

RELATIVISM (AND EXPRESSIVISM) AND THE PROBLEM OF DISAGREEMENT

James Dreier
Brown University

Many philosophers, in different areas, are tempted by what variously goes under the name of Contextualism, Speaker Relativism, Indexical Relativism. (I'll just use Indexical Relativism in this paper.) Thinking of certain problematic expressions as deriving their content from elements of the context of use solves some problems. But it faces some problems of its own, and in this paper I'm interested in one in particular, namely, the problem of disagreement. Two alternative theories, tempting for just the same kinds of expressions as Indexical Relativism is meant to handle, promise to solve the problem of disagreement. I'll argue that they do not live up to their promise. At the end of the paper, I'll ask what exactly disagreement amounts to, and I'll canvass some purported solutions.

1. Indexical Relativism and the Problem of Disagreement

The View

Indexical Relativism is the view that moral expressions are indexicals. The property picked out by a predicate like 'wrong' is determined by the context of utterance; likewise, the facts stated by declarative moral sentences. The relevant feature of the context, according to the simplest version,¹ is the speaker's *moral system*. Here 'moral system' is a technical term. We can think of it as a formal object: a moral system will assign an intension (that is, an extension in each possible world) to each moral predicate.² If we focus on the meanings of the moral expressions, we can say likewise that they are functions taking contexts to intensions, or contents; they are Kaplanian *characters*.³ The particular feature of the context that the moral characters operate on is the speaker's moral sensibility or outlook.

Before I get to the objection, let me make explicit a couple of features of Indexical Relativism that I'll be discussing later.

First, the main motivation for the view is to explain Internalism, or the Practicality Requirement. Moral judgment has an intimate, conceptual or 'internal' connection to motivation, or at least so it seems to many. One good explanation for this connection would be that each person's moral judgments are logically attached to that person's motivational emotions. Suppose, as Indexical Relativism claims, your use of 'wrong' picks out, by its very semantics, a property of actions of which you disapprove (or would disapprove under idealized cognitive conditions). This seems like the right kind of connection to motivation.⁴

Second, according to Indexical Relativism there are no moral properties or moral facts. It's true (according to Indexical Relativism) that each use of a moral predicate picks out a property, and each assertion of a moral sentence purports to state a fact, but there is nothing particularly moral about the properties or facts that get picked out. They are, presumably, ordinary natural or artificial properties and facts, a bit cumbersome to state in plain descriptive language but with no whiff of anything especially moral. Compare: when I say that it rained yesterday, my assertion picks out a perfectly good fact (or falsehood), but the fact is not an *indexical* fact. It is the same fact I could pick out by means of yesterday's date. It is the way of stating the fact, rather than the fact, that is indexical. And we could say the same thing about the thought I sincerely expressed when I said that it rained yesterday; I was thinking about some rain on a certain date, but I was thinking about it in a particularly indexical way (from the point of view of *now*, maybe). All of this is received wisdom about indexicals. Indexical Relativism applies it to moral language and thought. Take Alastair, a committed utilitarian (and let's suppose that he would be a utilitarian in light of all the natural facts, too, and in reflective equilibrium). When he says that imprisoning a person we know to be innocent is sometimes morally permissible, the property he is attributing to the imprisonment of the innocent is a natural property, one we could denote by purely descriptive language, namely, the property of being among the happiness-maximizing alternatives. The moral quality attaches to the way he said it and the way he was thinking about it, not to the property or fact.

Apparatus

The technical machinery used by Indexical Relativism is familiar enough. The semantic values for moral expressions are *characters*, à la Kaplan, which means they are functions from *contexts* to *contents*. The content of a sentence is a proposition. The content of a predicate is a property. Sentences and predicates have contents on any given occasion of use, but the linguistic items have no fixed content of their own. There is no fixed content of 'wrong' or "Cannibalism is wrong", just as there is no fixed content of "I live in New York". But on each

occasion of use, given a context, those things do get contents. The content is what is said when someone asserts the sentence and what is believed when someone's belief can be naturally expressed by the sentence.

The *character* is the semantic value of the expression; it is what you have to know in order to know what the expression means. The *content* is what the expression says. These are technical terms, and what I have just said about them is said in ordinary language, and is no doubt a bit vague and squishy; it will serve well enough for now.

But all of this is familiar enough from Kaplan.

The Problem of Disagreement

Suppose I say

Withdrawing from Iraq is not wrong.

and Alastair says

Withdrawing from Iraq is wrong.

Then we are disagreeing. This fact, that we are disagreeing, is a piece of data. At least, by saying what we have said we are *intending* to disagree; I will assume here that it is not plausible for a theorist to say that we have failed to disagree even though we intended to do so.

But Indexical Relativism gets this wrong. It says we are not disagreeing. For according to Indexical Relativism, I have attributed a certain property to withdrawing from Iraq, and Alistair has denied that withdrawing from Iraq has *some different property*. Alistair's moral system and mine are not the same, and the properties assigned to 'wrong' by its character are different in the different contexts. So Indexical Relativism has a false implication.

Compare a case of genuine, uncontroversial indexical sentences. I say

I do not live in Colorado.

and Alistair says

I live in Colorado.

We do not disagree. The character of 'I' assigns to its pronoun a different content in each context. The person I say does not live in Colorado is not the person Alistair says lives in Colorado. So there is no disagreement; the contents of what we say are consistent. This is the way indexicals work.

I think there are some defenses available for an Indexical Relativist.⁵ For instance, it seems clear that sometimes an indexical will pick out a content that comes from the overlap of the contexts of speakers engaged in a conversation.

You and I can disagree about where something is when you say “It’s here” and I say “It’s not here”, even though we do not share a location, so long as there is an overlapping area determined by the conversational context. However, I am not here concerned to defend Indexical Relativism, and in any case once all the defenses have been mounted I think there is some residual Problem of Disagreement left unanswered.

This problem is an old one. It was identified by G. E. Moore, for example.⁶ Insofar as Indexical Relativism is a theory about what we say, how our language (and thought) works, it is subject to the Problem of Disagreement.⁷

2. Expressivism

Expressivism is a well-known theory of moral (and in general normative) language, so I won’t need to explain it in detail. I will be focusing on Allan Gibbard’s version in *Thinking How to Live*.⁸ There is one complication that I have to deal with up front. Gibbard’s theory of *moral* language is complicated; he explains it in terms of more basic normative vocabulary, and then the basic normative vocabulary gets the directly Expressivist treatment. So to be (even approximately) true to Gibbard, I have to use examples of normative judgment that aren’t moral. So let’s use this one.

Anthony ought to give battle.

Famously, Expressivists decline to say what this means by giving truth conditions, content, or character. Instead they tell us what someone, say Cleo, is doing when she says “Anthony ought to give battle.” What she does is to express a certain state of mind: a *planning* state, according to Gibbard. Planning states are a bit like intentions, a bit like preferences; they are in that conative category, ‘ready for action’; they are plans for what to do. In general, when somebody tells you that in a circumstance, C, you ought to ϕ , she is expressing her plan to ϕ in C.

Expressivism is supposed to have a large advantage over Indexical Relativism: it solves the Disagreement Problem. Let me mention four features of Expressivism briefly, and then turn to the solution.

1. Why, if it is her plan to ϕ in C, does she say that *you* ought to ϕ in C? Because whatever it is about you that is relevant to whether you are to ϕ in C gets built into C. Features of you that we would not ordinarily count as part of your ‘circumstance’ are still part of C; for instance, the fact that you are in an angry mood is a part of C, or your being an only child. So the plans, the elements of the planning state, are like conditional preferences. They are what get expressed by the basic *ought* statements.

2. *Expressed*, and not reported, notice. In saying that Anthony ought to give battle, Cleo does not say that she is in a certain planning state. She rather expresses that state. Her normative assertion bears to the state the relation that an assertion of descriptive fact bears to the speaker's belief in that fact. It is the sincerity condition, not the truth condition. It is important to Expressivism that we do not understand it to be assigning to normative statements some truth conditions that involve the speaker's state of mind. Sincerity conditions are not contents.
3. So what are the contents of normative statements, according to Gibbard? There is a quick answer and a longer answer; I can give the quick answer now and defer the longer answer to the next subsection. Expressivists think that nothing helpful can be said in the way of truth conditions for normative statements. One can just repeat the statements when asked for the truth conditions; a deflationary answer is always available.⁹ But the helpful way of explaining the meanings of normative statements, according to Expressivism, is to tell you what people do by using them, not what in the world would make them true.
4. Expressivism is well designed to explain Internalism, the Practicality Requirement.¹⁰ For a plan is, ordinarily, a motivational state; conation is a psychological plan's primary function. The internal connection between normative thought and motivation, then, is found in the Expressivist account of meaning: the internally connected state just is the state expressed when the judgment is sincere.

Now to Disagreement. When Cleo says, "Anthony ought to give battle," and Brutus says "Anthony ought not to give battle," they disagree. The Expressivist explanation is that they disagree *in attitude*, as Stevenson put it; they disagree over *what to do*. (Stevenson says that a normative disagreement is like the disagreement between two friends when one says "Let's go to the cinema tonight" and the other says "No, let's go to the symphony.") This is disagreement in plan. It is not spelled out in terms of contrary truth conditions, but in terms of clashing psychological states or conflicting advice. Normative statements have no truth conditions but for the deflationary kind, and contrary deflationary truth conditions do not *explain* disagreement but merely record it.

The Expressivist Problem of Embedding

There is a famous problem for Expressivist theories, called the problem of embedding. I'll outline this problem and Gibbard's solution, with some focus on the formal apparatus he uses. Then I'll explain why I think the solution is not complete, and what it would take to complete it. It turns out that the obstacle to completion is a new Disagreement Problem. The problem is to say what disagreement is.

Expressivism tells us what a normative sentence means by telling us what state of mind a person expresses by sincerely asserting the sentence. But that is not enough. For sentences can be used legitimately in other ways than by asserting them (even when the speaker is entirely sincere). A normative sentence can occur as the antecedent or consequent of a conditional, as a disjunct, in an indirect discourse, and in all sorts of other ways, embedded in a larger semantic matrix. Expressivism has to tell us how the normative sentence contributes to the larger meaning in such cases, if it is to tell us what the sentence means. And it does not have available the most traditional resource for doing so, since it denies that truth conditions play any explanatory role in the theory of meaning (of normative sentences, at least).

Now, some people think that far too much has been made of this problem.¹¹ They think, for example, that once we know what ‘if’ means, and what the antecedent and consequent of a conditional mean on their own, then we’ll have no problem understanding the conditional, and that this has nothing much to do with truth conditions. For instance, when you understand what it is to accept “Anthony ought to give battle”, you can understand a conditional of the form,

If P, then Anthony ought to give battle.

as an inference ticket. To accept it, the conditional, is be prepared to infer “Anthony ought to give battle” from P.

I’m somewhat sympathetic to this line. But for now, I am going to take the Embedding Problem seriously. I’ll now sketch Allan Gibbard’s solution.

Apparatus

The basic idea is to assign formal objects to normative sentences that can play the role that more familiar formal objects play in the more familiar semantics of descriptive sentences. In place of sets of possible worlds, we can use sets of *hyperplans*. Suppose someone had a complete view about how the world is; then her belief could be represented by a possible world. But nobody is like that. Even philosophers are not so opinionated. So, our doxastic states are instead represented by sets of possible worlds. As we become more opinionated, the set shrinks.

A hyperplan is to human plans as a possible world is to human beliefs. It is an unimaginably detailed contingency plan, with a course of action planned out for every possible circumstance.¹² Since real people are undecided about what to do in most possible situations (just as we are undecided about most details of what is the case, which world is ours), our planning states can be represented by sets of hyperplans.

As we know, sets of worlds are also good candidates for (representations of) truth conditions, so they can be semantic values of sentences. When sentences are

combined by truth functional connectives, their semantic values can be combined by set theoretic operations to yield the value for the new sentences. And logical validity can be represented by set theoretic relations, too: an inference from a set of sentences to a conclusion is valid just in case the intersection of the semantic values of those sentences is a subset of the value of the conclusion.

In the same way, sets of hyperplans can function as the semantic values of normative sentences. Then complex sentences can have values compounded from the values of their atomic parts in just the way complex descriptive sentences do, by set theoretic operations. And validity can be characterized for normative arguments just as it is for descriptive ones.¹³

There is something elegant about Gibbard's apparatus. Descriptive sentences are answers to questions about what the world is like; they get represented by sets of incredibly specific ways the world could be. Normative sentences are answers to questions about what to do; they get represented by sets of incredibly specific plans for what to do. What is the world like? One of these! What should I do? One of those! How natural.

We can think of the sets of hyperplans as normative *contents*. They are contents of sentences and propositional attitudes, alike (which is a good thing, since we attribute propositional attitudes by means of sentences that are supposed to give the content of the attitudes). So, now I have given the longer answer I mentioned above, on p. 83, to the question of what are the contents of normative statements. Instead of just giving deflationary truth conditions, we can mention these formal objects. I think this is somewhat helpful. But, let me assert without much defense that this answer is entirely consistent with a deflationary answer. In effect, the formal objects position their sentences in the web of inference. When you know which sets the sentences have as their semantic values, what you know is which things can be inferred from which. Or do you know more than that? I'll return to this difficult question later.

There is one last twist. Gibbard needs a way to combine normative contents with descriptive contents. For one thing, some sentences are conjunctions with one normative conjunct and one descriptive one. The semantic values of these had better not be the intersection of a set of worlds with a set of hyperplans; that intersection would, of course, be empty, so a normative/descriptive conjunction would have the semantic value of a contradiction. But anyway, lots of what we say and think is laden with both plan and belief, with advice and description. I could tell you that Nell is guilty of wrongful killing, or that Joseph is an evil dictator. Fortunately, the combination is easy enough. We can assign to every sentence a set of ordered pairs, $\langle w, p \rangle$, with w a world and p a hyperplan. Take a purely descriptive statement: Mars is red. Its set will be the set of all $\langle w, p \rangle$ such that Mars is red at w . No restriction on p , of course. So the Mars-is-red worlds each get paired up with *every* hyperplan. Likewise for purely normative statements. The content of the sentence, "Everyone ought to give battle when the prospect promises a greater chance of personal happiness," will pair every hyperplan that includes giving battle under those circumstances as a subplan,

with every possible world. But the statement, that Nell ought not to do what John has just done, will get a set of ordered pairs including each world in which John has done something (and Nell exists) with a plan not to do that something when in Nell's circumstances.

I hope this sketch is clear enough. I have omitted many details, and I am counting on the general scheme being somewhat familiar. To sum up: Expressivism solves the Disagreement Problem; it has its own distinctive problem, the Embedding Problem; Allan Gibbard's solution to the Embedding Problem involves an apparatus of formal objects, sets of ordered pairs, that play the formal semantic role that sets of worlds play in possible world semantics.

I have a bunch more to say about Expressivism and its apparatus, but before I say it I will introduce a second alternative. I'll explain how this alternative shares a foundation with Expressivism and Indexical Relativism; I'll sketch how it is supposed to solve the Disagreement Problem; I will introduce its formal apparatus. Once that's done, we can get to the hard part.

3. Genuine Relativism

I take the term 'Genuine Relativism' from Max Kölbel.¹⁴

The main claim for which I want to argue is the claim that there is a significant difference between two broad forms a relativist thesis can take: that of *indexical relativism* and that of *genuine relativism*.

Indexical relativists locate all relativity at the level of sentences, while genuine relativists claim that there is relativity also at the level of utterances and the contents or thoughts thereby expressed. Indexical relativists about, say, morality will hold that moral relativity is essentially a matter of moral sentences expressing different contents on different occasions of use. Moral sentences are thus very similar to indexical sentences in that the context of utterance determines which content is expressed by any utterance of them. Thus the same moral sentence can express one content and be true in one context of utterance, while it may express a different content and be false in another context . . . Genuine moral relativists do not claim that moral sentences behave generally like indexical sentences. They say that moral sentences express the same contents in all contexts of utterance (unless they are indexical for the usual reasons), but that these contents have their truth-values relatively, i.e. vary in truth-value with parameter of evaluation.

The idea that *contents* have truth values relatively sounds pretty strange. We often think of contents as truth conditions, after all, and if a theory insists that they are something other than truth conditions, we are apt to feel a bit lost. But Genuine Relativism has gained quite a bit of popularity. John MacFarlane and Andy Egan each apply it to some philosophically interesting areas of thought and talk, and they (and Kölbel) say enough in support to make the view worth a close look.¹⁵

According to Kölbel, moral statements are true or false only relative to a *perspective*, which for our purposes means they are only true or false relative to a moral system. So far, Kölbel's view doesn't differ from Indexical Relativism. But the Genuine Relativist doesn't say that the moral system (perspective, context) in the context of use determines a particular proposition, which is then simply true or false. Instead, he lets the relativity, the context sensitivity, continue into the proposition expressed by the sentence. Suppose Smart and Kant (to use one of Kölbel's examples) each assert

Punishing an innocent person to prevent great public harm is wrong.

Since Smart and Kant (we suppose) have different moral systems, Indexical Relativism tells us that they have expressed different propositions; Smart has attributed one property to punishing an innocent, Kant has attributed a different one. Genuine Relativism, on the other hand, tells us that they have expressed the same proposition. That proposition may be true relative to Kant's system and false relative to Smart's.

This view is, initially at least, hard to understand. We are tempted (at least I am) to think that for a proposition to be true relative to you and false relative to me couldn't mean anything if it didn't mean that you *believe* it is true and I *believe* it is false. This is definitely not what Genuine Relativism is saying. To see what it *is* saying, let me turn to John MacFarlane.

MacFarlane points out that we are familiar with a couple of kinds of Genuine relativity already. First, we all accept that there is evaluator-relativity with respect to *worlds*. Suppose Jane says

The earth has exactly one moon.

Now imagine a merely possible world in which the earth has two moons, and in which June says

The earth has two moons.

Jane and June each expressed a proposition, and their propositions are contraries: they cannot both be true. But what Jane says is true in the actual world and false in June's merely possible world; what June says, on the contrary, is false in the actual world and true in June's world. Which world is the right one? That's a silly question. So, Jane's proposition is true at our actual world and false at June's possible world; is it *true*, or is it *false*? Again, it's silly to insist that it be one or the other. A proposition can be true at one world and false at another.

A second kind of relativity, though slightly more controversial, is familiar enough not to be written off as bizarre and incomprehensible. It is relativity to time. Suppose I say,

The United States has a black President.

And suppose Dolly Madison once said,

The United States has never had a black President.

It is plausible that we have expressed contrary propositions, and that Dolly's was true-at-her-time while mine is true-at-my-time. We can certainly say comfortably,

What Dolly said used to be true, but it isn't any longer.

Compare this to June's proposition. I can comfortably say,

That would have been true in June's situation, but in fact it isn't true.

That is, I can quite happily say *of these propositions* that they are true at some context other than mine, and then say that they aren't true, where I'm using my own context to fill in the missing relatum (in one case a world, in the other a time).

MacFarlane:

Taking this line of thought a little farther, the relativist might envision contents that are "sense-of-humor neutral" or "standard-of-taste neutral" or "epistemic-state neutral," and circumstances of evaluation that include parameters for a sense of humor, a standard of taste, or an epistemic state. This move would open up room for the truth value of a proposition to vary with these "subjective" factors in much the same way that it varies with the world of evaluation. The very same proposition — say, that apples are delicious — could be true with respect to one standard of taste, false with respect to another.¹⁶

Suppose we agree that this makes sense. What is the advantage of Genuine Relativism over Indexical Relativism? Genuine Relativism is supposed to solve the Problem of Disagreement. Take MacFarlane's example of judgments about what is *funny*. These are Genuinely Relative to a sense of humor; what's funny relative to one sense of humor may not be funny relative to another, and there is no absolute funniness. Suppose we both hear a rather tasteless joke, and I say,

Tasteless, no doubt, but at least it's funny.

You say,

It's not even funny.

Do we disagree? An Indexical Relativist about the funny would, of course, have to say we don't. But a Genuine Relativist can say that we do. For I accept, and

you reject, the *proposition that the joke is funny*. And this seems to be sufficient for disagreement: one of us accepts and the other rejects the very same proposition.¹⁷

There is quite a bit more to be said about disagreement a la Genuine Relativism. Before I start to say it, I want to set up the formal apparatus of Genuine Relativism. Then, in the following section of the paper, I'll muse over the similarities of the three sets of apparatus: one for Indexical Relativism, one for Expressivism, one for Genuine Relativism. To peek ahead: the upshot of the musings is going to be that the formal structures of the three semantics are similar, indeed isomorphic, and to appreciate the differences among the views we have to look elsewhere. After that, I'll return to Expressivism for one section, explaining a new problem that pops up in the wake of Gibbard's Expressivist semantics. And then, finally, I'll return to the Genuine Relativist account of disagreement, arguing that its shortcomings mirror those of Expressivist semantics.

Apparatus

As a starting point, the simplest apparatus for Genuine Relativists is the one used by Andy Egan.¹⁸ Recalling David Lewis's account of belief *de se*, Egan suggests that the values of some predicates are not properties but what he calls 'centering features'. Centering features combine with objects of predication to form centered propositions, which are the objects of self-locating beliefs.

Centered propositions first. Just as propositions can be thought of as sets of possible worlds, centered propositions are sets of *centered* worlds. Centered worlds are to possible worlds what maps containing a "You are here" arrow are to ordinary maps you'd find in an atlas. Formally, they are just worlds together with a context, which might be just an <individual, time> pair. Sets of them are (or represent) centered propositions. Some beliefs, Egan thinks, are attitudes toward centered propositions rather than toward centerless ones (Lewis thought so, too). For instance, the belief I might articulate by saying "It's late afternoon and the air is getting colder" is an attitude toward the set of centered worlds whose center is at a spot in space and time where (and when) it is late afternoon and the weather is getting colder. It contains some centerings of the actual world, of course, and lacks others. And it contains some centerings of many merely possible worlds, and lacks other centerings of those same worlds.

We can add moral systems into the contexts (the ordered tuples that get paired with worlds to form centered worlds), if we like; or we can just let the individual in the context supply his or her own moral system. Then we have formal objects for MacFarlane's relativism, and Köbel's. A moral sentence like "It is wrong to punish an innocent" has a truth value only relative to a world — this much is already familiar; but more, it has a truth value only relative to a *centered* world. It contains all and only those centered worlds centered on moral systems according to which punishing innocents *in that world* is wrong.

Let me sum up this section. Genuine Relativism agrees with Indexical Relativism on the claim that the semantic values of moral sentences have an extra

parameter: they need a moral system along with a possible world to determine a truth-value. It disagrees with Indexical Relativism about how and where the extra element is supplied. Indexical Relativism says that a moral sentence has a character, and that on an occasion of use it will express a plain vanilla proposition, perhaps a set of possible worlds, which can be simply true or false (at a world). Genuine Relativism says that the content the sentence delivers on an occasion of use is still relative. That content is not a plain vanilla proposition (set of worlds), but rather a centered proposition (set of centered worlds). The content itself is true or false only relative to a context (in particular the moral system in the context).

When you and I each assert an indexical sentence, we (often) say different things; when the assertions are sincere we (often) *believe* different things (I believe that *I* am the tallest philosopher in the room, while you believe that *you* are). Indexical Relativism extends this idea to moral sentences. So it runs into the Problem of Disagreement: we say different things by assertion of the same sentence, and likewise I can deny something different from what you assert, when I assert the syntactic negation of the sentence you assert. Genuine Relativism claims that in certain areas of language, besides indexicality we have another form of extra indexing, one that gets carried into the things said and the things believed. These are the centered propositions. When I negate the sentence you assert, I am denying *what you said*, and so we disagree.

4. Three Formal Apparatus

Here are the three models, the formal semantics for the three accounts of moral language.

Indexical Relativism assigns to each moral sentence a Kaplanian character, namely, a function from contexts to contents. The important feature of the context is a moral system. To a sentence that predicates ‘wrong’ of a subject, for example, the function assigns a proposition saying that the subject has a certain property, P: and P is the property the moral system in the context assigns to ‘wrong’. A person (in a context) has a moral system that is determined by the person’s moral attitudes, possibly in some idealized form.

Expressivism assigns to each moral sentence a set of <world, hyperplan> pairs. When a sentence says that a person in a circumstance *ought* to act a certain way, its semantic values includes the set of all pairs for which the hyperplan includes a plan to act in that way in those circumstances in that world. These ‘factual-normative contents’ then play the role that sets of possible worlds play in possible world semantics.

Genuine Relativism assigns to each moral sentence a centered proposition, namely, a set of centered worlds (or <context, world> pairs). Moral predicates contribute centering features, which are formally represented as functions from contexts to properties.

These formal models are so similar that they can appear to be almost trivial variants of one another. Even my description so far is enough to make out a certain similarity. But, as I will now explain, they are even more similar than they appear so far.

The model for Expressivist content and the model for Genuine Relativism rely on collections of worlds, and they are beholden to possible world semantics. The model for Indexical Relativism includes nothing (said here) about possible worlds, and is independent of possible world semantics. Indeed, Kaplan semantics is an alternative to possible world semantics; its propositions are Russellian structured objects (though I have made no use of this fact in the presentation).

What if possible world semantics is defective? For some purposes, it is too 'coarse grained'. Possible world semantics models every necessary proposition with the same formal object, and also assigns the same formal object to every impossible proposition, and to every pair of propositions true at the same world. I will not worry about whether this is a serious defect. For some purposes it is a definite disadvantage (I believe that $2 + 3 = 5$, but I do not believe that there is any even number greater than two that isn't the sum of two primes, and I don't believe that every even number greater than two is the sum of two primes). So, for some purposes we want some other objects to represent contents, not sets of possible worlds.

Whatever the defects of possible world semantics, the models that adopt its apparatus (by using sets of <world, hyperplan> pairs or sets of centered worlds) will inherit those defects. Fortunately, this is easy to fix.

Expressivist content first.¹⁹ We note first that the family of sets of <world, hyperplan> pairs is isomorphic to the family of functions from hyperplans to sets of worlds. Indeed, for mathematical purposes a set of <world, hyperplan> pairs *is* a function from hyperplans to sets of worlds. Take some such set; we construct for it a function. The function takes hyperplans as arguments, and for each hyperplan the value of the function is the set of all worlds that hyperplan is paired with in the set. (The set doesn't determine any particular function from hyperplans to worlds, because each hyperplan is paired with many worlds. Gather those worlds together into a set; this is the value of the function at that hyperplan.)

This scheme matches Gibbard's sets of pairs onto functions from hyperplans to sets of worlds. And the re-interpretation is natural enough. Take the sentence,

Anthony ought to give battle.

We are assigning to this sentence a function from hyperplans to sets of worlds. The function will take a given hyperplan to the set of worlds in which Anthony is in one of the circumstances the hyperplan says to give battle in.

A set of worlds is, in possible world semantics, a proposition. If for some purpose we aren't satisfied with possible world semantics, we can substitute whatever conception of a proposition we like. Suppose we like Russellian

propositions. Then we can assign to each normative sentence a function from hyperplans to Russellian propositions. There is nothing special, from the point of view of the general semantic scheme, about sets of worlds. The only reason to pair hyperplans with worlds, in Gibbard's semantics, is to draw on the framework of possible world semantics. But if some other framework is better, we can simply adjust Gibbard's scheme to that framework.²⁰

Now the semantics assigns to each sentence a function from hyperplans to propositions. And this is in effect a Kaplanian character, so long as a hyperplan is a feature of a context. Indexical Relativism takes moral systems as the relevant feature of the context; hyperplans are so close, formally speaking, to moral systems that the difference seems to make no difference.

As to the Genuine Relativism model: sets of centered worlds are (isomorphic to) functions from centers to sets of worlds, and the centers are contexts. So again, the semantic values of the kinds of sentences in question are really Kaplanian characters; that is, they are the same formal objects.²¹

MacFarlane doesn't say much about what formal objects propositions (relative or absolute) are. Instead he focuses on their truth and assertibility. But it is clear that for him, some propositions are true (or false) only relative to a context of use and a context of assessment. It is the context of assessment, for MacFarlane, that gives Genuine Relativism its punch. Without it, the proposition won't give us truth conditions (thought of as functions from worlds to truth values). So Genuine Relativist propositions, the values assigned to sentences, are truth conditions minus contexts; each proposition will yield truth conditions when provided with a context (of evaluation). So they are (well represented by) functions from contexts to truth conditions.

The three kinds of metaethical theories, then, assign (more or less) the same formal objects as semantic values of moral sentences. Of course, the objects are *called* different things by the three theories ('characters'; 'factual-normative contents'; 'centered propositions'), but *that's* not a substantive difference. If there is a substantive difference it has to show up in what work the theories do with their formal objects.

Notice that all three theories agree that the objects assigned to sentences as semantic values are also appropriately connected to belief states that a person might express by uttering the sentence assertively. Good thing! Indexical Relativism will not allow that the semantic value is the *content* of the belief so expressed. Contents are derived from characters by means of contexts. The other theories do call the objects 'contents'. But what difference does this make? Objects that have their place in a formal structure are useful to index states of mind. We identify a type of psychological state by tagging it with an index. And Indexical Relativism does agree with the other two theories on this score: there is an interesting taxonomy of psychological states according to which two states (two of yours over time, perhaps, or one of yours and one of mine) belong to the same taxon when we express them by means of sentences with the same character. No difference there.

Below I'll give some details of Genuine Relativism to draw out the distinctive role it gives to its formal objects. For now I'll focus on two purported differences between Genuine Relativism and Expressivism, on the one hand, and Indexical Relativism on the other.

First Difference: Indirect Discourse

As I said, Genuine Relativism and Expressivism call the function-from-contexts-to-propositions, 'contents', while Indexical Relativism doesn't. The name doesn't matter. But our ordinary, intuitive grasp on contents, I think, is via indirect discourse. I can tell you *what someone said* and *what she believes*, and when I do this the *content* of the sentence I use is supposed to be the same as the content of the saying or belief. This is what's behind calling some things 'contents' in a theory. So we can get some traction by looking at indirect discourse.

But discussion of the accounts the various views give of indirect discourse would take us too far afield. One difference comes out immediately in the accounts of disagreement. Aside from that one, I want to stipulate that there may be interesting differences among the theories on this score, and leave it at that.

Second Difference: Disagreement

It is one of the stated aims of Genuine Relativism that it improve on Indexical Relativism's handling of disagreement.

The relativist's central objection to contextualism is that it fails to account for the possibility of disagreement in subjective discourse—for our sense that when I say that carrots are delicious and you deny this, we are genuinely disagreeing with each other, and not making compatible claims about our respective tastes. If we are to adjudicate between contextualism and relativism, then, we must first get clear about what it is for two people to disagree.²²

We know what the Problem of Disagreement is: there is no proposition, according to Indexical Relativism, which you believe and I disbelieve when you sincerely express your belief by saying that something is, say, '*wrong*', and I express mine by saying the same thing is '*not wrong*'. Note that the quotation marks are necessary. I, your author, cannot (according to Indexical Relativism) describe the doxastic and conversational clash by saying,

I think the invasion was wrong, but he thinks it is not wrong.

For by saying that I would be reporting you as believing that the invasion lacks the property that 'wrong' is assigned by *my* moral system, *my* context. Indexicals can't be used that way. Suppose I phone my wife and ask where her W2 form is.

She: I'm sure it's not here.

I: Okay, it must be here.

Plainly I cannot summarize by saying, "I think the W2 form is here but Johanna thinks it is not here." Johanna and I do not disagree.

Genuine Relativists claim that they have a solution. There *is* something for us to disagree about, they say, only it isn't a set of worlds (it's not the question of which world we are in). Andy Egan writes:

I am partial to a picture of mental and linguistic content according to which the role of mental states and linguistic representations is to distinguish between possibilities. My beliefs distinguish between the possibilities that I take to be candidates for actuality and the ones that I rule out, my desires distinguish between the possibilities that I hope for and those that I dread, and my assertions distinguish between (roughly) those possibilities that I'm asking you to rule out and those that you're free to leave open.

If we like this possibility-sorting picture of content, then it's very natural to represent contents as sets of possible worlds. The content of a belief, desire, or assertion is the set of worlds where things are as they're believed, desired, or asserted to be.

My beliefs, desires, etc. with possible-worlds content draw distinctions between ways the world might be, while my beliefs, desires, etc. with centered-worlds content draw distinctions between situations that I might be in.²³

So, in a moral disagreement, you can believe what I disbelieve, you assert what I deny: a centered proposition. Max Kölbel agrees:

[E]very thinker *possesses* a perspective, and moreover everyone ought not to believe contents that are not true in relation to their own perspective. On this basis, it is clear why I can't come to believe what you said without needing to change my mind: what you have said and what I have said cannot both be true in relation to the same perspective. Thus, given that I ought not to believe something that is not true in relation to my perspective, I should not come to believe what you have said without changing my mind.²⁴

The objects of belief, in Kölbel's picture, are not truth-conditions; they are truth-conditions minus a 'perspective'. Sets of centered worlds (with the centers singling out the perspective) or whatever centered propositions one prefers will do the trick. And we disagree, Kölbel says, because there is an object you believe which I cannot (or anyway should not) believe, since I believe something that is incompatible with it: both cannot be true in the same perspective.

Formally speaking, Gibbard's Expressivism works out similarly. Take a simple case in which the sentence you use to express your planning state is the syntactic contradictory of the one I use to express mine. The contents of our respective judgments, then, will be sets of factual-normative world pairs, and

the two sets are complements; so they are related just as ordinary descriptive contradictions are related. It is incoherent to accept both of these propositions. As Gibbard puts it, each planning state ‘rules out’ some combinations of descriptive fact and normative planning. Contradictories will together rule out all factual-normative possibilities, leaving me in the unhappy state of having ruled out every contingency plan (indeed, any pair of contraries will have a null intersection in Gibbard’s scheme, so all such pairs leave me in the unhappy state). Disagreement is recaptured in Gibbard’s semantics by the same sorts of formal objects as we see at work in Genuine Relativism’s maneuver. The contents of the two statements and beliefs are contradictories, which cannot be coherently accepted together.

In the following sections, I will argue for what I hope has occurred to the reader: the formal moves do not really solve the problem. My argument will proceed as follows. In the next section, I will point out that there is a deficiency in Expressivist semantics: it does not have an adequate semantics of negation. It turns out, I argue, that in order to make good on this defect Expressivism has to give an independent account of what it is for one state of mind to disagree with another. The semantic model, therefore, has not *solved* the disagreement problem so much as presupposed that it has an answer. Then I’ll argue in the section after that, that Genuine Relativism has just the same problem. And in the last section I’ll look at some suggested solutions to this problem and assess their prospects.

5. Expressivism and Negation: the new Disagreement Problem

I said that Expressivists claim an advantage for their view over Indexical Relativism: that they have a plausible account of disagreement. And, in simplistic terms, their account is supposed to be this: that when you and I have a normative disagreement, we disagree not in belief but in some other attitude. We disagree in what we plan, or, in other versions, in our attitudes toward various possibilities. We disagree over *what to do*.

The semantic model, in which the disagreeers believe or assert contradictory contents (in the form of factual-normative propositions, centered ones) is not a substitute for the main explanation, it is important to see. Expressivists recognize that just giving us a bunch of formal objects, assigned to the various sentences, doesn’t do the explanatory work that needs to be done. The real explanation comes in their story about the attitudes: planning, emotive, motivational.

There is, I think, some intuitive sense in what Expressivism says here. To put it crudely, when you shout “Booo!” for the Yankees and I shout “Hooray!”, that seems like a *kind* of disagreement. Is it the right kind?

The situation is somewhat complicated. To expose the complication, let’s look at the Expressivist story about the following sentences.

Judith thinks you ought to write the invitation by hand.

Now consider three ways of adding negation.

It is not the case that Judith thinks you ought to write the invitation by hand.

Judith thinks it is not the case that you ought to write the invitation by hand.

Judith thinks you ought not to write the invitation by hand.

To see the difference: the first is true if Judith has no view whatsoever about invitations; the others aren't. The second is true if Judith has the considered view that it doesn't matter whether you write the invitation by hand; the third is not.

The problem is that there aren't three planning states for the three 'negations' to ascribe to Judith. Maybe she simply has no plan for how to write invitations (in your circumstances); maybe she has a definite plan *not* to write them by hand. These, I think, must be what are ascribed by the second and third negations. But what is ascribed by the second? This is the Negation Problem.²⁵

One way to see the root of the problem is to focus on the fact that there are (intuitively) three deontic statuses for a brand of invitation-writing to have. It could be required; this is the status Judith takes it to have according to the negation-free attribution. It could be forbidden; that's what Judith thinks according to the last of the attributions. Or it could be *merely permitted* or *optional* — permitted without being required. That's what Judith thinks according to the middle negation option. And according to the first negation option, she has no view, and note that this is not the same as having the view that hand-writing invitations is optional. Someone who has never heard of etiquette or invitations doesn't have the view that handwritten invitations are permissible. So, leaving aside the situation in which Judith has no view, there are three statuses but only two sorts of planning states: planning to write them, and planning not to write them.

What sort of state is represented by a set of factual-normative worlds, some of which include the plan to write invitations by hand (in your circumstance) and others of which include the plan to print them on your laser printer? This might seem to be a 'permissive' state. But it isn't. It's an undecided state. Compare the belief represented by a set of worlds, in some of which our galaxy has an odd number of stars and in others of which it has an even number of stars. A person in such a state has no definite view about how many stars are in our galaxy. She is agnostic. The analogous planning state is also agnostic. Again, suppose someone is in the pure doxastic state (about the stars), and then learns more (implausibly!) and decides that the number of stars in the galaxy is odd. She has not changed her mind, but only resolved an indecision. Similarly, when a planner changes from the state represented by a set of plans, some planning to write by hand and some planning to print, and moves to a state whose representation includes only plans to write by hand, she has come to a decision and not changed her mind. This shows that she has not switched from regarding the printed invites as

permitted to regarding them as forbidden (since that would be a change in view, not a resolution of indecision).²⁶

Later I'll sketch out some proposed solutions. For now, though I note that the hitch in the semantics points directly at a problem about disagreement. An Expressivist might say, "Well, what's happening when someone thinks it is *permissible* to print out the invitations is that she is *disagreeing* with everyone who thinks it is required to write them by hand. And this disagreement is revealed in the wording: she believes it is *not* the case that one ought to write them by hand; the *not* signals disagreement. Whereas when we say only that she does *not* believe that one ought to write the invitations by hand, we are signaling only our own disagreement with someone who has a different view about Judith's state of mind."

This would be a good answer, but only if we could be given an explanation of what it means for Judith to disagree with a plan. We had a suggestion on the table: to disagree with a plan is to have an incompatible plan. (Whether this can be properly thought of as disagreement remains to be seen; I have been assuming that it can.) But that cannot be the suggestion at this stage. For the incompatible planning state is expressed by the last attribution, and we are now to suppose that the second also attributes to Judith a planning state that disagrees with one that the unnegated attribution assigns to her. So the problem is that we have no explanation of what it is for one state of mind to disagree with another.

To be clear: the problem is not with the formal apparatus. It's true that Gibbard's sets of factual-normative pairs cannot represent all three normative statuses, but presumably some other formal objects could be wheeled in, with more structure and so better able to distinguish statuses. The question is about what states of mind the representations would be mapped onto. And my point is that to find enough, we need an answer to the question, what is it for two states of mind to be in disagreement? If we knew that, we could just suppose that the state of believing it is not the case that one ought to φ is the state of disagreeing (and no more) with the state of believing that one ought to φ . This is the new problem of disagreement.

6. Genuine Relativism and the New Problem

Genuine Relativism offers up some formal objects for you to accept and me to reject, and it says that in such cases we disagree. But that is not enough. Just as Expressivism's collections of factual-normative pairs cannot answer the explanatory question, neither can Genuine Relativism's centered propositions.

According to Andy Egan, the contents of our judgments are often sets of centered worlds (the sets themselves are centered propositions, awaiting a center to deliver a proposition assessable for truth). These sets, as he puts it, "draw distinctions between situations that I might be in" (and he could have added,

“or that you might be in”). Our centered beliefs may draw such distinctions — each of us taking himself to be in a particular sort of situation — even if we agree exactly about what the world is like. And he thinks believing that Sydney is nearby has a content of this type, since it represents to the believer which kind of situation he is in.²⁷ Egan writes,

[I]f I am in Canberra and you are in Boston, and we are both to be maximally well-informed, we must both agree that Sydney is near Egan. We ought not to agree about whether Sydney is *nearby*.²⁸

But this example illustrates my point. Suppose that collections of centered worlds are a good way of capturing the ‘content’, in some reasonable technical sense, of a person’s state of mind, and also ‘what is said’ in asserting a sentence that might also be used to express the state of mind. This supposition leaves it open whether two believers whose sets of centered worlds stand in some formal relation of exclusion to one another can be thought of as *disagreeing*. And his own example shows that the question about disagreement isn’t settled by the model. For when Andy thinks, *Sydney is nearby*, and you think, *Sydney is not nearby*, you and he most definitely do *not* disagree. The plain facts of the matter, stated in centerless terms, are that Andy is near Sydney and you are not, and once this is understood there is no residual ‘proposition’, centered or otherwise, that can be the nexus of your disagreement.

John MacFarlane recognizes that a semantic object you accept and I reject cannot be the criterion of disagreement, even when the object is called a proposition. He considers this criterion for when two parties disagree:

Accept/Reject: There is a proposition that one party accepts and the other rejects.

But he rejects the criterion.

This can be seen most clearly when we relativize propositional truth to parameters besides just worlds. Consider, for example, tensed propositions, which have truth values relative to world/time pairs. One such proposition is the proposition that Joe is sitting. (Do not confuse this with the proposition that Joe is sitting now, or at any other time: the tensed proposition is, in Kaplan’s terms, “temporally neutral.”) If you asserted this proposition at 2 p.m. and I denied it at 3 p.m., we have not in any real sense disagreed. Your assertion concerned Joe’s position at 2 p.m., while my denial concerned his position at 3 p.m. So accepting and rejecting the same proposition cannot be sufficient for genuine disagreement.

Lest anyone be tempted to save **Accept/Reject** by denying that propositions can be “temporally neutral,” the point can be made just as well with eternal propositions (with truth values relative to worlds but not times). Just as **Accept/Reject** can serve as a criterion for disagreement about tensed propositions only when the acceptance and rejection take place at the same time, so it can

serve as a criterion for disagreement about eternal propositions only when the acceptance and rejection take place in the same world.

Consider Jane (who inhabits this world, the actual world) and June, her counterpart in another possible world. Jane asserts that Mars has two moons, and June denies this very proposition. Do they disagree? Not in any real way. Jane's assertion concerns our world, while June's concerns hers. If June lives in a world where Mars has three moons, her denial may be just as correct as Jane's assertion.²⁹

On the other hand, Max Kölbel *does* claim that Genuine Relativism accounts for disagreement.

Another difficulty of indexical relativism was the fact that it had to give a counterintuitive account of moral disagreements. According to [Indexical Relativism], when I sincerely utter ['Blair ought to go to war'] and you sincerely utter 'It's not the case that Blair ought to go to war', what I said is not incompatible (in the right way) with what you said. I can just come to believe what you said without needing to change my mind. There is no such problem in the case of genuine relativism. However, I shall need to introduce one further, normative aspect of this theory in order to show how this works: every thinker possesses a perspective, and moreover everyone ought not to believe contents that are not true in relation to their own perspective. On this basis, it is clear why I can't come to believe what you said without needing to change my mind: what you have said and what I have said cannot both be true in relation to the same perspective. Thus, given that I ought not to believe something that is not true in relation to my perspective, I should not come to believe what you have said without changing my mind.³⁰

Ragnar Francén has explained nicely why Kölbel's explanation is unsatisfactory.

Kölbel is right that on his view, when Kant says that punishing innocent Irwin is wrong and Smart says that that this is not so, the propositions they assert cannot both be true at any one circumstance of evaluation (moral perspective). But this does not mean that they disagree. In analogy with Jane's and June's assertions, Smart's and Kant's assertions concern different circumstances of evaluation, different moral perspectives. It might very well be that they agree that punishing innocent Irwin is wrong at Kant's moral perspective and that punishing innocent Irwin is not wrong at Smart's moral perspective.³¹

Francén adds (using 'speaker relativism' for our 'Indexical Relativism'),

This is very similar to the problem speaker relativism has with disagreements. The problem for moral speaker relativism is that it makes moral assertions made by speaker's with different moralities be about different things (express different propositions), and that they therefore do not disagree in the intuitive sense when they are involved in moral disputes. Kölbel's variant of relativism makes moral assertions made by speakers with different moralities concern different moralities.³²

The idea of a proposition's 'concerning' one circumstance or another is not spelled out rigorously, but MacFarlane's use of it, adopted by Francén, seems fairly intuitive. A temporally neutral proposition asserted at a time concerns that time, in a sense I can't spell out myself but feel I can grasp pretty well; somewhat less clearly I understand a sense in which June and Jane make assertions each concerning her own world. Maybe the simplest way to make the point is that we are completely comfortable with the idea that each of them is quite correct and each speaks truly.³³

If we like the way Francén describes things, we might say the new problem isn't new after all. But the focus has changed. Instead of a phenomenon of the semantic model, the problem now seems to have to do with saying why the cases that the model counts as disagreements should be thought of as something robustly in conflict. That was the problem for Expressivism, too: say why a pair of attitudes toward the world should count as disagreement, rather than just difference.

In the last section, I'll look at how an account of disagreement might go. I warn the reader that this last section is fairly negative. I am more confident about what won't work than I am about what will or even might.

7. Explaining Disagreement

My claim is that neither Genuine Relativism nor Expressivism really has a better answer to the disagreement problem than Indexical Relativism, and that the difficulty is revealed to be one of what exactly disagreement *is*. John MacFarlane more or less agrees:

The relativist's central objection to contextualism is that it fails to account for the possibility of disagreement in subjective discourse — for our sense that when I say that carrots are delicious and you deny this, we are genuinely disagreeing with each other, and not making compatible claims about our respective tastes. If we are to adjudicate between contextualism and relativism, then, we must first get clear about what it is for two people to disagree. This question turns out to be surprisingly difficult to answer. Although the answer given below will be incomplete, I think it does shed a little light on what the relativist must say if she is to do better than the contextualist in securing genuine disagreement.³⁴

Here is his formal account of disagreement.

[T]wo parties disagree (as assessed from context C) if

Can't Both Be Accurate (Relative to C). (a) There is a proposition that one party accepts and the other rejects, and (b) the acceptance and the rejection cannot both be accurate (as assessed from C).

He then adds:

Perspectival Accuracy An acceptance (rejection) of a proposition p at a context C_U is accurate (as assessed from a context C_A) iff p is true (false) at the circumstance $\langle W_{C_U}, S_{C_A} \rangle$, where W_{C_U} = the world of C_U and S_{C_A} = the standard of taste of the assessor at C_A .

These together provide an account of disagreement. **Perspectival Accuracy** is a necessary component since otherwise we would not understand “accurate (as assessed from C)” in **Can’t Both Be Accurate**.

But MacFarlane is not satisfied with this explanation.

But what does it mean to say that a speech act or mental state is accurate “from a perspective” or “relative to a context of assessment”? The relativist needs to say something about the practical significance of claims of assessment-relative accuracy. How does it matter in practice whether a speech act or mental state is accurate relative to one context of assessment rather than another? What turns on this?

I would put things this way. From Egan, we know that centered propositions *can* represent contents of judgments about what is nearby. We can then define **Perspectival Accuracy** for propositions like the proposition that Sydney is nearby. It would turn out that Egan’s acceptance of the proposition that Sydney is nearby is inaccurate as assessed by me today, even though I know he accepted it when he was in Canberra (so it was accurate as assessed by him then). That’s no problem; **Perspectival Accuracy** is, after all, a purely technical notion, a defined term. Now, my judgment that Sydney is not nearby and Egan’s judgment that Sydney is nearby **Can’t Both Be Accurate**, in the technical sense. So the criterion says we disagree. But we do not.

It is pretty clear what MacFarlane has to say about this example. He has to say that our actual practice does not support relativist semantics (and the relativist account of disagreement) for ‘nearby’ sentences. And this shows that the real question about disagreement has to do with the stuff, the ‘practice’, that makes the formalism relevant and helpful (in some cases and unhelpful and irrelevant in others).

One approach, famously employed by Brandom, is to characterize the practice in terms of conversational norms. MacFarlane suggests something like this.

This, then, is the practical significance of the classification of assertions into accurate and inaccurate:

Accuracy and Challenges. Accuracy is the property we must show assertions to have in order to vindicate them in the face of challenges, and it is the property we must show others’ assertions not to have if our challenges are to be justified.

And here is how the account works when accuracy is relativized.

Making the relativitization explicit, we can see that there is work for a relativized notion of accuracy to do:

- one is entitled to challenge an assertion when one has good grounds for thinking that the assertion was not accurate (relative to the context of assessment one occupies in issuing the challenge), and
- a successful response to such a challenge consists in a demonstration that the assertion was, in fact, accurate (relative to the context of assessment one occupies in giving the response).³⁵

I have no fundamental objection to this sort of explanation, but what I want to say is that it is pretty radically incomplete. Take the rules of some Brandomian challenge/response games, and see what kinds of tokens can intelligibly play the roles of *assertions*. (I don't mean what kinds of sounds or shapes of symbols, of course, but what kinds of speech acts, as it were, conceived independently of the rules.) Empirical statements can apparently fill the role just fine — the game in which I make some observationally testable hypotheses and someone challenges me, is perfectly intelligible. Now let an expression of pain sit in the spot that assertions occupy. Suppose we introduce a sentence allowed to be asserted when the speaker has a headache. Well, we *could* play this game. I shout, “It is throbbly in here”, and you challenge. You assess “It is throbbly in here” from your own perspective and, pain free, find it badly defective (inaccurate); you find you were entitled to your challenge. I now roll my eyes, place my head in my hands, grope around for Ibuprofen, thus demonstrating that from my perspective my assertion was entirely accurate; I have successfully (from my perspective) met the challenge. We could play this game. But it would be a bad game.

Again, MacFarlane is apparently worried about almost the same thing:

This can look like a pretty silly game. Why do we play it? Assuming we do have assessment-sensitive expressions in our languages, why do we have them? What would we be missing if we did not?³⁶

But this is not quite *my* point. The game as we were playing it seems not to make any kind of conversational sense at all. In this way it is quite different from other kinds of pointless conversational games we might play. For example, some people think that quite a lot of philosophy is a pointless game; people sometimes point out that thinking about what *would* have happened had Al Gore been declared President by the U. S. Supreme Court is a pointless exercise; and so on. But these games are conversationally intelligible even if their critics are right about their pointlessness. When you say that human beings are four-dimensional and I say we are three dimensional, this at least makes sense as a disagreement even if there isn't anything either of us can say to budge the other from his position (and we know it in advance). If I insist that America would have been

the victim of six deadly terrorist attacks had Al Gore been declared President and you opine that we would have had eight years of peace and prosperity, this is intelligible even though we quickly see that there is no prospect of convincing one another and no sufficient agreement even in what we'll count as evidence. But when I insist that it is throbbing in here and you reply that it isn't the least bit throbbing, all that is happening is that I have a headache, and you do not, and we both know this . . . and there is no intelligible sense in which we disagree.³⁷

So I want to say that there are preconditions that have to be met before a game of challenge and response can be thought of as a conversation in which the assertions are (for want of a better word) propositions. And one precondition, maybe the only one, is that the state of mind expressed (as an Expressivist likes to put it) is the kind of thing that we can disagree with.³⁸

Solutions

MacFarlane's explanation goes roughly like this. The challenge/reply game has as its point to produce a feeling of controversy (by comparison with a conversation in which each of us merely reports our own attitudes). The feeling of controversy is uncomfortable, so conversations that produce it will make each of us take some interest in removing its cause. And in practice this will mean reaching some kind of *coordination* of our attitudes, so that they converge toward a single outlook. Of course, this point could not possibly be had by a challenge/reply game in which we seem to dispute questions of which cities really are nearby — we could not coordinate our locations conversationally, even if there were some reason to want to do so.

This story is remarkably close to C. L. Stevenson's account of the use of moral language.³⁹ In ethical discussion, according to Stevenson, we express our moral attitudes, but we also attempt to get others to share them (thus his paraphrase, so easy to mock, of "This is good": I approve of this, do so as well!). Stevenson had no persuasive account of how on earth my saying such things could ever influence you to change your moral outlook.⁴⁰ Maybe if the challenge and the controversial feel are uncomfortable, I'd consider adapting just to remove the discomfort.

But somehow this story seems to get the cart before the horse. What we want to know is why my state of mind, when I think roller coasters are fun, disagrees with yours when you think they aren't fun. The answer cannot be the controversial and challenged feel I get when I learn that your state of mind is different. I will feel challenged only if I can think of your state as in disagreement with mine. And why am I supposed to do that?

Paul Horwich wonders, too.

I am sympathetic to the following suggestion of Allan Gibbard's. When someone genuinely accepts something and someone else accepts its negation there is a

sense of substantive (“not merely verbal”) disagreement; there is conflict, a clash, a feeling that the other person is somehow in bad shape . . .

This seems right as far as it goes; but one might hope for a somewhat deeper understanding. What is the nature of the clash that is manifested in contradictory beliefs but absent in [cases where psychological states merely differ]? Clearly it won’t do to say that it consists simply in one person believing something and the other denying it. For we are trying to explain why the states are taken to qualify as beliefs in terms of their being seen as linked with some deeper conflict, or in terms of our sense that others would be better-off sharing our own such states.⁴¹

I wonder about that, too. That is, it seems to me that once we start getting worried about why certain kinds of differences between your state of mind and mine count as disagreement while others are ‘mere differences’, it is easy to wonder the same about paradigmatic beliefs. Horwich has a suggestion.

Here’s a sketch of a possible answer (— again, it’s close to Gibbard’s). An essential property of our faculty of belief — its *raison d’être* — is the role it plays in determining how we are inclined to act. And the conflict associated with contradictory beliefs consists in their potential, through inference, to engender conflicting desires and decisions. If I disagree with you about the truth of some empirical proposition, <T>, then that can easily result (via theoretical reasoning and given other premises) in our disagreeing about the truth of some more directly action-guiding belief, <If A is done then X will occur>. And if we both want X to occur then one of us will, on that account, be in favor of A being done, and the other won’t. We might even come to blows! So can one see how divergent empirical beliefs might correlate with a practical tension.⁴²

The point here is to see whether Expressivism can make out a helpful sense in which differences in attitude — or *plan* — are to count as disagreements; once that sense is made out, Horwich suspects (as do I), the infamous difficulties that are supposed to arise in making sense of logic, inference, embedding, will be soluble.

Suppose that this story is along the right lines. What are its implications for emotivism? If the fundamental function of basic “ought” sentences is to express the speaker’s desires, will that imply that “ought” pronouncements could not articulate states relevant to decision and could never reflect genuine clashes — and so can’t qualify as expressions of belief? The answer would appear to be no. On the contrary, insofar as normative pronouncements tend to be associated with desires and decisions, then they manifest, in a peculiarly immediate way . . . , the feature that marks certain declarative pronouncements as expressions of belief.⁴³

Plans (to update the reference to Gibbard) do not even need the aid of fundamental or universal desires to lead to differences in action, of course. And if the aim of normative conversation is to coordinate actions, the ‘clash’ between plans that differ will strike us as in need of ironing out. By contrast, there is no

urgency to coordinating our “views about whether it’s throbbing in here,” that is to say, our headaches (and lacks thereof).

But there is a problem with this suggestion. To see what it is, let’s spell out how the ‘clash’ works in a typical example of normative disagreement à la Gibbard. Here’s the Good Case: You judge that saving the whales is what we ought to do while I judge that it is not worth doing and our resources ought to be directed elsewhere. So my aims are thwarted to the extent that you succeed, and conversely. This is a practical conflict of the clearest sort. But here is the Bad Case. There is only one dose of painkiller left and we each have a headache. You judge that people ought to forswear pharmaceutical relief from pain and tough it out, while I think you are mistaken — people ought to embrace the pain relief offered by medically tested drugs. Now my aims are met to the extent that you manage to act on yours. There is no practical clash. And indeed the clash would come precisely if we *agreed* that each of us ought to try to grab the Ibuprofen. Why, then, do we disagree in the first situation and agree in the second? The Practical Clash test gets the wrong answer (or at least risks getting it wrong) whenever the norms about which we could disagree are agent-centered norms; it is guaranteed to match our intuitive judgment about disagreement only when the subject is agent-neutral normativity.⁴⁴

Toward a solution?

So what is disagreement?

I can only point in a general direction. Start with some attitudes that we are comfortable thinking about as in disagreement with one another. Maybe we can only go so far in saying *why* they count as in disagreement; that’s acceptable, I’m suggesting, so long as we aren’t in serious doubt *that* they are. If we can identify some, then we’re off and running; our job will be to explain the puzzle cases in terms of the comfortable ones.

Here’s my paradigm: *preference*. Preference, as we ordinarily think of it, is unlike desire in an important way: your preferences can be incoherent, while your desires can be only conflicted. For instance, you might desire to present your views in front of a large and critical audience, but at the same time be terrified of doing so; you may want to drive across the country but want also to stay out of nauseous situations; and so on. When we are conflicted like this, we have to work out how to balance conflicting desires, but there is no necessity to be rid of any of them. On the other hand, suppose you prefer taking the last Thai dumpling off the plate to leaving it for someone else, but also prefer maintaining a polite status to acting rudely, and then you realize that the only way to maintain a polite status is to leave the remaining dumpling for someone else. This won’t do; your preferences (together with your beliefs) are incoherent. And there are many other, fancier ways of having incoherent preferences (they might fail to be transitive, or you could run afoul of a dominance constraint,

for example). When someone's preferences are incoherent, and she notices that they are, something's gone wrong. She 'has to change', in something like the sense that a believer 'has to change' her beliefs when she notices that they are inconsistent.

I don't have a satisfying explanation of why preference is subject to coherence constraints. It seems to me to have something to do with the fact that preference is a kind of model of choice or intention. It's too simple to say that preferences are dispositions to choose, since we have preferences that could not possibly be alternatives of our choices (like our preferences about the weather, for instance, or some alternatives that would be spoiled if we chose them, like the preference I have that someone throws a surprise party for me). But preferences seem to be in the same general family as choices, so that the constraints that limit coherent choice are inherited by choice's relative. This is nowhere near a theory, I know.

Now for the second stage. Maybe normative statements express preferences (or some other attitude similarly connected to choice). Then they could be subject to coherence constraints, too. And a pair of attitudes could be in disagreement if holding them together is incoherent. Of course, there is no incoherence in your preferring soup to salad and my preferring the converse; the idea is rather that I'll find your attitude unacceptable just in that I could not add it to my own without changing my mind about something or other.

This approach promises to help solve the Negation Problem.⁴⁵ You can have a settled view and still prefer neither writing invitations by hand nor printing them out; you could be indifferent. So indifference is a good candidate for Judith's attitude when she denies that you ought to write by hand but also that you ought *not* to write by hand. And being indifferent between a pair of alternatives is not the same as having no settled view about them at all (compare being indifferent between two brands of cola, on the one hand, with being undecided between having a soft drink and drinking plain water on a hot day), so there is no difficulty in distinguishing the attitude reported by "Judith thinks it is not the case..." and "It is not the case that Judith thinks..."

The kind of solution I'm trying out promises to carry over fairly well to the Relativisms. The idea is that disagreement resides not in the contents by themselves, but in the conditions under which it is appropriate to assert the sentences in question. For an Indexical Relativist, the analogy would be with John Perry's meander through the supermarket with a leaky bag of sugar: when he's in a position to assert, "I'm the one with the leaky bag," his new state of mind makes a substantial difference to his plan of action, even though he hasn't come to believe a new content, because the assertibility conditions of first-person sentences is different from those of third-person sentences that express the same proposition.⁴⁶ And similarly, the kinds of states that Expressivism claims to be *expressed* by normative sentences will be said by Indexical Relativists to be among the conditions of sincere assertion for those same sentences. The superficial conflict in assertions when you deny what I have asserted, then, counts as genuine disagreement because of what it shows about our attitudes (if we are

sincere). In short, the account of conflict of attitudes can be adopted by Indexical Relativism. (I don't have enough of a grip on Genuine Relativism to see how to generalize the solution. I imagine the story would connect disagreement in states of mind that constitute acceptance of a (centered) proposition, to disagreement between the acceptances.)

The Threat

Suppose we can't find any solution that satisfies. What would be wrong, in that case, with taking *disagreement* as our primitive? We could just stipulate that for certain kinds of beliefs and assertions, one person's rejecting what another accepts will count as disagreement; for others, not. After all, we are not in any real doubt that there is such a notion.

The problem is that if we cannot say anything by way of explanation, then we are hostage to the possibility that the intuitive notion of disagreement that we rely on, that we are taking for our purpose as primitive, is not friendly to the kinds of theories I am considering. Maybe the intuitive notion is this: when we can see, by our native grasp of our language, that your rejection and my acceptance of this certain sentence counts as disagreement, that is because we have a prior grasp on the idea that some sentence really express real propositions, while others 'merely' express our attitudes, or pick out our individual position (not necessarily shared by others) in some kind of logical space of 'outlook'. And this will spoil the projects of Relativism and Expressivism alike. If we can say nothing independently about disagreement, we have no assurance that the correct account is compatible with the explanatory priority that these theories give to planning or sensibility. That is a real possibility — at least for all I have shown. And it is reason enough to keep looking, even if the best attempts so far to explain what disagreement amounts to, are failures.

Notes

1. I describe a more complicated version, intended to fit more closely to our actual use and in large part designed to meet some of the Disagreement Problem, in "Internalism and Speaker Relativism", *Ethics* 101.1 (1990): 6–26.
2. The point is that a moral system in this sense is not a state of mind; of course, which moral system is yours is determined by your intentional states. To keep things close and parallel with Expressivism, we could say that your moral system takes 'wrong' to the class of things that you disapprove of in their various possible circumstances; a more plausible view would complicate matters but follow the same rough idea.
3. "Demonstratives", in *Themes From Kaplan*, Almog, et al., eds., Oxford (1989).
4. The truth in Internalism, I think, is more complicated. And the best version of Indexical Relativism will design a character for moral terms to match up

with the complications of Internalism. I take some first steps in “Internalism and Speaker Relativism”, and add some further thoughts in “Relativism and Nihilism”, *Oxford Handbook of Ethical Theory*, D. Copp (ed.), Oxford (2005).

5. See my “Internalism and Speaker Relativism”.
6. *Ethics*, London: Library of Modern Knowledge (1912). No doubt it’s a lot older than Moore.
7. There is a second problem that seems to be related: the problem of Indirect Reports. I’ll mention this issue below, but I cannot address it in this paper.
8. Cambridge, Mass.: Harvard University Press (2003).
9. Of course, Indexical Relativism does *not* endorse the most straightforward deflationary answer to the truth condition question. When you ask me for the truth conditions of Alistair’s assertion of the sentence “Cannibalism is wrong”, I cannot tell you that it is true just in case cannibalism is wrong, according to Indexical Relativism, because when I assert the sentence it has different truth conditions. This is obvious enough; compare other indexical sentences. When asked for the truth conditions of Churchill’s sentence, “I cannot forecast to you the action of Russia”, I had better not tell you that it was true iff I cannot forecast to you the action of Russia.

However, a deflationary answer is still available; see, e.g., Hartry Field’s “Deflationism about Meaning and Content”, esp. §10 (pp. 134–6 in *Truth and the Absence of Fact*).

10. See W. D. Falk, “‘Ought’ and Motivation”, *Proceedings of the Aristotelian Society*, 48 (1947): 492–510 for the Internalist idea; M. Smith *The Moral Problem*, Blackwell 1994, for its transformation into the Practicality Requirement.
11. Horwich, “The Motive Power of Evaluative Concepts”, forthcoming in his collection, *Truth — Meaning — Reality*. R. M. Hare makes just the same point in “Meaning and Speech Acts,” *The Philosophical Review* 79 (January 1970): 3–24.
12. Note that a hyperplan is a formal object; the psychological state of planning is represented by a set of hyperplans, just as the psychological state of believing is represented by a set of possible worlds. In both cases there can be ambiguity: when I ask what Cynthia’s plan is, you can cite the formal objects, saying things like *in case of rain she will take a cab*, and it would sound like a joke if I replied, No, Cynthia’s plan is a state of her brain. The same ambiguity lurks in talk of belief. Cynthia’s belief is *that it will rain*; Cynthia’s belief is a state of her brain; but that it will rain is not a state of Cynthia’s brain.
13. In “Expressivist Embeddings and Minimalist Truth”, *Philosophical Studies* 83:1, 29–51, I called Gibbard’s solution the NutraSweet solution, because it was designed to fit into ready-made apparatus, namely possible world semantics, in a way reminiscent of the way NutraSweet was designed to fit into our ready-made apparatus, namely our taste buds.
14. “Indexical Relativism versus Genuine Relativism”, *International Journal of Philosophical Studies* 12 (2004): 297–313.
15. Neither MacFarlane nor Egan endorses a Genuine Relativist view of metaethics. Still, both of them think that Genuine Relativism helps to solve the Problem of Disagreement, so the remainder of my discussion is relevant to their views even though I’m talking about ethics and MacFarlane and Egan are more interested in, e.g., predicates of taste and epistemic modals.

MacFarlane doesn't consider all the "Genuine Relativisms" discussed here to be Genuine. For MacFarlane, a theory is Genuine Relativist only if it includes *assessor relativity*; see below in section

16. "Relativism and Disagreement" pp. 21–2.
17. It *seems* so; according to MacFarlane's view, though, that you accept the proposition I reject does not entail that we disagree. See section 6, below.
18. Egan uses the elements of this formal semantics for analysis of talk of secondary qualities, of predicates of taste, and of epistemic modals; see his "Secondary Qualities and Self-Location", *Philosophy and Phenomenological Research* 72.1 (2006): 97–119. I think Egan does *not* endorse Genuine Relativism for moral thought and talk.
19. See my "Transforming Expressivism", *Noûs* 33.4 (1999) : 558–572 for details.
20. For further development, including the explanation of how the functions work in composition and inference, see my "Transforming Expressivism".
21. Egan notes:

It's important to notice the difference between centering features and predicates with hidden indexicals. When I attribute a property to something using a predicate with a hidden indexical, I'm still attributing a *property*, and so I'm still expressing a possible-worlds proposition. It's just that which property I attribute to things with a use of the predicate varies from context to context. If we had a predicate that expressed a centering feature, sentences in which it occurred (in the usual way) would express *centered-worlds* propositions. ("Secondary Qualities and Self-Location", n. 31, p. 109)

But this is not a difference between formal objects; it is a difference in how they are used in the theory. I'll return to this point shortly.

22. MacFarlane, "Relativism and Disagreement", *Philosophical Studies* 132 (2007), 17–31.
23. Egan, "Secondary Qualities", fn 31, p. 109.
24. "Indexical Relativism versus Genuine Relativism", 307.
25. See my "Negation for Expressivists" in *Oxford Studies in Metaethics*, R. Shafer-Landau (ed.), Oxford (2006), for a fuller development of the Negation Problem.
26. Thanks to Terry Horgan for some discussions that made me see that I needed to clarify this point.
27. Egan's model builds on Lewis's model for attitudes *de se*; the formal resemblance will be obvious to those familiar with Lewis.
28. "Secondary Qualities" p. 114.
29. "Relativism and Disagreement", pp. 22–3.
30. "Indexical Relativism versus Genuine Relativism", p. 307.
31. Ragnar Francén, "Metaethical Relativism: Against the Single Analysis Assumption", *Acta Universitatis Gothoburgensis* (PhD. dissertation), Gothenburg, Sweden (2008): 112.
32. *Ibid.*, p. 112.
33. It is tempting to put it this way: what Jane said is true and so is what June said. But this is a tricky issue — what June said, after all, is not *actually* true, and so I can't really say that it's true. It would be true if the world were like the world in the story of June. Likewise, what you said about Joe at 2:00 is not true, *though it*

was at the time you said it. I don't insist on this way of talking, but it does seem fairly natural to me.

34. "Relativism and Disagreement", p. 18.
35. *Ibid.*, pp. 28–9.
36. *Ibid.*, p. 29.
37. Related: suppose that whenever I have an occurrent belief that there is no recursive and complete axiomatization of arithmetic, I get a headache. Noticing this disposition, shall I conclude, "If arithmetic is incomplete then it is throbbing in here"?
38. See Gibbard's *Thinking How to Live*, Harvard University Press (2003), especially Chapter Four, for development of the idea that disagreement in states of mind is the hook on which to hang a full blown semantic theory.
39. "The Emotive Meaning of Ethical Terms", originally in *Mind* 1937; reprinted in Stevenson, C. L., *Facts and Values*, Yale University Press (1963).
40. His best try, I think, was to point out the parallel of a 'persuasive declaration', as when a parent tells a child, "*We* do not approve of lying." But moral discussion among peers is in this way unlike the didactic declarations of parents to their children.
41. "The Motive Power of Evaluative Concepts", *op cit.*
42. *Ibid.*
43. *Ibid.*
44. A similar problem arises for MacFarlane's suggestion, I think, though I won't rerun the argument in detail. MacFarlane suggests that we'll find disagreement when we have an interest in coordinating attitudes; attitudes that won't coordinate well count as being in disagreement. But this simply isn't right when the attitudes are centered preferences, desires, and so on.
45. I show how in "Negation for Expressivists".
46. See John Perry, "The Problem of the Essential Indexical", *Nous* 13 (1979): 3–21.