mental state than we are, he is – according to (*Mentalism*) – making claims with different meanings than ours, and 'in order to engage in real agreement or disagreement, we must mean the same thing by our terms; otherwise we are simply talking past one another' (Merli 2007b, 26). If the meaning of moral claims comes from the attitude they express, the question of whether or not interlocutors engaged in an apparently genuine disagreement are in fact talking past one another is entirely hostage to *both of them having the right kinds of attitudes*. This suggests there is a tension between our common-sense intuitions about moral univocity and whatever theoretical commitments an expressivist might make in specifying the attitude moral claims function to express.

The MAP generalizes, too. Metaethicists of any stripe must deal with the potential conflicts between (1) their theoretical commitments, whatever they may be, regarding the foundation of meaning for moral claims, and (2) our common-sense judgments about ethical univocity, which seem to systematically outstrip these commitments. David Merli (2007a, 2007b, 2009) points out that this gives us a way to undermine accounts of genuine disagreement across the metaethical spectrum; let us call it *Merli's Strategy*:

> Once a theory puts its chips on one particular aspect [essential to all] moral judgment, it must struggle to account for cases of what appear to be moral judgments without that favored characteristic. Our views about what counts as a real moral claim and what does not, and about what counts as genuine moral dispute, seem

to be sensitive to a cluster of features, any one of which can be removed without destroying the crucial intuitions. This leaves us with a sort of counterexample-generating algorithm: once a theory provides a delineation of what is really central to moral judgment, we can go to work constructing problematic cases. (Merli 2007b, 28)

The danger that the MAP presents for the expressivist is clear and compelling. In order to give a plausible, informative account of moral language, expressivists need to specify the mental states that moral judgments express. This has to be detailed enough to show why these judgments are unique, but the more specific the account is, the more exposed it is to problems of conative variation. The takeaway is that this leaves the expressivist vulnerable to Merli's Strategy: our intuitions about whether or not someone should be counted as making moral assertions are orthogonal to our theoretical considerations about what attitudes they may be expressing.

4. An inferential approach

In this paper, I want to explore a different deflationary approach, one that promises to short circuit the problem by abandoning the expressivist's theoretical commitment to conative attitudes. If we deny that the meaning of a moral claim comes from the mental state it expresses, it looks like the problem won't get any traction. We won't have any obligation to specify a particular attitude that undergirds all moral discourse, because univocity won't depend on universal expression of such an attitude.

We can do this by applying Robert Brandom's general inferential account of meaning to moral claims. The inferentialist shares with the expressivist her deflationary commitments, but instead of adopting an expressivist account of the meaning of moral claims, she explains the meaning of such expressions in terms of their location within a network of inferential relations.[7] Roughly speaking, the meaning of a claim is a function of what inferences would justify its assertion, and also of what further claims we should assent to if we accept it.

We can use a well-worn metaphor here, and compare the inferential norms that are meaning-constitutive for moral claims to the general norms governing the play of a game like chess. Chess pieces and chess moves only count as such against the background of general rules for the game. What transforms a piece of wood into a knight, say, are the moves it can make, and also its part within a game that is defined by certain norms. We define the game in terms of these norms, and identify pieces by the moves that are permitted within the context of these norms. If you want to know what a knight is, you have to understand the rules of chess.

With this metaphor in hand, we can think of the meanings of moral assertions in terms of the rules governing moves in a game. The meaning of moral claims will be identified by the import these 'moves' have in conversation, against a

background understanding of the broader norms for moral discourse itself. So, to understand what a moral assertion means, we need to think about the network of inferences it's caught up in – the implications and entailments, consistencies and inconsistencies that it's connected to. Think again of the claim that eating meat is wrong. We identify the meaning of these words by looking at the rules that govern their use – just the same way we identify the chess pieces.

Adam's assertion: *'Eating meat is wrong.'*	
Rules	Sample inferences
A person's moral assertions reveal commitments, sometimes about general moral principles we hold …	*We should avoid harming sentient creatures*
… and sometimes about the ways that those principles intersect with conclusions we've reached about matters of fact	*Animals are sentient creatures*
These commitments may be prompted by multiple lines of evidence …	*Anyway, eating meat is bad for the environment*
… and, when expressed, can themselves prompt us to worry about hitherto unconsidered moral possibilities	*Maybe I should become a vegan*
Finally, and essentially, our moral convictions aren't just beliefs we form about the way the world is – they lead us to form resolutions to act in a different way, to be a different sort of person, and to form new reactions (such as resentment, punishment, or praise) to others' behavior. Adam's assertion has implications for what he's going to do next	*I'm going to start making my friends feel really guilty during meals!*

You can see that these inferences neatly divide into two categories: What justifies Adam's assertion, and what sort of other beliefs and commitments it justifies (indicated above in the dark background). Following Brandom (1994), we can understand these relations as the 'upstream' inputs that entitle us to make a claim, and the 'downstream' outputs that such a claim licenses, respectively. Again, one of the essential rules of playing this language game is that moral assertions have direct implications for what we decide to do. Upstream, we make inferences from our principles and beliefs; downstream, moral assertions can lead to new types of behavior.

Our ability to keep track of the inferential import of one another's claims only makes sense, Brandom argues, within the socially-embedded practice of asking for and giving reasons for our assertions. This is the language game we play together that enables meaningful discourse:

> In asserting a claim one not only authorizes further assertions, but commits oneself to vindicate the original claim, showing that one is entitled to make it. Failure to defend one's entitlement to an assertion voids its social significance as inferential warrant for further assertions. It is only assertions one is entitled to make that can serve to entitle others to its inferential consequences. (Brandom 1983, 641)

Taken as a metaethical proposal, inferentialism urges us to understand the meaning of our moral claims by locating their role within an interlocking structure of linguistic responsibility and license. We undertake a commitment to answer challenges to our assertions, to give reasons that justify them, and if

these challenges are met (or are never issued), we enjoy entitlement to make further inferences from these assertions, and to share license for these further inferences with others. The meaning of a moral expression falls out of its place within such a practice.

As a deflationary approach, inferentialism denies that we need to cite moral properties or relations in order to explain the meaning of moral assertions. Michael Williams' (2010) meta-theoretical explanation of meaning in deflationary theories gives us a framework to explain the meaning of expressions without adverting to their truthmakers. To give such an account, he argues, we must first look to the functional roles characteristically played by the terms of a discourse – for moral discourse, we consider terms such as 'ought', 'wrong', 'good', 'bad', and so on. The inferential norms that govern our use of such terms will be those that enable the terms to fulfill these functional roles. And these roles in turn are best understood in terms of the overall pragmatic aim that moral discourse serves in our lives. This gives us an order of explanation for the meaning of moral terms that eschews both the notion of correspondence to truthmakers and the expressivist emphasis on mental states: The meaning of a moral term like 'bad' is given by the inferential norms that govern it; these norms make sense in light of the broader pragmatic roles that are particular to our moral thought and discourse.

A thorough exploration of the function that moral discourse has in human life lies outside of the scope of this paper, but the most likely explanation of the usefulness of moral language will tie it to the role it plays in facilitating cooperative social behavior. For a highly social, language-using species there is an obvious advantage to a discourse that regulates uncooperative behavior: it promotes our collective chances for evolutionary success.

> By providing a framework within which both one's own actions and others' actions may be evaluated, moral judgments can act as a kind of "common currency" for collective negotiation and decision-making. Moral judgment thus can function as a kind of social glue, bonding individuals together in a shared justificatory structure and providing a tool for solving many group coordination problems. (Joyce 2006, 117)

Moral discourse offers us a potential counterpoint to the selfish impulses that would drive individuals to uncooperative behavior in situations when cooperation would be mutually beneficial. It also gives us a platform for negotiating social behavior by appealing to agreed-upon principles.

Once we've shown that the inferential roles associated with moral expressions arise from the pragmatic role they play (rather than the truthmakers they represent) no metaphysical or epistemic accounting is necessary. This means that a deflationary inferentialism shares the same advantages expressivism enjoys over inflationary metaethics. There is a *prima facie* case for favoring this inferential conception of meaning over the expressivist's, though, owing to its relative simplicity. As we saw above, to give an adequate account of why some moral assertions express incompatible content, an expressivist must first

explain what sorts of mental states these assertions express, and then go on to explain how it is that such attitudes ground a kind of pragmatic clash that gives rise to incompatibility. The inferentialist loses the attitude, and so can take a more direct route, simply accounting for the inferences that moral assertions license without taking an explanatory detour through the mental states being expressed. Indeed, much of the recent literature criticizing expressivism has focused on the tangled knot of suppositions and commitments regarding the nature of moral attitudes that the expressivist must untie in order to account for various semantic relationships between moral claims.[8] By taking an inferential approach, we can cut cleanly through this knot, because inferentialists don't need to show anything about attitudes being expressed or their relationship to meaning and inferential import.

Inferentialism doesn't locate the source of moral disagreements in a clash between attitudes, but instead in conflicts between the inferential import of the moral claims in question. Inferentialists can explain the same kind of disagreement-grounding pragmatic clash that expressivists locate in the conflict between non-cognitive attitudes, by noting the behavior-guiding aspects of moral inferences themselves. Amy thinks eating meat is permissible; Adam argues it's wrong. They have a genuine disagreement because the inferences they're committed to endorse types of behavior that are incompatible.[9] Adam is against eating meat; Amy is okay with it.[10] This is the source of their disagreement.

Because inferentialists can account for univocity without referring to particular attitudes, the MAP doesn't get traction. If we can gesture convincingly at the sorts of inferential norms that serve the function of moral discourse, we can then go on to explain univocity in terms of sharing these norms: interlocutors are engaged in genuine disputes just when the moral assertions they make share the appropriate inferential properties; their terms share meanings just when they're bound by the same inferential rules.

5. The problem for inferential accounts

At first glance, then, we might think that the inferentialist's stance towards the MAP should be a lot like the atheist's stance towards the Problem of Evil: in both cases, the problem dissolves because neither the atheist nor the inferentialist posits the entities that raised the issue in the first place. But we're not out of the woods yet. Inferentialists argue that the meaning of moral expressions is ultimately explained by our rules of use for them, so just as the expressivist owes an account of the mental states expressed by moral assertions, the inferentialist has to give an account of the inferential rules that govern these assertions. In the following sections, I'll consider three different conceptions of these rules: *dispositional*, *regulatory*,[11] and *normative*. The dispositional conception of inferential rules, we'll see, leads to an inferential version of the MAP. The regulatory conception faces a dilemma: it either has the same problem as dispositionalism

or it relies on a naturalistically suspect notion of the norms governing use. In the penultimate section of this paper I will argue that the final approach, normativism, gives a uniquely satisfying response to worries about inferential variation.

5.1. Dispositional

Dispositional forms of inferentialism hold that it is either speakers' actual use of terms or their dispositions to employ terms in certain ways that establish the meaning-constitutive rules that govern our use of these concepts. Ned Block, for example, understands the inferential role of an expression as its causal role

> in reasoning and deliberation and, in general, in the way the expression combines and interacts with other expressions so as to mediate between sensory inputs and behavioral outputs. (Block 1986, 93)

An explanation of the meaning of an assertion is therefore reducible here to natural facts about an agent's psychology – her disposition to make certain inferences. The rules that govern inferences are rules of thought in an individual's head.

The problem with this is that our judgments of what qualifies as a moral claim are not exhausted by the circumstances and consequences a speaker might endorse for her assertion.[12] Something can qualify as a moral claim even if it doesn't comport with a particular kind of dispositional pattern. Take for example the case of Thrasymachus,[13] who in his dialogue with Socrates argues that justice is no great moral good – that indeed the unjust are to be admired. Thrasymachus defines justice as the advantage of the stronger. If Socrates and Thrasymachus are evoking fundamentally different inferential dispositions when they use the term 'justice', and the meaning of the term comes from the inferences it is connected to, this seems to entail that Socrates and Thrasymachus are not in fact having a dispute, since they're using (the Greek word for) 'justice' to mean different things.

This is a variation on Merli's Strategy, at the level of inferential rule instead of attitude. The MAP for expressivists works because we recognize the plausibility of genuine disagreements between interlocutors who express different sorts of attitudes. If we read the rules of inference off of the behavioral significance individuals attach to moral expressions, and if it is these rules that fix the meanings of our moral assertions, then inasmuch as different speakers use their terms with different inferential commitments, they mean different things with their words. Our judgments of what qualifies as a moral claim are not exhausted by our considerations of the sorts of upstream and downstream inferential implications we considered in Section 4. Something can qualify as a moral claim even if it doesn't comport with a particular kind of inferential pattern. The MAP has returned in a new guise, this time as a worry about variance not in conative attitude but in inferential significance.

5.2. Regulatory

On a regulatory account, the meaning of our expressions is given not by the inferences we are disposed to make, but by the inferential rules which govern our claims. Assertions mean the same thing because they are regulated by the same set of rules. Adopting this account of inferential meaning would undermine worries about eccentric speakers like Thrasymachus – or the amoralist – who deploy moral language without the disposition to make the normal sorts of inferences. What qualifies their use of moral expressions as univocal with our own is a mutual liability: even if we do not make the same inferences from our claims, we *should*.

A regulatory conception hold that claims of meaning are justified by appeal to the meaning-constitutive norms that regulate our use of moral expressions, instead of extant use-regularities or dispositions. Consider again the metaphor from chess: engaging in a norm-governed practice like chess is a matter of being answerable to these rules. Someone can count as playing chess even if she isn't disposed to follow all of its rules – it's enough that she's held liable to the norms of a public rule-governed practice. A player who does not know how to capture *en passant* plays badly, but plays nonetheless. Likewise, someone can count as using moral language even if he doesn't have all of the appropriate inferential dispositions. So, when the amoralist says something like, 'Murder is wrong', we take his assertion to mean the same thing as it would coming from our own mouths, even though he says it without the right kind of attendant behavioral dispositions. This is because the publically held standards of correct use for that assertion include (among many other things) a defeasible motivation to avoid murder.[14]

But this approach to meaning is also problematic. On this account, our assertions mean the same thing because they are governed by the same rules. But how are we to conceive of these rules? Where do they come from? The regulist faces a dilemma: these norms are either constituted in some way by a community's inferential patterns, or they aren't. The latter option resurrects the problem of variation; the former undermines the inferentialist's deflationary advantages.

We can highlight the problems with these options if we consider the possibility of intercultural or diachronic disagreements. Consider for example the following regrettable passage:

> If a man has sexual relations with a man as one does with a woman, both of them have done what is detestable. They are to be put to death; their blood will be on their own heads. (Leviticus 20:13, New International Version)

We take ourselves to be in disagreement with this verdict. But how does the regulatory view account for the shared meaning that is a precondition for genuine disagreement? When faced with eccentric individuals like Thrasymachus or the amoralist, the regulist can defer to community-wide inferential norms that such individuals are violating. But it's quite plausible in this case that the Levitical

author is not using idiosyncratic inferential patterns; his moral judgments are a reflection of the inferential standards of his community.

Here is the first horn of the dilemma, then. The moral expressions of the Levitical community answer inferential patterns that are different from our own. (For example, these patterns include justifications for the claim that we would not accept, e.g. that a vengeful God abhors homosexuality.) This is the same problem Thrasymachus presented, but pitched now at the communal instead of the individual level. Worryingly, it looks like the regulatory view would undermine our conviction that we actually disagree with this passage: after all, both we and the Levitical author are using moral expressions that are governed by different inferential rules, and in contrast to the case of the merely eccentric speaker, we can't easily explain this difference in terms of some common set of publically held norms to which we're all obviously liable. Merli's Strategy once again rears its ugly head.

On the other horn of the dilemma, we might conceive of these rules *transcendentally*, in such a way that the inferential rules come apart entirely from how people actually use (or used) language.[15] The regulatory view works by connecting meaning to the inferential rules to which we are liable. So, for it to render disputes like this one univocal, it needs to be the case that even though our own actual inferential patterns do not match up with those of the Levitical community, it's still the case that (somehow) they are both governed by the same set of inferential rules. But in appealing to rules that transcend linguistic dispositions, we're in danger of violating the naturalism that motivated our deflationary metaethics in the first place. The problems that arise with this conception of meaning echo the issues that plague realist accounts of moral truths: how are we to understand the truthmakers for these rules? Are they Platonic or natural? If the latter, what kinds of natural facts (if not use-regularities) constitute these normative facts? If the former, how is it that we language users come to enjoy the access to these rules that are necessary for meaningful communication? And how would we know that the same rules apply to the Levitical community that apply to our own?

6. Normativism: a deflationary account of meaning

Up to this point, we've been considering accounts of meaning that share an inflationary assumption: that there are some features of linguistic reality that, whatever they are, our meaning claims aim to represent. The idea is that if '*X* means the same as *Y*' is true, there's some fact about the world, some truthmaker in virtue of which the claim comes out true – whether these are facts about mental states, about the dispositions of individuals or communities, or facts about norms that somehow exist separately from those dispositions.

As we've seen with expressivism and the varieties of inferentialism proposed so far, theoretical considerations about the nature of these semantic truthmakers

puts us in a position of tension with our common-sense intuitions about what qualifies as a genuine moral disagreement. If we want to account for meaning in a way that grants an immunity to Merli's Strategy, we should change our tactics. Instead of first committing ourselves to a theoretical meaning-giving feature of moral discourse, and then checking for counterexamples that would defeat this conception, we should instead ask ourselves what is happening when we make judgements of univocity in the first place.

To proceed in this way is to make the same deflationary turn toward meaning that the expressivist took towards morality, and that we saw in Section 2 has also been pursued by deflationist approaches to epistemology, modality, and theories of truth.

> Just as a deflationary theory of truth denies that there is any deep and substantive answer to the question of what the property of truth consists in … deflationism about meaning denies that there is any special (non-semantic) property of *meaning F* to uncover the nature of, enabling us to say in general what meaning F consists in. (Thomasson 2015b, 189)

This shouldn't be taken as denial that there are facts about meaning, though. A deflationary approach to meaning does not entail that there are no facts of the matter about whether or not a given disagreement is univocal. Here again, the normativist can take inspiration from the expressivist, and marry her account to deflationary conceptions of truths and facts. The idea is that, on a deflationary reading, one can move from the sentence, '*X* means *Y*', to the sentence, 'It is true that *X* means *Y*', or 'It is a fact that *X* means *Y*', without adding any real content to the original claim. If we can make sense of the first of these claims, we can make sense of the last two also. The normativist (like the expressivist) can thereby make claims to semantic (or moral) truths without relying on semantic (or moral) truthmakers. Again, the basic deflationary approach is to prefer a *functional* account over a *representational* one. Instead of asking about the nature of these truthmakers, we should ask what we're doing when we make these sorts of assertions.

This is what the inferential normativist does for questions of meaning. According to normativism, 'semantic claims do not talk about 'normative entities' attached to expressions, they *prescribe* how to handle the expression in the proper way' (Peregrin 2012, 97). So, if we judge that some distant community uses a term with the same meaning as one of our own, we're not thereby describing some regularity or disposition or set of transcendent rules for that community. What we're doing fundamentally is not making a *description*, but instead a *prescription* – when we say two expressions, *X* and *Y*, mean the same thing, we are licensing the inferences attendant upon *X* to *Y* (and vice versa). This is because 'meaning talk is primarily used to provide normative guidance for inferential behavior' (Lance and Hawthorne 2008, 138).

> In translating, we are not … trying to describe a foreign community as it is in itself. Rather, we are trying to form one large community where previously there

were two. We are trying to make communication, discussion, and argumentation possible; … we are trying to make possible the sort of cognitive openness that is largely taken for granted in any unified speech community. (Lance and Hawthorne 2008, 63–64)

Perhaps chess gives us a misleading metaphor for how rules of use relate to meaning. After all, the norms governing chess aren't metaphysically mysterious. They are written in rulebooks. There's no difficulty explain how these norms inform actual play, either; it is simply a matter of understanding the kinds of sociological facts revealed in games of chess. But there is no rulebook for a language. Unlike chess play, our judgments of synonymy are sensitive to practical considerations, the sometimes-dynamic ways that our use of words connects to our ways of living.

If we need a game metaphor, we should think instead of a pickup game of beach soccer.[16] There is no rulebook for such a game. Our play itself informs the rules to which we hold ourselves and other players liable. If the beach players endeavor to state these rules, they will be attempting to precisify, adjudicate, or just make explicit what is already implicit in their practice. We count as players in virtue of being sensitive to the rules, but these rules evolve dynamically, and are apt to assimilate particular features of that day's game: have we been worried about offsides so far? Should we count that rock as out-of-bounds, or is it too close to the shoreline? This is a better metaphor for the relationship between meaning and rules of use; the meanings of expressions sometimes assimilate certain aspects of the evolution of the discourse in which the expressions appear.

The normativist and the regulist agree that our meanings are answerable to norms, but because normativism takes a deflationary approach to meaning itself, we needn't hold out for an account of the metaphysical grounds of these norms. Instead of asking what the norms are, we ask how we use them. So, for example, our understanding of how to appropriately use morally loaded terms like 'person' and 'fairness' is answerable to the conclusions that have come from an evolving moral discourse, e.g. about who should be counted as a person and why. We need an account of meaning that does justice to the way in which norms are liable to particular features of use. Such an account doesn't cash out in terms of truthmakers for meaning claims, but instead in terms of a sensitivity to the shifting pragmatic implications of our judgments of synonymy. 'A meaning-claim gains its normative authority … from the difference it would make to subsequent practice to respect and enforce the norms it expresses' (Rouse 2014, 35).

So, let's think again about the implications of a univocal treatment in moral discourse.

Adam: Eating meat is wrong.

Amy: <No, it's not wrong at all.>

Adam makes a moral assertion, with its attendant upstream and downstream implication, some of which I explicated above. Amy makes her own claim, and this claim carries its own network of implications. Now let the angle brackets indicate that she speaks a foreign language. Adam must consider whether to treat theirs as a genuine disagreement. If he does so, he is bringing those inferential networks into contact with one another, by endorsing a uniform treatment of the entitlements and commitments that come with their assertions. He is in effect inviting Amy into the language game he is playing. This then is an opportunity for both of them to demonstrate that their commitments to the respective claims were justified. In moral discourse, challenges like this can often be valuable platforms for the exchange of information and reasoning – and maybe for a cooperative exploration of the principles and beliefs that brought them to disagree in the first place.

Such a translation isn't just the identification of identical patterns of use; *it is the building of a communicative bridge*. It is a way to invite interlocutors into a mutual realm of persuasion, argumentation, and cooperation. This also shows what we do when we judge that our moral expressions are *not* univocal. We deny ourselves a platform for communication. We cut ourselves off from negotiating with one another and restrict the tools we can use to influence behavior – without moral discourse, we must resort to bribes, trickery, threats, or the outright use of force.

If Adam and Amy keep this bridge open, though, they maintain the possibility of persuasion. Perhaps he convinces her that his original assertion was justified, and so as a result, Amy takes on a belief that has a similar inferential profile to Adam's. This will have implications not just for what she comes to believe, but also on what she decides to do next. As we've seen, this is one of the special hallmarks of moral language. One of the things she might decide to do is spread the word to other parties, because moral discourse is a game we all play together: offering our thoughts on what's right and wrong, justifying, challenging, questioning, and adjusting these thoughts – building and partaking in the discourse that comes out of this practice. Again, all of this ethical discussion and thinking isn't just about changing and reaffirming our beliefs, but also essentially in deciding about how we're going to behave. In the moral language game, ideas and arguments can propagate into the zeitgeist and provoke important social movements. Moral discourse is a coordination device for social creatures, for whom cooperative behavior is crucial for survival. For Adam and Amy, as with all of us, the payoff of treating their disagreement as genuine is potentially huge. Recall Williams' suggestion that the norms governing a discourse are responsive to the function that discourse plays in our lives. If moral discourse functions to facilitate the negotiation of our behavior, as I've suggested, we can now see why it makes sense that our judgements of ethical univocity are so robust.[17]

None of this entails that the view I'm advocating will always recommend judgments of univocity, though. Perhaps we can take a cue from Merli, and try to

imagine situations that *wouldn't* recommend a univocal treatment. If we imagine a moral-ish discourse stripped of one or some of the rules I associated with our moral assertions in Section 4, we might find our intuitions pulled in different directions.[18] So, for example, we could learn of some community that regularly engages in a type of discourse that like our moral discourse, informs practical decision-making and prompts critical reflection, but unlike our discourse, does so a way that never (tacitly or otherwise) involved any appeal to overarching principles. We might wonder whether or not the thought and practice that has grown out of their discourse should be counted as moral thought and practice. This would amount to wondering whether or not we should be reading some of the assertions within that alien discourse as having the same inferential import as some of our moral assertions – to wondering if treating our assertions as inferentially uniform is worth the effort. But the takeaway from normativism about meaning is that there is no Archimedean point outside of these sorts of considerations to look for universal principles for synonymy.

At this point, the reader may worry that I've conflated two questions: 'Under what conditions are disagreements actually univocal?' and 'Under what conditions does it make sense for us to treat disagreements as univocal?'[19] I begin this paper by promising a response to the former question, the complaint goes, but instead only deliver a response to the latter. It's important to note the sense in which this worry is question-begging. It assumes that the only adequate explanation of sameness of meaning is inflationary – that we should only be satisfied when we've been given a general account of what makes it the case that two claims to mean the same thing. But this is just the assumption being challenged by a normativist approach, which argues that once we've given an account of when it makes sense to treat an argument as a genuine disagreement, there's nothing left to explain.

The payoff of this approach is that it gives us a response to the challenge posed by Merli's Strategy. This is because the challenge began by taking a given specification of the attitudes or rules that ground moral meaning, and then searches for counterexamples wherein the given specification cannot account for our intuitions of univocity. Taking a deflationary approach to meaning itself shields the normativist from this worry. Any time the normativist is asked 'Do *X* and *Y* mean the same thing, for these people, in this context?' she will not need to read the answer off from her theoretical commitments regarding truthmakers for meaning claims – since she doesn't have any. She instead attends to whatever might be pragmatically salient to the question. When we judge that a moral interlocutor means the same thing we do with his words, this is not because we've discovered that he's answering the same inferential norms that we are, but because we've identified his discursive practice in terms of its success or failure as an alternative approach towards the same goal our moral language has: the management of coordinated group behavior for social and selfish creatures. Because we see our expressions as aiming at the same discursive goal, we see

a reason to build a bridge between the inferences to which we are both liable. This reverses the inflationary order of explanation: Instead of positing inferential rules as theoretical meaning-giving entities to which our meaning claims aim to correspond – and then testing them against our intuitions about univocity – we recognize that our meaning claims and the inferential rules they reveal are a way of *making explicit* the practical judgements that inform these intuitions.[20] Merli's Strategy threatens other theories of meaning with potential counterexamples; this strategy is defanged by a normative approach, which transforms the threat into an invitation to imaginatively explore what matters to us.

Metaethical inferentialism shares the basic deflationary advantages of expressivism. We needn't posit the kind of relationship between representation and truthmaker that inflationary approaches do. For the metaethicist, that means we don't have to explain how our moral assertions lock on to some aspect of the world, and we don't have to give any account of moral properties. With the deflationary approaches, all of that goes away. If you want to understand how moral thought and discourse functions, the interesting work isn't to investigate moral truthmakers and how they fit in with all the other natural properties that science posits. The important thing is to investigate us, we hairless apes, and how we use language to figure out how to live together.

Normativist inferentialism in metaethics has the same kind of deflationary payoff for questions of meaning and genuine disagreement. This approach to meaning has the added advantage of undermining Merli's Strategy. The very intuitions that drive judgments of sameness of meaning (i.e. intuitions about whether or not a disagreement is a potentially fruitful opportunity for working out issues of how to live together) are the ones that make sense of moral discourse's functional role in our lives. Claims of synonymy endorse decisions to tie ourselves to common inferential rules for terms. These endorsements are normatively governed – not, as the regulist would have it, by reference to existent communal norms, or to norms that float free of patterns of use – but by the same pragmatic considerations that inform our intuitions about moral univocity. *We take disputants to be using the same terms when we see the negotiated assimilation of inferential commitment as one that is conducive to decisions about how to live together*.

7. Conclusion

My aim in this paper was to offer a solution to the problem of univocity for metaethical deflationists. This problem arose for expressivists once we recognized that their theoretical commitments regarding ethical meaning-makers cut orthogonally to our intuitions about genuine moral disagreements. I've pointed a way out of this problem: first, we recognize that an expressive account of meaning isn't our only option as deflationists. Instead of identifying meaning with the expression of a distinct moral attitude, an inferential approach sees the

meaning of an expression in terms of its place within a network of inferential patterns. Metaethical inferentialism sidesteps the MAP because the semantic account it offers doesn't depend on a specification of which attitudes are singularly moral.

I then considered the possibility of a revenge version of the problem for inferentialists. At its strongest, the question is: how are we to account for our intuitions of univocity between denizens of cultures where the use of moral terms seems to be governed by substantially different inferential rules? We can avoid this problem too, by taking a deflationary approach to meaning claims themselves. Normative inferentialists understand questions of synonymy as questions about whether or not the same inferential norms should govern a linguistic form, where this normative question can only be answered by reflecting on the pragmatic needs that would be served by an affirmative answer.

This approach gives us a new way to think about moral univocity. It accounts for the felt intuitions we have about genuine moral disagreements, and does so without positing any inflationary moral or semantic facts that our different discourses are somehow tracking or failing to track. We count disputes as genuine, not because there is some (moral or semantic) truthmaker to which the disputants' assertions are answerable, but because doing so licenses us to hold them liable to inferences that are conducive to moral decision-making. This perspective urges us to understand the semantics of moral disagreements pragmatically: the source of univocity does not come from moral facts waiting to be described (or, as the expressivist would have it, moral attitudes being expressed), but instead from the needs that ethical and semantic discourses answer – solutions to the problems of what we are to do and how we are to talk about it.

Notes

1. I will treat these two problems – explaining sameness of meaning and explaining genuine disagreement – interchangeably. For an alternate approach, see Plunkett and Sundell (2013).
2. See Merli (2007a).
3. Cf. Price (2011).
4. See Chrisman (2011) and Field (2009).
5. See Thomasson (2007).
6. See Horwich (1999).
7. See Bar-Bar-On and Chrisman (2009), Chrisman (2010).
8. See for example Schroeder (2008), Dreier (2009).
9. Of course, not all moral disagreements directly settle questions about what to do. Disagreements in such cases may not be about what we should do immediately, but might instead be about the principles that inform what we would do. See Warren (2013, ch. 5) for an inferential account of disagreement in such situations.
10. Accounting for the incompatibility between Adam's judgement that eating meat is forbidden and Amy's judgement that it is permissible is a noted difficulty for expressivists. See Warren (2015, especially pp. 2877–2879) for an argument that inferentialism is well-situated to solve this problem.

11. Brandom (1994) uses the terms 'regularist' and 'regulist' for these conceptions, respectively.
12. Here I repurpose David Merli's (2009, 540–545) arguments against Ralph Wedgwood's Conceptual Role Semantics. Though Wedgwood's metaethics is explicitly inflationary, Merli's arguments have equal force for a deflationary, dispositional version of metaethical inferentialism.
13. See Sturgeon (1986, 115–142).
14. This is a nice potential upshot of the inferential approach: a rather tidy resolution of the internalism/externalism debate about moral motivation. See Chrisman (2010, esp. 118–119).
15. See Lance and Hawthorne (2008, 186–187).
16. See Lance and Hawthorne (2008, 218), Thomasson (2015a, 250–251).
17. Similar points are made by Tersman (2006, Ch. 6) and Bjornsson and McPherson (2014), who argue that it makes good evolutionary sense that we allow for a lot of latitude between divergent interlocutors in our judgments of moral univocity.
18. Cf. Bjornsson and McPherson (2014, 9–10).
19. Thanks to an anonymous reviewer for pressing this point.
20. In the conclusion of his (2007b), Merli gestures at a possible avenue to solve the MAP:

> We might try to preserve our intuitive ascriptions [of moral univocity] by rejecting views that make participation in moral discourse hinge on any one member of the cluster of features that affect the attribution of moral concepts … According to one way of developing a view of this sort, facts about whether speakers share meanings are constituted, not tracked, by our best interpretations. In other words, our interpretive norms are fundamental. Hence there is no possibility of a gap between facts about what speakers mean and how they are best interpreted. (Merli 2007b, 54)

> Setting aside the worry that Merli is calling for our best interpretations to be taken as truthmakers for moral claims, I believe much of what I argue for here is consistent with this approach.

Acknowledgments

Notes on contributor

Mark Douglas Warren is assistant professor of philosophy at Daemen College, Amherst, New York. His research is in metaethics with a focus on inferentialism.

References

Bar-Bar-On, D., and M. Chrisman. 2009. "Ethical Neo-Expressivism." In Vol. 4 of *Oxford Studies in Metaethics*, edited by Russ Shafer-Landau, 133–165. Oxford: Oxford University Press.

Bjornsson, G., and T. McPherson. 2014. "Moral Attitudes for Non-Cognitivists: Solving the Specification Problem." *Mind* 123: 1–38.

Block, N. 1986. "Advertisement for a Semantics for Psychology." *Midwest Studies in Philosophy* 10: 615–678. Reprinted in Stephen P. Stich and Ted A. Warfield, eds., *Mental Representation: A Reader*, Oxford: Blackwell, 1994.

Brandom, R. 1983, November. "Asserting." *Noûs* 17 (4): 637–650.

Brandom, R. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Chrisman, M. 2010. "Expressivism, Inferentialism, and the Theory of Meaning." In *New Waves in Metaethics*, edited by Michael Brady, 103–125. Palgrave-Macmillan.

Chrisman, M. 2011. "Is Epistemic Expressivism Incompatible with Inquiry?" *Philosophical Studies* 159 (3): 323–339.

Dreier, J. 2009. "Relativism (and Expressivism) and the Problem of Disagreement." *Philosophical Perspectives* 23 (1): 79–110.

Field, H. 2009. "Epistemology Without Metaphysics." *Philosophical Studies* 143 (2): 249–290.

Gibbard, A. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.

Horwich, P. 1999. *Truth*. 2nd ed. New York: Oxford University Press.

Joyce, R. 2006. *The Evolution of Morality*. Cambridge, MA: MIT Press.

Lance, M. N., and J. Hawthorne. 2008. *The Grammar of Meaning: Normativity and Semantic Discourse*. Cambridge: Cambridge University Press.

Merli, D. 2007a. "Moral Convergence and the Univocity Problem." *American Philosophical Quarterly* 44 (4): 287–313.

Merli, D. 2007b. "Expressivism and the Limits of Moral Disagreement." *The Journal of Ethics* 12 (1): 25–55.

Merli, D. 2009. "Possessing Moral Concepts." *Philosophia* 37 (3): 535–556.

Miller, A. 2003. *Introduction to Contemporary Metaethics*. Cambridge: Polity.

Peregrin, J. 2012. "Inferentialism and the Normativity of Meaning." *Philosophia* 40: 75–97.

Plunkett, D., and T. Sundell. 2013. "Disagreement and the Semantics of Normative and Evaluative Terms." *Philosopher's Imprint* 13 (23): 1–37.

Price, H. 2011. *Naturalism Without Mirrors*. New York: Oxford University Press.

Rouse, J. 2014. "Temporal Externalism and the Normativity of Linguistic Practice." *Journal of the Philosophy of History* 8: 20–38.

Schroeder, M. 2008. *Being For*. Oxford; New York: Oxford University Press.

Stevenson, C. L. 1937. "The Emotive Meaning of Ethical Terms." *Mind* XLVI (181): 14–31.

Sturgeon, N. 1986. "What Difference does it Make if Moral Realism is True?" *The Southern Journal of Philosophy* 24 (Supplemental Issue): 115–141, cited in Merli (2007): 32–33.

Tersman, F. 2006. *Moral Disagreement*. Cambridge: Cambridge University Press.

Thomasson, A. L. 2007. *Ordinary Objects*. New York: Oxford University Press.

Thomasson, A. L. 2015a. *Ontology Made Easy*. New York: Oxford University Press.

Thomasson, A. L. 2015b. "Deflationism in Semantics and Metaphysics." In *Metasemantics*, edited by Alexis Burgess and Brett Sherman, 185–213. Oxford: Oxford University Press.

Warren, M. D. 2013. *Lightweight Moral Realism: Objectivity and Reasoning without Heavyweight Facts*. Open Access Dissertations. Paper 1138. http://scholarlyrepository.miami.edu/oa_dissertations/1138.

Warren, M. D. 2015. "Moral inferentialism and the Frege–Geach Problem." *Philosophical Studies* 172 (11): 2859–2885.

Williams, M. 2010. "Pragmatism, Minimalism, Expressivism." *International Journal of Philosophical Studies* 18: 317–330.