

Convergence in Plan

Moral judgments, Gibbard tells us, are plan-laden – fraught with ought, they implicate the directive, planning, side of our psychology as well as its prosaic, representational, side. Following Stevenson, Gibbard has emphasized that such plan-laden judgments put us in disagreements with one another that are no less profound than disagreements in purely prosaic belief. To this, many would add that these claims also allow for an explanation not only of the *possibility* of moral disagreements, but of their *pervasiveness*. If moral judgments are plan-laden, it can seem to be no wonder that Cleopatra and Antony differ in their moral judgments. Each’s judgments reflect plans – *choices* – about how to respond to possible situations, and the difference in their plans reflects only the fact that many choices are possible.

Gibbard’s expressivism, however, does not concern moral judgments only. It is a general claim about normative judgments of all kinds. Genuinely normative judgments, Gibbard holds, are all plan-laden in this way – this is the distinctive mark of the normative. And among such normative judgments, are judgments of linguistic meaning. If the plan-ladenness of moral judgments is what explains the depth and persistence of moral disagreements, then we might expect to see a similar depth and persistence of disagreement in each normative domain. But this is not obviously so. Indeed, it is obviously not so. Though there are disagreements about meaning, for example – and whether expressivism is true is certainly among them – this is against a backdrop of a great deal of consensus.¹

In this paper I want to explore what can be said, from an expressivist perspective much like Gibbard’s, about such consensus in normative outlook – about convergence in plan. What I will be looking for, are the sorts of factors that might lead us to expect a greater degree of convergence in plan, and evidence as to whether these factors are present in greater numbers in the case of judgments of linguistic meaning than in the case of moral judgments. I will distinguish between three main kinds of engine of convergence – high-octane, medium-octane, and low-octane. High-octane engines of convergence guarantee perfect convergence

¹ Nicholas Laskowski [ms] makes a similar observation about the normativity of speech acts in his APA author-meets-critics comments on Cuneo [2014], which he touches on in Laskowski [2017]. This paper is an extended attempt to work out the implications of Laskowski’s remark.

over some domain, at least among rational and reflective thinkers. When these are present, there is a guarantee given meaning that every thinker is rationally committed to taking the same view, or at least to not denying it. Low-octane engines of convergence are driven by empirical or historical assumptions – convergence over some domain could be driven simply by psycho-social or etiological factors. And medium-octane engines of convergence aspire to something middling. They appeal to intrinsic features of some planning questions in order to explain why some answers are more natural than others.

I High-Octane Convergence

Some normative convergence is pervasive and rationally compelling. According to Gibbard, for example, every rational thinker is committed to accepting the claim that the normative supervenes on the natural, and similarly, to accepting the claim that there is some natural property that constitutes what it is to be wrong. The arguments that Gibbard gives for these claims are, as I will put it, high-octane. They put convergence over these claims in the same category as convergence over truths of logic. Given the meanings of all of the terms involved in these claims, every thinker is rationally committed to the accepting them, unless they fail to have the relevant concepts at all.

High-octane explanations of convergence are a particularly essential tool in any normative expressivist’s toolkit. Broadly speaking, what high-octane explanations of convergence do, is to establish a kind of *analyticity* for certain claims. But we need to be careful about exactly what this means. Normally, a claim is understood to be analytic just in case the meanings of its terms guarantee it to be true. But Tappenden [1992] has taught us how to relax this conception: we might take analyticity to be the status that a sentence has when the meanings of its terms guarantee it not to be false.

But expressivist meanings are not truth-conditional in nature; they do not guarantee any claims to be true or even not to be false – at least not directly. So what expressivist meanings are capable of guaranteeing directly is that some claims are ones that everyone (or perhaps, everyone who understands their meanings) is rationally committed to *accepting*, or – on the analogue of Tappenden’s generalization – at least rationally committed to not denying.² These expressivist analogues of analyticity play the right role for analyticity, within an expressivist view – they establish analytic claims as ones that can be taken for granted (on the first formulation), or at least as undeniable (on the more relaxed formulation).

Expressivist-friendly analyticity, as I have said, is an essential tool for expressivists. Not only can it be used to characterize the status of supervenience and the natural constitution of normative properties, as

² Compare Schroeder [2010b], which I apply this expressivist conception of analyticity to the paradox of the liar.

Gibbard explores in *Thinking How to Live*, but it can be used to characterize the validity of arguments. The claim that it is impossible for any argument of the form *modus ponens* to have true premises but an untrue conclusion can be shown to satisfy the expressivist analogue of analyticity.³ With care, we could extend Gibbard's system to prove that every thinker is committed to accepting this claim, and similarly, in the system of bifurcated attitude semantics, developed in my *Being For*, there is a relatively straightforward proof that it is rationally undeniable.⁴ But once we have this result on board, since we are ourselves rational thinkers, we may take the transcendental turn. Since every rational thinker is committed to accepting this claim (or at least, to not denying it), we should accept it, too, on pain of irrationality. And since we do, let us assert it: it is impossible for any argument of the form *modus ponens* to have true premises but an untrue conclusion. That is how we earn the right not just to some expressivist substitute for validity, as noncognitivists have sought at least since the criticisms of Ross [1938], but to the real thing.

But even before we take the transcendental turn and assert, as theorists, the claims about validity, supervenience, and property constitution, our standing to make these claims, for the expressivist, is rooted in the high-octane explanation of why we should expect convergence. We should expect convergence over these matters for the same reasons that we should expect convergence over matters of logic or other analyticities – because anyone who disagrees is either making some mistake about meanings, or making some mistake about following through on their own rational commitments. So it is no wonder that we observe a striking asymmetry between the great divergence amongst those who accept utilitarianism and those who do not, and the great convergence over normative supervenience. The latter is not incompatible with the idea that expressivism helps to explain the existence of deep disagreements – rather, it is just what you would expect, given that some claims are bound to be analytic, and supervenience is (according to many, at least⁵) plausibly one of them.

2 Low-Octane Convergence

High-octane explanations of convergence explain part of the contrast between topics of great normative disagreement and topics of great consensus. But they only explain part of it. There are many matters of significant consensus that do not lend themselves at all to such high-octane explanations.

³ Compare Schroeder [2010a], chapter 10.

⁴ See Schroeder [forthcoming].

⁵ It is worth pointing out that I myself do not think that supervenience is analytic, since that is not always obvious to readers of my [2005], [2007], [2014b], or Schmitt and Schroeder [2012].

For example, for most of human history, it was a matter of great consensus that duties to people far away are less stringent than duties to the nearby. This is hardly the sort of thing that we would expect to be analytic – indeed, it is almost certainly false. Similarly, there is great consensus that it is wrong to torture someone for fun. But though in contrast to the former claim, this is almost certainly true, it seems like an important *substantive* truth. People who deny this claim aren't linguistically mistaken – they are *evil*. So the claim that it is wrong to torture people for fun isn't the right kind of thing to be an analytic truth – and so it's not the right kind of thing to be subject to a high-octane explanation of convergence.

Fortunately, low-octane explanations of convergence can come to the rescue. Though there is no linguistic confusion or failure to follow through on their own rational commitments that is exhibited by people who deny that it is wrong to torture for fun or that duties to people who are far away are less stringent than duties to people who are nearby, it is no surprise that such people are rare. For moral views are constituted by patterns of norm-acceptance (or by plans for what to do in a range of counterfactual circumstances), and because which norms we accept (or which plans we adopt) have strong implications for how we act, moral views are subject to powerful evolutionary forces. Altruistic behavior toward neighbors received ample evolutionary payoffs under the historical conditions of human evolution in small hunter-gatherer societies, but not so altruistic behavior toward the distant needy, and so it is no wonder that stronger moral attitudes toward helping the nearby were selected for, without selecting such strong attitudes toward helping the distant needy. Similarly, torture for fun is a paradigmatically non-cooperative activity with no direct evolutionary payoff. So the evolutionary payoffs of cooperation would naturally select against it. Obviously, these explanations could be tightened considerably, and I am merely gesturing toward how such explanations might go.

Low-octane explanations of moral convergence may also be cultural. Widespread consensus could be the result of influential films or novels, of popular trends, or even of the bare fact of unfamiliarity with certain ways of life. Low-octane explanations of moral convergence may be historical, etiological, psychological, or sociological. What they have in common, is that they offer contingent explanations which aspire to explain something less than high-octane explanations of moral convergence.

Appreciating the diversity of the range of possible low-octane explanations of moral convergence is important, in order to temper our sense of what it is reasonable to expect about patterns of disagreement in some domain, given an expressivist account of that domain. Low-octane explanations show why there may be a wide variety of explanations for why not everything seems to be up for grabs, just because the meanings of the terms in some domain do not settle what views it is rational to hold in that domain.

But as our examples of low-octane explanations of convergence clearly illustrate, low-octane explanations have severely limited power. It is at least as plausible that evolutionary considerations explain why throughout human history, the view that our duties to those nearby are more stringent than our duties to the distant needy has been nearly universal as it is that they explain why human history has been dominated by the view that it is wrong to torture for fun. But to many of those of us who reflect on the matter in the twenty-first century, when there are easy ways of helping people in all corners of the globe, it is far from obvious that distance could possibly affect the stringency of our moral obligations.⁶ The fact that we now understand why many people have thought otherwise does nothing at all to cast the common view in a favorable light, let alone to make it seem compelling or natural, in any way. It explains without rationalizing.

I take away from this example the observation that low-octane explanations of convergence in normative outlook are weaker than high-octane explanations along more than one dimension. They are less powerful in that they typically explain less convergence – whereas high-octane explanations of convergence can explain convergence among every rational and reflective thinker, low-octane explanations of convergence will typically only have the right structure to explain tendencies or predominant patterns. But more strikingly, low-octane explanations of convergence provide less than high-octane explanations, because their explanations give us causes without rationalizations. Understanding a low-octane explanation for convergence will never make the converged-on view seem compelling or even appealing, in its own right.

3 Expressivism, Meaning, and Judgment Internalism

Let us return, then, to our test case that motivated this inquiry – the normativity of meaning. To say that meaning is normative is to say that meanings are “fraught with ought,” and given the expressivist treatment of the normative, that means that they are in Gibbard’s terms plan-laden. To hold that a word has a certain linguistic meaning is to plan for how to use it, or for what standards to hold others to, for its use.

Gibbard’s [2012] primary motivation for accepting the normativity of meaning thesis is as an answer to the problem of the underdetermination of meaning, as explored particularly extensively by Kripke [1982]. The core of this problem is that meanings are infinite in the distinctions that they make, but the meaning-constituting facts are finite. For every pattern that extends the totality of past finite patterns of use out into the future, there are infinitely many alternatives to that pattern that respect the totality of past use equally well. This is no problem at all, Gibbard contends, if meaning is something that we bring to the world, rather than something that we find there.

⁶ Compare especially Singer [1972].

Once we adopt expressivism about meaning, moreover, our account of meaning turns in on itself. Since expressivism is itself a claim about meaning, that means that expressivism itself is a thesis fraught with ought. Endorsing expressivism, either about meaning or about morality, therefore, consists in accepting a plan for what to do with the words ‘means’ or ‘ought’, or what standards to hold others to, in their use of these terms.

The fact that expressivism turns in on itself in this way could potentially have striking implications for Gibbard’s original metaethical expressivism, in turn. Gibbard’s original formulation of metaethical expressivism, in *Wise Choices, Apt Feelings*, was committed to a particularly strong form of judgment internalism – on that view, it is literally impossible to think that stealing is wrong without being in a norm-acceptance state that would motivate you not to steal. And sincere speakers will always be in the mental states that their assertions express, so sincere speakers who assert, ‘stealing is wrong’ will always be in such a norm-acceptance state. But given that expressivism itself is a plan-laden thesis, the disagreement between cognitivists and expressivists is itself a plan-laden one. So speakers who plan to use moral language in the cognitivist way may be *sincere* in their assertion of ‘stealing is wrong’, in that they are in the mental state which they themselves *take* it to express, without being in any state of mind with any intrinsically motivating properties. Such speakers would behave much as supposed counterexamples to judgment internalism are alleged to behave.

Of course, those of us who endorse metaethical expressivism will still say that these speakers are no counterexample to the thesis that everyone who *understands* the meaning of ‘stealing is wrong’ and who asserts it sincerely will be in a planning or norm-acceptance state that would motivate them not to steal. And so we will still get the letter of Gibbard’s original, very strong, version of judgment internalism. But the spirit of this claim how has the potential to be substantially watered down by the fact that claims about what it takes to understand what a sentence means, at least in general, are now substantive, planning, questions.

So while we still get to say that the cognitivist who is unmotivated by her moral judgments is no counterexample to judgment internalism, because she does not understand the meanings of moral words, we *may* now also allow that it betrays no misunderstanding of how to use the term ‘means’, to come to her conclusion about the meanings of moral terms rather than ours. This disagreement is, in some sense, to be expected, precisely *because* it is a kind of disagreement in plan.⁷ So our criticism of the character who sincerely asserts, ‘stealing is wrong’ but has no motivation not to steal is now on a par with our criticism of the character

⁷ The disagreement is particularly to be expected, if the dispute between expressivism and cognitivism is itself one of the cases in which patterns of use underdetermine meaning – a claim which I think Gibbard himself rejects, but which I think could be accepted along with most of his other commitments. Thanks to Billy Dunaway for discussion.

who believes that our obligations to the nearby are more stringent than our obligations to the distant needy – it is grounded in a substantive planning error.

My own view is that this is one of the subtlest and most wonderful possible upshots of Gibbard's arguments in *Meaning and Normativity* – a striking virtue that has the potential to arise when Gibbard's views about meaning are put together with his views about moral language, so long as we allow (which Gibbard may not) that the dispute between cognitivists and expressivists is itself one of the questions left open for planning. Although metaethical expressivism like the variety that Gibbard advanced in his earlier work, particularly in *Wise Choices, Apt Feelings*, offers a powerful account of the practical force of moral language, one glaring worry that we should always have had about it is that its account is *too* powerful, because it is committed to a stronger claim about what this practical force amounts to than can plausibly be defended, given what we know about the vast range of actual motivational upshots of moral judgments among the diverse range of thinkers and speakers in the real world.

The idea that metaethical expressivism is itself a plan tempers this thesis, without strictly weakening it. It still comes out as true (according to the proponent of both expressivism about meaning and metaethical expressivism) that anyone who understands the meaning of 'stealing is wrong' and sincerely asserts it will be in a planning state that has the right structure to motivate them not to steal, but there will be speakers who exhibit no misunderstanding whatsoever of the meaning-determining facts who can sincerely assert 'stealing is wrong' *with its usual meaning* but have no motivation whatsoever not to steal. This *is* a concession to motivational externalists, but in contrast to other concessions – such as Michael Smith's [1994] suggestion that the proper formulation of judgment internalism includes a restriction to agents who are 'practically rational' – it does not water down judgment internalism too much for it to still bear weight in supporting metaethical expressivism. So it is the right kind of concession to make, for an expressivist who seeks to use judgment internalism in its traditional role of helping to motivate metaethical expressivism. It strengthens the hand of the metaethical expressivist, by heading off the worry that he is committed to an unacceptably strong form of judgment internalism.

4 Convergence in Plans for Meaning

As I said, in my view the implication that I have just been discussing of the normativity of meaning thesis for the defensibility of metaethical expressivism is one of the most striking and powerful upshots of Gibbard's normativity of meaning thesis. It is just one example of how powerfully the more general perspective of the

normativity of meaning can lead us to re-think how we understand the commitments of metaethical expressivism, without undermining those commitments.

But now I want to make one important observation about how we get this upshot: we get it because we were assuming that because the question of whether cognitivism or expressivism is true for moral terms is a planning question, both answers are in some sense optional, and hence to be expected. This is just on a par with the assumption that since the question of whether utilitarianism or deontology is true is just a planning question, both answers are in some sense optional, and hence to be expected. In both cases, we get the expectation of disagreements from the diagnosis of the underlying nature of the issue as a planning question.

I've been arguing that it can be a *virtue* of Gibbard's package of commitments that it opens up the possibility of allowing that the disagreement between cognitivists and expressivists is reasonable. But even if we don't say this – even if we conclude that cognitivists somehow get things more deeply wrong than either deontologists or consequentialists, and so this disagreement is not, fundamentally, reasonable, we must still recognize that it *exists*. It is an example of a deep and substantial disagreement about meanings. Like all disagreements about meanings, there is some trivial sense in which whichever party is incorrect does not even understand the meanings of the words at stake in the debate, but this sense is trivial, and does not obscure the fact that this disagreement really exists.

But that brings the puzzle that I am pursuing in this paper into full view – *some* disagreements about meaning do exist, and are quite arguably intelligible and natural, including the disagreement between expressivism and cognitivism about moral claims. But *many* such disagreements are not at all intelligible or natural. The view that 'steal' means to give a gift is just a plan about how to use the word 'steal', but we don't observe widespread disagreement about whether 'steal' means to give a gift; on the contrary, this view is not merely false, but absurd. So the challenge for the advocate of expressivism about meaning is to explain how the planning nature of meaning judgments could create space for disagreement in a very restricted range of cases, at the same time as we observe that the vast majority of possible plans about what words mean seem to be not just false, but totally absurd, and endorsed by no one, actual or imaginable.

So far, we have seen two models for explanations of why we might observe a convergence in plan, despite the fact that something is itself a planning question. But neither of these models does quite what we should want of an account of the vast range of agreement that we see over linguistic meaning. On the high-octane model, we get explanations of convergence because some planning questions turn out to have analytic answers – answers to which every planning agent is implicitly committed, or at least committed not to denying. But the claim that 'steal' does not mean to give a gift does not seem like the right kind of thing to

be analytic – if this is not obvious, it helps to remind yourself that this sentence only mentions, but does not use, the word ‘steal’.

In contrast, on the low-octane model, all that we get are causal explanations of patterns among speakers’ plans. And as we’ve seen, these kinds of causal explanations don’t *rationalize* the views that they explain. But the claim that ‘steal’ does not mean to give a gift is not just common, or even just universal – it is *compelling* to anyone acquainted with the use-facts surrounding ‘steal’. And it is compelling in a way that ought to be graspable by anyone who understands how meaning works.

We need, therefore, a third model for explanations of normative convergence – one that can hope to explain the strong rational pressures toward convergence about meaning and other topics of intense and pervasive normative agreement, without subsuming this agreement to the analytic. We need to see what a medium-octane explanation of convergence might look like.

5 Medium-Octane Convergence

What we need from a medium-octane explanation of convergence is that it rationalize, and not merely explain, convergence in plan. The more forcefully it rationalizes this convergence, the stronger its claim to be able to make good on the prediction that every rational and reflective thinker will have plans that satisfy the relevant constraint, and hence to predict extensive actual convergence, given minimal assumptions about real thinkers’ actual levels of rationality. The answer, I take it, is that some plans are better than others. Not better merely in the sense expressed by a second-order plan to make some plans rather than others, but better in some unavoidably recognizable way – that can be appreciated by anyone, no matter what else they plan.

This is a high standard. In order to be bad in a way that can be appreciated by anyone, no matter what else they plan, a plan would essentially have to be self-frustrating, or at least conditionally self-frustrating, given pretty minimal conditions. That is what I will now argue medium-octane explanations of convergence in plan can provide.

Consider the case of tic tac toe. Some plans about how to play tic tac toe are better than others. There is no mystery about why there is so much convergence among plans about how to play tic tac toe, for the point of these plans is to solve some problem – how to *win* or at least *avoid losing* at tic tac toe. And given the rules of the game, some plans are straightforwardly dominated by others. Finally, the way in which some of these plans dominate others with respect to this goal is easily discovered and widely known. So it is no wonder that these are the strategies on which everyone converges, and it is no wonder that these strategies seem rationally compelling to anyone who shares the goal of winning or at least not losing at tic tac toe.

Not every move at tic tac toe is mandated by the winning strategy. The first move of the second player, for example, is fairly unconstrained by the goal of not losing. Most of us, therefore, adopt a plan that is indifferent about what move to make at this stage of the game, but restrictive about what move to make at later moves of the game. But another conceivable plan is to always mark the top left corner at this stage of the game. This plan is no more frustrated by the goal of not losing than the more permissive plan. And similarly for the plan to always mark the bottom right box at this stage of the game. But though there are *several* possible plans for how to play tic tac toe that are not dominated, conditional on the goal of not losing, the vast majority of such plans are ruled out. They are guaranteed to do worse than some other strategy at the goal of not losing.

So if hypotheses about linguistic meaning are like strategies for how to communicate, as strategies in tic tac toe are strategies for how to win or at least not lose, then the background facts about patterns of use that according to Gibbard, following Kripke's Wittgenstein, underdetermine linguistic meaning, could set sharp constraints on the success of these strategies. If your goal is to successfully communicate with your audience, then it takes only minimal observation of the pattern of use of 'steal' in order to observe that successful communication will be difficult, if you use it to mean 'to give a gift'.

Using 'steal' to mean 'to give a gift' is not intrinsically doomed to failure – it could succeed in an environment very different from ours – and even in our environment, it works just fine if successful communication is not one of your goals. But given minimal background facts about other speakers of English, it straightforwardly fails at the goal of successful communication, and it does so in a way that is relatively transparent – obvious enough to even minimally reflective thinkers that it should be no wonder that any thinker who formulates views about linguistic meaning as part of a plan to communicate will not plan in this way.

6 Optional Meanings

Meaning, Kripke's Wittgenstein reminds us, is infinite, but use facts are finite. The use facts underdetermine how to go on well after our use facts have died out, and even an intention to go on in one way rather than another helps only if something makes it the case that one's intention has one content rather than another. Hypotheses about linguistic meaning that diverge over the infinity of cases underdetermined by the use facts, therefore, are predictably going to be unconstrained by the use facts with respect to the goal of successful communication. No hypothesis about linguistic meaning that diverges only over this range of cases can get

one into trouble, given the goal of successful communication. And so these alternative hypotheses might each be a reasonable plan about how to use language.

This result makes good on the idea that the use facts underdetermine the meaning. After all, if there was only a *unique* plan that fit with the use facts, facts about linguistic meaning could simply be facts about that plan, and so meaning would not be underdetermined by use, after all. But this kind of result can also be extended, in order to make sense of evolving uses of language over time. For example, it is now clear that ‘water’ is a natural kind term, which picks out the same chemical kind in every possible world. But it is not obvious that this was always the case. The discovery of the chemical constitution of water could have been a semantic choice point – before which none of the use facts determined how to go on in using the word ‘water’ under the multifarious possibilities of philosophers’ imaginations. If that was so, then before the advent of modern chemistry, multiple plans for how to use ‘water’ could have been consistent, given ordinary speakers’ knowledge of the use facts, with the goal of successful communication. But even if this is right about the past, we have now crossed a semantic choice point, and our own plans for how to use ‘water’ must correspond to a broader set of constraints. This is just what we should expect, if use constrains meaning but underdetermines it – more use could more closely constrain meaning over time, yielding less and less scope for reasonable disagreements about meaning.⁸

This view can make sense of why the vast majority of views about linguistic meaning, like the vast majority of strategies for playing tic tac toe, are not just rare, but unreasonable, because they frustrate the goal of linguistic communication, which is the point of meaning things by our words. And it can do so while leaving open a range of permissible plans about what to mean – by and large, plans that go beyond the range of the existing use facts. So what, then, of the dispute between cognitivist and expressivist theories of the meaning of normative terms?

Earlier I endorsed the idea that both cognitivism and expressivism may be reasonable plans about the linguistic meaning of normative terms as a way for Gibbard to soften the commitments of the form of judgment internalism that underlies his norm-expressivism. If, even while endorsing the expressivist plan for what to mean, we allow that the cognitivist plan is another reasonable plan, then we can allow that although every agent who understands the meaning of ‘ought’ is motivated in accordance with their ‘ought’ judgments, among these are agents who fail to understand what ‘ought’ means only because they endorse a different plan for what to mean with it. This makes much more intelligible how these agents could fail to be motivated by

⁸ By this of course I don’t mean that use or dispositions to use are a hard constraint – merely that pressures toward success in plans about what to mean come from dispositions and patterns of use, and so if there are more developed dispositions and patterns of use, they will create more pressures on our plans for what to mean.

their moral judgments, making the resulting form of judgment internalism more palatable. And this explanation requires the assumption that the disagreement between cognitivists and expressivists is itself reasonable.

I'm not yet sure, however, on the form of the medium-octane convergence strategy being considered here, whether the disagreement between cognitivists and expressivists is the right sort of thing to be stably reasonable. It contrasts with the case of the plus/quus distinction and the case of possible past semantic indecision about whether 'water' was a natural-kind term in that it applies to cases that have already been considered – indeed, with which we are all, as speakers, highly familiar. So for the disagreement between cognitivists and expressivists to be stably reasonable, both patterns of use need, intuitively, to be persistent and common, even in the face of the fact that we recognize these features of each other's use. Much more would need to be said, in order to achieve clarity about why this continues to be a stably reasonable thing to disagree in plan about, even though it is not reasonable to disagree with someone who uses 'Mary' to refer to a different person than you do about to whom it refers.

So I'm not sure that my solution, in the form of a medium-octane strategy for predicting convergence, gets us everything that I would have liked to have, on behalf of the view in the neighborhood of Gibbard's that I would most have liked to have been able to defend. But perhaps I have tried to get too much, and the moral is that we should let go of the idea about using expressivism about meaning in order to soften the force of the kind of judgment internalism to which Gibbard is committed. But the point remains, I think, that the space of strategies is rich, for expressivists to make sense of a wide variety of differences in the space of reasonable disagreement.⁹

References

- Cuneo, Terence [2014]. *Speech and Morality: On the Metaethical Implications of Speaking*. Oxford: Oxford University Press.
- Gibbard, Allan [1990]. *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press.
- _____ [2003]. *Thinking How to Live*. Cambridge: Harvard University Press.
- _____ [2008]. *Reconciling Our Aims: In Search of Bases for Ethics*. Oxford: Oxford University Press.
- _____ [2012]. *Meaning and Normativity*. Oxford: Oxford University Press.

⁹ Special thanks to Nicholas Laskowski for the idea to pursue this line of thought, to Billy Dunaway and David Plunkett, for the opportunity to pursue it and for enormously helpful comments, and especially to Allan Gibbard, for giving us all so much to think about.

- Kripke, Saul [1982]. *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press.
- Laskowski, Nicholas [ms]. 'Comments on Terence Cuneo.' Comments delivered at the Pacific Division meeting of the APA, April 2016.
- _____ [2017]. 'Speech and Morality: On the Metaethical Implications of Speaking, by Terence Cuneo.' *Journal of Moral Philosophy* 14(6): 781-784.
- Ross, W.D. [1938]. *Foundations of Ethics*. Oxford: Oxford University Press.
- Schmitt, Johannes, and Mark Schroeder [2012]. 'Supervenience Arguments Under Relaxed Assumptions.' *Philosophical Studies* 155(1): 133-160. Reprinted in Schroeder [2014a].
- Schroeder, Mark [2005]. 'Realism and Reduction: The Quest for Robustness.' *Philosophers' Imprint* 5(1).
- _____ [2007]. *Slaves of the Passions*. Oxford: Oxford University Press.
- _____ [2008]. *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.
- _____ [2010a]. *Noncognitivism in Ethics*. New York: Routledge.
- _____ [2010b]. 'How to Be an Expressivist About Truth.' In *New Waves in Truth*, edited by Nikolaj Jang Pederson and Cory Wright. New York: Palgrave MacMillan, 282-298. Reprinted in Schroeder [2015].
- _____ [2014a]. *Explaining the Reasons We Share: volume 1 of Explanation and Expression in Ethics*. Oxford: Oxford University Press.
- _____ [2014b]. 'The Price of Supervenience.' In Schroeder [2014a].
- _____ [2015]. *Expressing Our Attitudes: volume 2 of Explanation and Expression in Ethics*. Oxford: Oxford University Press.
- _____ [forthcoming]. 'The Moral Truth.' Forthcoming in the *Oxford Handbook to Truth*, edited by Michael Glanzburg.
- Singer, Peter [1972]. 'Famine, Affluence, and Morality.' *Philosophy and Public Affairs* 1(3): 229-243.
- Smith, Michael [1994]. *The Moral Problem*. Oxford: Basil Blackwell.
- Stevenson, C.L. [1937]. 'The Emotive Meaning of Ethical Terms.' *Mind* 46(181): 14-31.
- Tappenden, James [1993]. 'The Liar and Sorites Paradoxes: Toward a Unified Treatment.' *The Journal of Philosophy* 90(11): 551-577.