Non-Standard Features

by

Andrew M. Egan

B.A. Philosophy
University of Wisconsin-Madison, 1993

M.A. Philosophy
University of Colorado-Boulder, 1999

SUBMITTED TO THE DEPARTMENT OF LINGUISTICS AND PHILOSOPHY IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2004

© 2003 Andrew M. Egan. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper
and electronic copies of this thesis in whole or in part.

Signature of Author: _

Department of Linguistics and Philosophy
December 18, 2003

Certified by: _

Stephen Yablo
Professor of Philosophy
Thesis Supervisor

Accepted by: _

Vann McGee
Professor of Philosophy
Chair, Committee for Graduate Studies
ABSTRACT:

The dissertation is composed of three papers on properties and their relatives. "Second-Order Predication and the Metaphysics of Properties" argues that giving a happy account of second-order predication motivates us to identify properties with functions from \( <\text{world, time}> \) pairs to extensions rather than with the sets of their instances. "Secondary Qualities and Centering Features" offers a characterization of the elusive distinction between primary and secondary qualities. "Appearance Properties" argues that Sydney Shoemaker's proposal for reconciling intentionalism with the possibility of spectrum inversion without error fails in its details, but there is a nearby proposal which may well succeed.
Acknowledgements:

Thanks are due to very, very many people for helpful discussions of the ideas in this dissertation—far more than I could list here. I owe especially large debts to Tyler Doggett, Adam Elga, Liz Harman, and Daniel Stoljar, and to the members of my committee: Alex Byrne, Ned Hall, Steve Yablo, and Bob Stalnaker.
Contents:

Second Order Predication and the Metaphysics of Properties.........................9
Secondary Qualities and Centering Features..............................................41
Appearance Properties?.............................................................................75
Second-Order Predication and the Metaphysics of Properties

Introduction

If you’re doing systematic metaphysics, problems for, and revisions to, one part of your theory will tend to have ramifications elsewhere. This paper is about a problem for Lewis’s theory of properties. It’s also about what happens to the Lewisian system when we revise the theory of properties in order to avoid the problem. So I’ll be concerned to do two things in what follows. First, I’ll argue that Lewis’s theory of properties has a fatal problem with accounting for second-order properties—the properties of properties. Second, I’ll trace the ramifications of the revision through the rest of Lewis’s system.

1. The Problem

An attractive nominalist strategy is to identify properties with the sets of their instances—not their actual instances (the property of being round isn’t the property of being blue, even if all and only the round things are blue)—but all of their instances, actual or not.¹ This is the theory that Lewis (1986a, 1983b) endorses.²

¹ Sets, or classes? Well, Lewis says that he means sets (Lewis 1986a:50 fn37). But this runs into problems: being a set, for example, seems like a perfectly good property, but there isn’t (on the standard ways of doing set theory) a set of all sets. So if properties have to be sets, being a set turns out not to be a property after all. (Classes run into other problems—being non-self-membered seems like a perfectly good property, too.) Since I won’t be concerned with these problems, and nothing will hang on the set/class distinction in what follows, we needn’t resolve this question here. (Thanks to Vann McGee for pressing me on this.)

² There are a couple of extremely contentious bits of Lewisian metaphysics that I’m going to treat as if they were uncontroversial. The first of these is Lewis’s modal realism. The other is that some sort of reductive nominalist theory of properties is right. The sorts of questions that are addressed in the remainder of the paper will still arise for many philosophers who reject one or both of these theses. First, while the questions only arise if you’re some kind of modal realist, but they don’t depend on merely possible worlds being the big concreta that Lewis takes them to be. Second, while anti-nominalists won’t want to identify properties with set-theoretical entities built out of possibilia, properties still need to be intimately associated
It's a straightforward consequence of this theory that things that exist in more than one world can't have any of their properties accidentally. Here is why not:

Suppose (contra Lewis) that people aren't worldbound—that they exist in more than one world. Elmer is a philosopher, but he might have been a plumber instead. So he's only accidentally a philosopher. Though he's a philosopher in the actual world @, he's not a philosopher in some other possible world w. That is: in @, Elmer has the property, being a philosopher, while in w, Elmer lacks the property, being a philosopher. Suppose also (with Lewis, this time) that properties are sets of possibilia. Something has the property just in case it's a member of the set, and lacks the property just in case it's not a member of the set. So since Elmer is a philosopher in @, he must be a member of the set, being a philosopher. Since Elmer is not a philosopher in w (since, in w, he lacks the property, being a philosopher), Elmer must not be a member of the set, being a philosopher. So Elmer must both be and not be a member of the same set. Contradiction.

The reason why this kind of case isn't a problem for Lewis is also straightforward. Lewis takes philosophers to be worldbound, and to have their modal properties in virtue of the behavior of their (distinct) counterparts in various other worlds. And while Elmer can't both be and not be a member of the same set, there's no problem about Elmer being a member of some set that not all of his counterparts are members of.

with some such entities. Whatever properties are, they at least need to determine, for example, a class of all of their possible instances.

Now consider a parallel, though less familiar, case. Elmer has a favorite property. It’s being green. But Elmer is fickle—he might have favored some other property instead. So being green might not have been Elmer’s favorite property. In fact, being green might not have been anybody’s favorite property. So being green is somebody’s favorite property, but it might not have been. Here’s another way of saying the same thing: being green has the property, being somebody’s favorite property, but only accidentally. This is a case of second-order predication—of attributing properties to properties. More specifically, it’s a case of contingent second-order predication—of attributing accidental properties to properties. Since being somebody’s favorite property is itself a property, it’s the set of its (actual and possible) instances. Its instances are properties, so being somebody’s favorite property is a set of properties.

Let @ be the actual world, in which being green is Elmer’s favorite property, and let w be a world in which Elmer (along with everyone else) has turned his attentions elsewhere. Since being green is somebody’s favorite property in @, it must be a member of the set, being somebody’s favorite property. Since being green is not anybody’s favorite property in w, it must not be a member of the set, being somebody’s favorite property. Contradiction. And since properties aren’t worldbound and therefore don’t have distinct counterparts in other worlds, counterpart theory can’t help us here.4

So here’s the argument: If properties are the sets of their instances, then properties can’t have any of their properties accidentally. But properties plainly can have some of their properties accidentally—witness being somebody’s favorite property, playing the

4 Mark Heller (1998) has proposed a theory on which properties do have distinct counterparts in other worlds (in response to a very different problem). This sort of theory will, obviously, avoid the problem that I’m raising here.
pain role, and being the semantic value of 'red', all of which are accidental properties of any property that has them. So properties aren’t the sets of their instances.\(^5\)

2. Objections and Responses

Isn’t there some sort of paraphrase strategy available that’ll make this problem go away? I don’t think that there is. At the very least, there’s not one that’s worth the trouble, given that there’s another solution available (presented in the next section) that gives a smooth treatment of the problem cases. Notice that it’s a constraint on such a solution that it deny that any second-order properties are ever had contingently. As long as there are any accidental second-order properties, properties just can’t be the sets of their instances. So the responses that attempt to retain the Lewisian picture of what properties are will all be strategies for doing away with accidental second-order properties and finding something else to do the work of making what look like attributions of accidental properties to properties come out true, despite the absence of the relevant accidental properties.

One thing that we might try, in order to make all of the properties of properties necessary, is world-indexing all of the second-order properties.\(^6\) So there’s no such property as being somebody’s favorite property—instead there’s a family of world-

\(^5\)Some more examples of accidental second-order properties: being instantiated, being coinstantiated with being green, and being the subject of extended philosophical debate. Some slightly fancier ones: the second-order properties you get by lambda abstraction from statements of contingent natural laws, and the second-order properties you get by dropping one of the quantifiers over properties in a Ramsey sentence, like the ones that (we hope) provide the characteristic functional roles of mental properties like being in pain (being in pain is the property G such that there are properties F\(_1\), F\(_2\)…F\(_n\) such that…).

\(^6\) Since properties are arbitrary sets of things, there won’t actually be a clean distinction between first-order and second-order properties, since there will be properties that could be had by, for example, both being green and my desk. These properties actually needn’t even be especially weird or unfamiliar: being mentioned in this footnote seems to do the trick. Still, there will be some properties that can only be had by particulars, and some that can only be had by properties—call these the pure first-order and second-order
indexed properties, *being somebody’s favorite property at w₁, being somebody’s favorite property at w₂*, etc. These properties are had necessarily if at all, so we needn’t say that properties have any of their properties accidentally. We can explain the appearance of having accidental properties by noting that *being green* has some, but not all, of these world indexed properties—a fact naturally expressed by saying that *being green* is somebody’s favorite property in some worlds, but not in others.

This move fails because it assigns the wrong contents to sentences involving second-order predication. When I say, ‘*being green* is somebody’s favorite property’, two things have happened: I’ve attributed some property to *being green*, and I’ve said something (metaphysically) contingent. But suppose the property I’ve attributed to *being green* is a world-indexed property. Then I haven’t said something contingent. What I’ve said must be either necessarily true or necessarily false. So since the contents of sentences like ‘*being green* is somebody’s favorite property’ are contingent, the second-order properties attributed in them can’t be world-indexed.

---

7 Or even, ‘*being green* is Elmer’s favorite property’. I focus on the case of ‘quantified’ properties like *being somebody’s favorite property* not because the points only go through for these sorts of properties, but because showing that they go through for properties like *being Elmer’s favorite property* requires a fair amount of house-to-house fighting about the details of counterpart theory.

8 It’s possible for the objector to make a further move at this point. She could bite the bullet, accept that the contents of such sentences are necessary, and attempt to explain the appearance of contingency by a diagonalization strategy. Even if second-order properties are all world-indexed, in different worlds the predicate ‘*is somebody’s favorite property*’ will express different properties (in w₁ it expresses the property, *being somebody’s favorite property in w₁*, in w₂ it expresses the property, *being somebody’s favorite property in w₂*, etc.) and so ‘*being green* is somebody’s favorite property’ will express different propositions depending on which world it’s uttered in. In every world, the proposition expressed by local utterances of ‘*being green* is somebody’s favorite property’ will either be necessarily true or necessarily false, but since the propositions expressed will be different in different worlds, we will get world-to-world variation in whether local utterances of ‘*being green* is somebody’s favorite property’ are true or false. I think that this move is unattractive. For one thing, it gets the facts about entailment wrong. The content of my utterance of, ‘*being green* is somebody’s favorite property’ ought to be entailed by the proposition that *being green* is Foghorn’s favorite property, but ought not to be entailed by the proposition that Foghorn is a rooster. If the diagonalizer is right, the content of my actual utterance is necessary, and so is entailed by everything. (Thanks to Sally Haslanger for discussion of this point.) Anyway, there’s another theory of
Another tempting strategy is based on the very plausible idea that the only accidental features of properties are which *relations* they stand in to other things. This suggests a way to avoid the problem—we could get the right sentences to come out true by exploiting the unproblematic contingent *relations* between properties and other things, while denying that they have any troublesome accidental *properties*.9

Tempting though it is, this strategy doesn’t work, because the contingent relations *are* problematic—accepting the contingent relations also commits us to accepting the accidental properties.

It’s non-negotiable that there are contingent facts about which relations properties stand in to other things (including other properties). If we’re allowed to move from the contingent facts about which *relations* properties stand in to contingent facts about which *relational properties* they have, then the jig is up, since the Lewisian theory of properties can’t allow that there are any contingent facts about the properties (relational or otherwise) of things that don’t have distinct counterparts.

We plainly can move from the fact that Elmer is hunting Bugs (that Bugs stands in the *being hunted by* relation to Elmer) to the fact that Bugs has the relational property, *being hunted by Elmer* (and therefore that he has the property, *being hunted by somebody*). And if we can do that, we ought to be able to move from the fact that *being

---

9 This move may look like a non-starter, because of the existence of contingent relations between properties (*being systematically coinstantiated with, being related in lawlike way* L to, etc.). Perhaps this really is a fatal problem with the proposal. It might be, though, that whenever there two properties stand in a contingent relation to one another, it’s because the properties stand in certain contingent relations to
green is Elmer’s favorite property to the fact that being green has the relational property, being Elmer’s favorite property, and therefore has the property, being somebody’s favorite property. And since it’s contingent that being green is anybody’s favorite property, it’s contingent that being green has the property, being somebody’s favorite property. Now we have our counterexample to the claim that properties are the sets of their instances: a property that has a property, and has it contingently.\(^\text{10}\)

Notice that any story we tell in order to avoid attributing accidental properties to properties will require us to say that the move from relations to relational properties is illegitimate. We’ll be forced to say, for example, that being green can stand in the being the favorite property of relation to Elmer without having the relational property, being Elmer’s favorite property. I suspect that this is incoherent. But in any event, whether it’s coherent or not, it’s a desperate move.

Lewis says some things that suggest a different line of response. In On the Plurality of Worlds, Lewis says that ‘[a] universal can safely be part of many worlds because it hasn’t any accidental intrinsics,’\(^\text{11}\) and ‘[i]f indeed there are no accidental worldbound things. In any case, whether this is true or not, there is an independent, fatal problem for the proposal.

\(^\text{10}\) Here is almost the same argument, cast more formally: \(\text{'Rab}' \text{ is true iff } \lambda x (Rx b) a \text{ is true. And } \lambda x (Rx b) a \text{ is true iff a has the relational property that's the semantic value of } \lambda x (Rx b). \text{ So if (a) there are some contingent sentences of the form } \text{'Rab}' \text{ where one of the relata is a property, (b) we can do the usual kind of lambda abstraction on sentences of the form } \text{'Rab}' \text{, then properties can't be the sets of their instances. (Since we get some contingent facts about the properties of properties.)}

The properties we get by this procedure are ones like being Elmer’s favorite property, rather than being somebody’s favorite property, but (a) those properties will do just fine as counterexamples (see note 8), and (b) we can also get the ‘quantified’ properties if we want them by adding another innocent-seeming step to the argument, in which we existentially generalize before doing the lambda abstraction: \(\text{'Rab}' \text{ is true only if } \exists y (Rxy) \text{ is true, which is true iff } \lambda x (\exists y (Rxy)) a \text{ is true. So in order for } \text{'Rab}' \text{ to be true, a certain object must have the property expressed by } \lambda x [\exists y (Rxy)] a.\)

\(^\text{11}\) Lewis (1986: 205 fn6).
intrinsics to raise a problem, then overlap confined to the sharing of universals seems entirely innocent'.

It looks like the claim is that there's an important difference between accidental intrinsic properties and accidental extrinsic properties—there's a big problem if would-be transworld objects (like universals) have accidental intrinsics, no problem if their only accidental properties are extrinsic or relational.

But the mere fact that the relevant properties aren't intrinsic can't make a difference. The problem is that (a) there are things without distinct counterparts (namely properties) which plainly have (second-order) properties, and have them accidentally, and (b) Lewis's theory of properties cannot accommodate this fact. That the second-order properties are extrinsic is beside the point; the thing that's making the trouble is just that they're properties. When Lewis says that properties are the sets of their instances, he's offering a perfectly general theory of properties, not a theory that's only supposed to apply to some restricted subset of all the properties that there are. The claim that properties are the sets of their instances is supposed to apply to all properties, not just the intrinsic ones.

12 Lewis (1986: 205). The other relevant passage is Lewis (1983b: 11, fn5). In these passages, Lewis is talking about universals, which would be points of overlap between worlds—the very same thing would be part of more than one world, and not by having different parts in each. This is unlike Lewisian properties, which are not meant to be points of overlap. They do draw their members from many different worlds, and so it's natural to say that they have parts in various different worlds (especially for Lewis—see Lewis (1986b), (1990)). However, at least for first-order properties, a property's part in one world will be wholly distinct from its part in any other world. But this difference between properties and universals is not relevant to our concerns. What's important to generating the problem that Lewis is worried about for universals, and that I've claimed actually arises for Lewisian properties, isn't overlap, but the absence of distinct counterparts in different worlds. The problem is that universals, being present in more than one world, are always their own counterparts. So, it can't be that a universal is a member of some set and one of its otherworldly counterparts is not. This feature—the lack of distinct counterparts—is a feature that universals and Lewisian properties share.

13 Steve Yablo (1998) says some similar-sounding things.
In light of this, I think that Lewis’s (very brief) discussion of accidental intrinsics is best understood not as a new line of argument, but as a gesture in the direction of something like the ‘get rid of the properties and make do with the relations’ strategy discussed above.

Summing up: we can’t paraphrase away accidental second-order properties without saying implausible things about the connection between relations and relational properties and/or assigning the wrong contents to sentences involving second-order predication. We also can’t appeal to the bare fact that the accidental properties of properties are all extrinsic in order to solve our problem. Fortunately there is another theory of properties right around the corner, in the same spirit as Lewis’s theory, which handles second-order predication without incident.

3. The Replacement Theory

Perhaps surprisingly, there’s no parallel problem for a theory that identifies properties with functions from possible worlds to extensions (where extensions are just sets of things).¹⁴

Here’s how the functions account handles the fact that being green is somebody’s favorite property, but only accidentally: At @, being green is somebody’s favorite property. At w, it’s not. The second-order property, being somebody’s favorite property, is a function that, for each world taken as argument, delivers as value an extension, which will be a set of properties. Call the value of being somebody’s favorite property taking @

---

¹⁴ This is potentially surprising because it’s tempting, if you’re only thinking about the properties of worldbound things, to take the two theories to be notational variants. Of course, Lewis didn’t take them to be mere notational variants. One of the problems about offering Lewis a properties-are-functions theory as
as argument ‘A’, and call the value taking w as argument ‘B’. Since being green is somebody’s favorite property in @, being green must be a member of A. Since being green is not anybody’s favorite property in w, being green must not be a member of B. No problem. All we get from this is the unsurprising result that A≠B; that being somebody’s favorite property delivers different extensions for @ and w.

So while there’s a problem about contingent second-order predication if we take properties to be the sets of their possible instances, there’s no problem if we take them to be functions from worlds to extensions. I conclude that Lewisians should take properties to be functions from worlds to extensions, rather than the sets of their possible instances.

This lets us avoid the problem about accounting for the accidental properties of properties while still telling a uniform story about first-order and second-order predication: something has a property F at a world w iff it, or its counterpart at w, is a member of the value of F taking w as argument. It has F essentially iff it has F at every world (or at every world at which it has a counterpart), and it has F accidentally iff it has F at @, and has the complementary property (it’s in the antiextension of F) at some other world.15

We can still build properties out of antecedently familiar things, and we can still have plentiful, arbitrary, gerrymandered properties, and a distinction between the more and less natural properties. It’s as easy to have arbitrary, gerrymandered functions as it is to have arbitrary, gerrymandered sets of things, so there’s no difficulty in accommodating extremely unnatural properties. There’s also no more of a problem singling out the

---

15 A way to avoid the problem about contingent second-order predication is that he offers an argument against identifying properties with functions in Lewis (1986a). We’ll get to this in the next section.
natural properties if we take them to be functions than if we take them to be the sets of their instances. The natural properties can still be the ones sharing of which makes for genuine similarity.\textsuperscript{16}

So taking properties to be functions from worlds to extensions does away with the problem about contingent second-order predication, and preserves the benefits of Lewis’s theory. So far, so good.

4. Functions, Properties, and Relations

Despite these benefits, there’s a worry about the acceptability of this substitute theory of properties. Lewis offers an argument that we ought not to take properties to be functions from worlds to extensions. Here I paraphrase his argument:

It’s tempting to say that there are properties that things don’t have or lack \textit{simpliciter}, but only relative to this or that other thing. Since I’m thirsty at some times and not at others, I don’t just have (\textit{simpliciter}) the property \textit{being thirsty}. Instead I have the property \textit{being thirsty} relative to some times, but not relative to others. Since the road is surfaced in some places but not in others, the road doesn’t have (\textit{simpliciter}) the property, \textit{being surfaced}. Instead it has the property \textit{being surfaced} relative to some locations, but not relative to others. Since Ted is the father of Fred, but not of Ed, he doesn’t have (\textit{simpliciter}) the property, \textit{being a father}. Instead he has the property \textit{being a father} relative to some people (like Fred) but not relative to others (like Ed). Similarly, since nine numbers the planets in the actual world but not in every possible world, nine doesn’t have (\textit{simpliciter}) the property, \textit{numbering the planets}. Instead it has the property \textit{numbering the planets} relative to some worlds, but not relative to others.

If you think that things don’t have or lack their properties \textit{simpliciter}, but only relative to this or that time, location, person, or world, then you should think that properties are functions from various kinds of things to extensions. You should think that \textit{numbering the planets} is a function from worlds to extensions, \textit{being thirsty} is a function from world-time pairs to extensions, \textit{being a father} (on one reading) is a function from individuals to extensions, and \textit{being surfaced} is a function from locations to extensions.

But we shouldn’t think any of this, because we should think that things have or lack their properties \textit{simpliciter}. A ‘property’ which is had only relative to this or that other thing is, whatever formal apparatus we use to describe it, not a \textit{property} but a \textit{relation}. So what the proponent of identifying properties with functions from things of one sort or another to extensions is proposing is really \textit{doing away} with genuine

\textsuperscript{15} The complication about having the complementary property in some world (rather than just failing to have the property in some world) is there in order to allow things whose existence is contingent to have some of their properties essentially.

\textsuperscript{16} A complication: for modally extended objects, sharing of natural properties will make not for similarity \textit{simpliciter}, but for a sort of world-indexed similarity.
properties and leaving us with only relations in their place. But it's obvious that things do have properties—it isn't all relations. And so we should have a theory of properties that allows things to have or lack them simpliciter, rather than merely relative to this or that other thing.17

At a first pass, this is an appealing argument. But its appeal is based on the assumption that functions from locations, people, world/time pairs, and worlds to extensions are all in the same boat, and will sink or swim together as candidates to be the properties. Since it's clear that functions from locations to extensions, or from people to extensions, can't be properties, it must be that they all sink. But the various candidates that Lewis mentions aren't all in the same boat. Not all functions are equally good candidates to be properties.

Lewis tells us, 'in order to say what a meaning is, we may first ask what a meaning does, and then find out what does that.'18 This isn’t a special principle about meanings. We can use the same procedure for other kinds of things, too. What properties do (among other things, but first and foremost) is provide semantic values for predicates.19 And what the semantic values for predicates do is determine an extension at each world. (More carefully: they combine with the semantic values of names, quantifier phrases, etc. to yield propositions. If we take propositions to be functions from worlds to truth values, then the semantic values of predicates need to determine an extension for each world)

If that's what properties do, then functions from worlds to extensions do it better than sets of instances. Such functions can supply the semantic values of predicates that apply to properties, while sets of instances can supply semantic values only for predicates that apply exclusively to worldbound individuals.

17 From Lewis (1986a:52-53). All of the examples are his.
Functions from locations or people to extensions aren’t good candidates to be the properties because they aren’t good candidates to be the semantic values of predicates. They aren’t good candidates to be the semantic values of predicates because they don’t do the thing that the semantic value of a predicate needs to do: determine an extension at each world. Functions from worlds to extensions are, unsurprisingly, ideally suited to the role of determining an extension at each world. So they’re very good candidates to be the properties. It’s not being a function that makes functions from locations, people, etc. to extensions ineligible to be properties. It’s being the wrong kind of function.

5. Some Consequences

Counterpart Theory

If we accept the revised theory of properties, we lose one of the arguments for counterpart theory. One reason to be a counterpart theorist is to avoid a problem about contingent first-order predication. If we take properties to be the sets of their actual and possible instances, then we have to be counterpart theorists, since otherwise the modally extended tables, chairs, and philosophers couldn’t have any accidental properties. Once we take properties to be functions from worlds to extensions, we no longer have this reason to adopt a counterpart theory instead of a theory of transworld individuals.

Of course, there are other reasons to be a counterpart theorist. So the revision of Lewisian metaphysics I’ve suggested certainly doesn’t require us to abandon counterpart theory. The surviving motivations—the implausibility of ordinary objects being only

---

20 What about functions from <world, time> pairs to extensions? How good are they as candidates to be the properties? We’ll come back to this in sections 6 and 7.
21 See Lewis (1971, 1986a) for some of them.
partly actual (or being points of overlap between worlds), skepticism about whether there are any deep facts about the essences of things, independent of our concerns or the particular ways in which we refer to them, a desire to identify statues with the clay that constitutes them, or persons with their bodies—seem to be the stronger, and the more philosophically interesting, motivations anyway. But still, it's an interesting fact that one of counterpart theory's selling points doesn't survive the revision. If the revised Lewisian is going to be a counterpart theorist, it won’t be because they’re compelled by their theory of properties. Instead it will be due to some slightly more elaborate, and less obviously conclusive, argument about essentialism, the implausibility of ordinary objects being that widely scattered, etc.

Events

One place where the revision makes things better for Lewis’s metaphysics is in the theory of events. Lewis (1986b) identifies events with properties of spacetime regions. It should come as no surprise that this view has trouble accounting for the fact that events have some of their properties accidentally. The problem is the one you’d expect. Lewis identifies events with properties of spatiotemporal regions. So, given his theory of properties, he identifies them with sets of possible regions. But this sort of modally scattered set isn’t the kind of thing that has different counterparts in different worlds, and so, if the only story about having properties accidentally is the counterpart account that applies to ordinary worldbound things, events can’t have any accidental properties. But of course they do have accidental properties, so there’s some fancy footwork to be done in order to make the theory plausible.
Much of ‘Events’ is spent in an attempt to give an account of what it is for an event to have a property accidentally—more carefully, of what it is that makes apparent attributions of accidental properties to events true. As Lewis leaves it at the end of the paper, the account is remarkably complex and counterintuitive, and, as Lewis notes, incomplete—there are still accidental properties of events that are left unaccounted for. This is one of the major blemishes on Lewis’s theory of events. It’s easy to think that a theory of events that needs to take such desperate measures in order to account for the straightforward fact that yesterday’s football game involved Brett Favre (featured 700 yards of total offense, came down to the wire, etc.), but might not have, can’t possibly be right.

We can now see that the difficulty about accidental properties that Lewis was struggling with in ‘Events’ is an instance of a much more general problem with his theory of properties. And since we have a solution to the general problem about properties, we also have a solution to the specific problem about events. None of the elaborate maneuvers in ‘Events’ are necessary if properties are functions from worlds to extensions. So we can keep the core of Lewis’s theory of events—that events are properties of regions of spacetime—without buying ourselves a big problem about how events can have some of their properties accidentally.\footnote{Note for aficionados of events: There are actually two problems for Lewis’s theory about the properties of events, and the revised theory of properties that I’ve suggested only solves one of them. (Though I think it’s the more serious one.) Lewis wants properties of events, in the end, to be explained in terms of the (more or less intrinsic, more or less natural) properties of the possible regions where the events occur. There shouldn’t be any ‘free floating’ properties of events. So one problem is: what’s the relation between the properties of events and the properties of the (actual and possible) regions in which they occur? The revision to the theory of properties is no help at all here. The other problem, though, is: can the Lewisian allow that, however events come by their properties, they have some of them accidentally? If the answer to this is ‘no’, then the theory is dead in the water. If the answer is, ‘yes, but only by saying a bunch of really
Let’s take stock of what’s happened so far. I’ve argued that it won’t do to identify properties with the sets of their instances, because if we do, then we can’t account for the fact that properties have some of their properties accidentally. I’ve also argued that the way to avoid this problem is to identify properties with functions from worlds to extensions. Accepting this revision has two interesting consequences: first, we lose one of the arguments for counterpart theory; and second, we avoid the biggest problem for Lewis’s theory of events. The rest of the paper is about an argument that we need to make a further revision—that we actually ought to identify properties with functions from <world, time> pairs to extensions—and the consequences of accepting it.

6. Worlds Enough, or Times?

Functions from people, locations, etc. to extensions are bad candidates to be the properties. Functions from worlds to extensions are better candidates. What about functions from <world, time> pairs to extensions? There’s an argument for identifying properties with functions that take <world, time> pairs rather than just worlds as arguments, parallel to the argument in section 1.

Here is a way that taking properties to be functions from worlds to extensions could get us in trouble: At a certain time \( t_1 \), Sylvester is sitting, and therefore has a bent shape. But Sylvester isn’t always sitting. At some other time \( t_2 \), Sylvester is standing on a rickety stepladder trying to reach Tweety’s cage. So at \( t_1 \), Sylvester is bent, and at \( t_2 \) Sylvester is not bent. That is: At \( t_1 \), Sylvester has the property, *being bent*, and at \( t_1 \) he

elaborate, counterintuitive stuff, then the theory is at least in serious trouble. And this problem goes away when we say that properties are functions from worlds to extensions, rather than the sets of their instances.
lacks the property, being bent. Suppose properties are functions from worlds to extensions. Then being bent is one such function—call it ‘F’. Then, since Sylvester is bent at \( t_1 \), he must be a member of \( F(\varnothing) \). And since Sylvester is not bent at \( t_2 \), he must not be a member of \( F(\varnothing) \). Contradiction.

If this argument sounds familiar, that’s because it is. It’s essentially Lewis’s argument from temporary intrinsics for the existence of temporal parts. (Rather, it’s essentially the first part of that argument—the rest is ruling out the other candidate solutions to the problem.) Lewis, of course, avoids this problem by adopting a metaphysics of temporal parts.\(^\text{23}\) Then there is no contradiction, because for Sylvester to be bent at a time \( t \) is for Sylvester’s temporal part at \( t \) to be bent—that is, for Sylvester’s temporal part at \( t \) to be a member of \( F(\varnothing) \)—and there is no problem about some of Sylvester’s temporal parts being members of \( F(\varnothing) \) while others are not.\(^\text{24}\) This would-be problem about Sylvester’s temporary properties, and its solution, are exactly parallel with the would-be problem about Elmer’s accidental properties, and its solution. Different times play the role of different worlds in the attempt to make trouble, and temporal parts play the role of otherworldly counterparts in dissolving the problem.

In the earlier argument, we were still able to make trouble by finding a sort of entity that had accidental properties, but lacked counterparts. Here again, we can revive


\(^\text{24}\) A complication: Sylvester has lots of different temporal parts that exist at \( t \). (The one that starts at \( t \) and ends three seconds later, the one that starts five minutes before \( t \) and ends five minutes after, the one that starts twelve seconds after Sylvester’s birth and ends at his death, etc.) Some of these are also bent at some times and not at others. So for ‘Sylvester’s temporal part at \( t \)’, we should read, ‘Sylvester’s minimal temporal part at \( t \)’; the relevant temporal parts are probably the instantaneous slices, that exist at exactly one time. (There’s a worry about whether or not these slices are really the right kinds of things to have the relevant properties, but we won’t worry about this for now.)
the problem by finding a sort of entity that has temporary properties, but lacks temporal parts. In fact, the same kind of entity will do the trick.

At $t_1$, Sylvester is seated (and therefore bent) and hungry (he’s hatching a devious plan to catch and eat Tweety). At $t_2$, after his plan has come to fruition, Sylvester is stretching (and therefore straight) and full. Then the property, *being bent*, is coinstantiated with *being hungry* at $t_1$ (because Sylvester is bent and hungry), but not at $t_2$. (Suppose that nobody else is both bent and hungry at $t_2$, either.) That is: *being bent* has, at $t_1$, the property, *being coinstantiated with being hungry*, but lacks it at $t_2$. Suppose that the property, *being coinstantiated with being hungry*, is a function from worlds to extensions—call it ‘$G’$. Since, at $t_1$, *being bent* is coinstantiated with *being hungry*, *being bent* must be a member of $G(\oplus)$. But since, at $t_2$, it’s not coinstantiated with *being hungry* (since at $t_2$, nobody is both bent and hungry) it must not be a member of $G(\oplus)$. Contradiction.

Appealing to temporal parts won’t help us in this case, since properties don’t have temporal parts. So we really do have a contradiction.

Just as the problem here mirrors our original problem, the solution here mirrors our original solution. There is no problem if we take properties to be functions from $<$world, time$>$ pairs to extensions.

This allows the extensions assigned to properties to vary with times, not just with worlds. The property, *being coinstantiated with being hungry*, can return an extension that includes *being bent* taking $<\oplus, t_1>$ as argument, but return an extension that doesn’t

---

25 Is this so clear? Well, it’s hard to see what the temporal parts of properties would be. The natural candidates to be the parts of properties, construed as sets of one sort or another (such as sets of instances, or sets of $<$world, extension$>$ pairs), are their subsets. But properties won’t, in general, have subsets that are well suited to play the role of temporal parts. (Especially if we’ve been convinced by the first part of
include being bent when it takes <@, t_2> as argument. Problem solved. Properties are functions from <world, time> pairs to extensions.

It would be nice if this argument was conclusive. Unfortunately it's not. One of the unsuccessful objections to the argument in the first section has a parallel here that isn't subject to the same criticism. Time-relativising the second-order properties, it turns out, fares better than world-relativising.

One of the responses discussed in section 2 was world-indexing second-order properties. We deny that there's any such property as being somebody's favorite property, and replace it with a bunch of world-indexed properties such as being somebody's favorite property in w_{453}. These properties are had necessarily if at all, so we're left without any troublesome accidental second-order properties.

This move failed because it assigned the wrong contents to sentences like 'being green is somebody's favorite property'. The parallel time-indexing move seems to fare better.

We can avoid the problem about the temporary properties of properties by denying that there are any. Time-index all of the temporary second-order properties: there's no such property as being coinstantiated with being hungry—instead there is a family of time-indexed properties, such as being coinstantiated with being hungry at t_i. In this case, though, there's no problem about assigning the wrong contents to sentences. The contents assigned to (particular utterances of) sentences like 'being bent is coinstantiated with being hungry' on this view are perfectly respectable—indeed,
intuitively exactly right. So we can resist the argument for taking properties to be functions from <world, time> pairs to extensions by time-relativising all of the (apparent) temporary second-order properties.

There are some worries about doing away with temporary properties in favor of their time-relativised cousins. One is that the resulting picture of the world looks suspiciously like one in which there’s not any change—there aren’t any properties that things have at some times, but lack at others. (This is also a worry about the temporal parts view—really, for any view that doesn’t, somehow or other, relativise property instantiation to times.)

There’s also another semantically driven objection to time-relativising properties. It would be nice if the semantic values of (at least most) predicates were constant across different contexts. This won’t be so if the semantic values of predicates are time-indexed properties. Predicates like ‘is red’ will need to express different time-indexed properties on different occasions of use.

These objections aren’t conclusive. I think that the cost/benefit assessment will, in the end, favor <world, time> functions, but the argument here certainly isn’t as open-and-shut as the case for the first revision. Suppose that you agree with me, and we accept that properties are functions from <world, time> pairs to extensions. What happens to the rest of Lewis’s system?

26 A first-pass semantics: if Daffy says, ‘being bent is coinstantiated with being hungry’ at t, then that occurrence of ‘is coinstantiated with being hungry’ expresses the time-indexed second-order property, being coinstantiated with being hungry at t, and so the sentence expresses the proposition that’s true in all and only the possible worlds in which being bent is coinstantiated with being hungry at t—that is, all and only the worlds in which, at t, something is both bent and hungry. This sounds like an extremely good candidate to be the content of Daffy’s utterance.

27 Of course there are temporal analogues of the modal notions of contingency and necessity, but it’s much less controversial (perhaps not controversial at all) to say that the propositions expressed by the relevant sentences are all timelessly true than it is to say that they’re all necessarily true.
7. Consequences of the Second Revision

The Puzzle of Change and the Argument from Temporary Intrinsics

Lewis argues that the best way to account for the phenomenon of (intrinsic) change—the fact that, for example, Sylvester can be bent at \( t_1 \) and not bent at \( t_2 \)—is to adopt a metaphysics of temporal parts. Sylvester is bent at \( t_1 \) because his temporal part at \( t_1 \) is bent, and he’s not bent at \( t_2 \) because his temporal part at \( t_2 \) is not bent. And while *Sylvester* can’t both have and not have the same property, there’s no problem about some of his parts having properties that other parts don’t. This, together with arguments against competing solutions, is the argument from temporary intrinsics.

Three things fall out of the discussion in the last section:

First, adopting a metaphysic of temporal parts is not a general solution to the problem of change. It’s a solution that works for particulars, but not for things like properties, which don’t have temporal parts.

Second, if we avoid the problem of change as it arises for second-order properties by taking properties to be functions from \(<\text{world, time}>\) pairs to extensions, there’s no need to adopt a metaphysics of temporal parts for particulars. Properties such as *being bent* (just like properties such as *being coinstantiated with being hungry*) can return different extensions for the same world at different times. So we don’t need to appeal to temporal parts to explain how Sylvester can be bent at some times and not at others, and the argument from temporary intrinsics loses its force.

Finally, across-the-board time-indexing also defangs the argument from temporary intrinsics. If we time-index shape properties like *being bent*, then there’s no

---

28 This sort of worry appears in Haslanger (1989a).
problem about Sylvester (the cat himself, not just his parts) having some, but not all, of the various \textit{being bent at }t\textit{ properties.}

In order for the argument from temporary intrinsics to be persuasive, the Lewisian needs a non-temporal-parts solution to the problem of temporary second-order properties that doesn’t generalize. If there’s no story about temporary second-order properties, then we haven’t solved the problem of change. If the story about second-order properties generalizes, then we’ll already have solved the problem of temporary first-order properties without any appeal to temporal parts.

I think that the role of \textit{intrinsicness} in the argument from temporary intrinsics is to rule out as general solutions certain strategies (like time-relativising) that seem to work for extrinsic properties. So it must be that, while we can, for example, time-index all of the extrinsic properties, time-indexing intrinsic properties like \textit{being bent} is unacceptable.

One reason why we might think this (not Lewis’s reason—see below) is that time-indexed properties are all \textit{relational}—they’re properties that we get from a relation (between objects and times) by fixing one of the relata (the time). If relational properties couldn’t be intrinsic, then we couldn’t say that shape properties were time-indexed, since shape properties are clearly intrinsic. This would give us a way to accept the time-indexing solution for temporary second-order properties (which are all extrinsic), but to deny that it generalizes—we can’t apply it across the board, because we’d be left without any intrinsic properties.

This argument is not persuasive. The claim that no relational property can be intrinsic has whatever force it has only because it’s so clear that relational properties where the other relatum is something like a person, a location, or a fire hydrant can’t be
intrinsic. It's a hasty overgeneralization from these kinds of examples that leads us to think that no relational property could be intrinsic. Lewis himself gives a counterexample: if Platonism is true, then participating in the form of Squareness is, though relational, still perfectly intrinsic.²⁹

Lewis’s official reason for rejecting time-indexing (and for rejecting views according to which properties are relations to times) is that it does away with genuine monadic intrinsic properties like being bent. None of the time-indexed properties can be identified with being bent, and the bent-at relation that holds between objects and times is obviously not monadic. So it looks as if being bent has been left out of the picture. And this is very bad.³⁰

Notice that this objection hinges on monadicity, not on intrinsicness. I don’t see any principled reason for thinking that it’s more objectionable to leave the intrinsic being bent out of the picture than it is to leave out the extrinsic, but equally monadic (and equally necessary for systematic semantics) being hunted by Elmer, or being somebody’s favorite property. If all monadic properties are on the same footing, then time-indexing won’t work as the conservative Lewisian’s non-generalizable solution to the second-order problem. Either Lewis’s objection works, in which case time-indexing isn’t a solution to the second-order problem, or it doesn’t, in which case the time-indexing move does generalize.

Noticing that there’s a problem about temporary second-order properties puts the argument from temporary intrinsics in a bad spot. For the argument to have force, we must have a principled reason for giving a disunified theory of change, on which we give

---

²⁹ Lewis (2002: 3).
³⁰ ibid: 4.
one account of the temporary properties of things with temporal parts and another for the temporary properties of things (like properties) that lack them. (Or maybe one account for intrinsic properties and another for extrinsic ones.) There doesn't seem to be any such reason. So we ought not to be moved by the argument from temporary intrinsics.

This isn't really a consequence of taking properties to be functions from \(<world, time>\) pairs to extensions. It's a consequence of the fact that we need to give an account of change in things that clearly don't have temporal parts, and there's no good reason to think that the account we give there won't generalize, and obviate the need to posit temporal parts in order to solve the problem of change in things like people, tables, and cats that plausibly do have temporal parts.

Still, the argument from temporary intrinsics isn't the only argument for temporal parts. Once we've adopted the second revision to the theory of properties, we don't need to appeal to temporal parts in order to solve the puzzle of change. But it's not as if the revision forces us to become endurantists. While the new theory of change is clearly compatible with endurantism, it's compatible with perdurance theories as well. What happens when we accept the revised theory of properties is that the problem of temporary intrinsics ceases to be an argument one way or the other.

Proper Subjects of Predication

We have a strong intuition that Sylvester's shape properties are intrinsic. One of the benefits of a four-dimensionalist ontology of temporal parts was supposed to be that it

31 See for example Lewis (1983c), Cartwright (1975), Balashov (1999), and especially Sider (2001).
allowed us to respect this intuition in a way that competing theories did not.\textsuperscript{32} There is an equally strong intuition that Sylvester’s shape properties are intrinsic properties of \textit{Sylvester}. The temporal parts view does not allow us to respect \textit{this} intuition. It’s Sylvester-stages, not \textit{Sylvester}, the temporarily extended continuant cat, that have the property \textit{being bent}. The best that the temporal parts solution can do for Sylvester is to give him the time-indexed property of \textit{being bent at t} in virtue of the shape of his t-stage. Considered as a whole, Sylvester doesn’t have any particular shape (except maybe something describable in 4-d geometrical terms).

This fact has been used to object to temporal parts theories.\textsuperscript{33} One benefit of the proposed revision to the theory of properties is that we can avoid this problem. It’s Sylvester, and not merely his parts, that has shape properties like \textit{being bent}.

\textit{Semantics}

Solving the problem of change by attributing temporary properties to temporal parts makes for an ugly semantics. Subject-predicate sentences will attribute properties sometimes to continuants, sometimes to their stages. The semantics will be particularly ugly if we, for example, time-index all of the extrinsic properties (alternatively, all of the second-order properties) and not the intrinsic (alternatively, the pure first-order) properties.

One good consequence of the revision is that it allows us to give a substantially cleaner semantics for predication. Suppose we solve the problem of temporary (first-order) properties by saying that properties like \textit{being bent} are really properties of person-
stages and cat-stages, rather than properties of temporally extended persons and cats. We still need to give an account of the cases where the properties being attributed are clearly properties of persons, not properties of stages. Doing this will require us to give a more complicated and disunified semantics than we'd like.

Here's one way it could go: Names like 'Sylvester' (and probably noun phrases in general) are ambiguous between continuant objects and their stages. When I say, 'Sylvester is bent', 'Sylvester' refers not to the whole cat, Sylvester, but to a Sylvester-stage. When I say, 'Sylvester is a cat', 'Sylvester' refers to Sylvester, the temporally extended continuant cat. When I say, 'Sylvester is a bent cat', 'Sylvester is a cat who is bent', or 'Sylvester is both bent and a cat', something fancy is going on. (For example, maybe the predicates are also ambiguous.)

Whatever the story is, it's going to be messier than we'd like. It would be nice if people's (and cats') names reliably referred to the relevant people and cats. It would be nice to keep the systematic ambiguity of predicates to a minimum. And it would be nice if the attribution of temporary properties worked the same way across the board—not by sometimes attributing time-indexed properties to whole objects, and sometimes attributing unindexed properties to stages. If we take properties to be functions from \(<world, time>\) pairs to extensions, we don't need to make either noun phrases or predicates ambiguous in order to deal with the problem of change, and we can have a uniform story—we never need to do any time-indexing, and we can always attribute the properties to the whole object.

34 How problematic this is, and how unique this problem is to temporal parts, is an interesting question. Chomsky (2000) discusses some similar examples that are independent of what we think about four-dimensionalism. We say things like, 'the tattered book on the table has been on the best-sellers list for
One seemingly less attractive consequence is that properties cease to be the kind of thing that, given an object, determines a possible-worlds proposition. On this account, properties will combine with objects to determine functions from <world, time> pairs to truth-values, not functions from worlds to truth values. So when we get the semantic values of subject-predicate sentences by composing the semantic values of the subject expression and the predicate expression, we won’t get possible-worlds propositions. We’ll get tensed propositions—functions from <world, time> pairs to extensions—things whose truth values can vary at different times within the same world.

It’s not clear what we should think of this consequence. If we were convinced that the semantic values of sentences had to be possible-worlds propositions, then it would be unwelcome. But it shouldn’t be that unwelcome, for two reasons. First, it’s not so clear that the semantic values of sentences have to be possible-worlds propositions rather than tensed propositions. Second, even if the contents of most sentences turn out to vary across times, we’ll still have some eternal sentences whose truth-values aren’t affected by the time.35

Attitudes De Se36

Lewis (1979) argues that attitudes like belief and desire are best thought of, not as attitudes toward propositions, but toward properties. The reason for this is that

twelve weeks now”. This requires some sort of special treatment, since only book-tokens are tattered, and only book-types make the best-seller list. (Though Chomsky puts these examples to a different use.)

35 It’s also not so clear that this view of the semantic values of predicates really commits us to any particular view about the semantic values of sentences. If we still want sentences to be associated with possible-worlds propositions, we can adjust the rest of our semantic (and maybe syntactic) theory in order to get some other constituent of the sentence to do the extra work.

36 Thanks to Bob Stalnaker for pointing out this consequence.
propositions—thought of as functions from worlds to truth values—won’t do as the contents of self-locating beliefs.\(^\text{37}\)

Suppose Pepe is lost in the library. This might be because he doesn’t know which world is actual—because, for example, he doesn’t know what the actual floorplan of the library is. But Pepe could still be lost, even if he knew everything there is to know about which world is actual. He could know the complete floorplan of the library, and even the location of every person, cat, and skunk in the library, and still not know where he is, because he could still fail to know which, of all of the creatures in the library, is him. (It helps, to make this sort of ignorance compatible with full knowledge of which world is actual, to suppose that it’s a very boring library full of amnesiacs, all having experiences indiscernible from Pepe’s.)\(^\text{38}\)

The same concerns arise for beliefs that, for example, my pants are on fire. I can believe that my pants are on fire without believing that Egan’s pants are on fire, and I can hope that someone turns a fire extinguisher on me right now without hoping that someone turns a fire extinguisher on Egan at 5:41pm. There also seems to be something that we both believe when I believe that my pants are on fire, and you believe that your pants are on fire. These facts are not easily accommodated on a view that takes attitudes like belief and desire to be attitudes toward possible-worlds propositions. Knowing which possible-worlds propositions are true can, at best, pick out a unique world as the one we inhabit. And like they say in the James Bond movie, sometimes (as when we want to know who or where we are) the world is not enough. We also need to know our place within it.

\(^{37}\) See also Perry (1979) for similar arguments (with a somewhat different conclusion), and Elga (2000, forthcoming), Perry (2001), and Stalnaker (1981, forthcoming) for more discussion and some interesting applications.

\(^{38}\) Well, a library that’s boring apart from the fact that it’s full of amnesiac skunks.
Second-Order Predication

We can treat self-locating attitudes as attitudes toward properties, rather than toward propositions. So (on Lewis’s theory of properties) when I believe that my pants are on fire, I believe that I am one of the possible individuals with burning pants. My attitude is toward a property—\textit{having burning pants}, and the content of the belief is that I am one of the individuals who instantiates the property.\footnote{As opposed to my attitude being toward a proposition—\textit{that Egan has burning pants}, and the content of the belief being that I am in one of the worlds in which the proposition is true.} We get essentially the same results if we accept the first revision and take properties to be functions from worlds to individuals.\footnote{It’s a bit trickier to state just what’s going on, though. The best way to see it is this: Functions $f$ from worlds to extensions are in one-one correspondence with sets of $<w, i>$ pairs (where $w$ is a world and $i$ an individual) such that $i$ is a member of $f(w)$. So when we take self-locating beliefs, desires, etc. to be attitudes toward properties, we can think of them as attitudes toward, not sets of worlds, but sets of $<\text{world, individual}>$ pairs. If individuals are worldbound, then the worlds won’t be doing any work—sets of individuals would do just as well. But this way of treating the contents of self-locating beliefs is compatible with (though it doesn’t mandate) taking individuals to be present in more than one world. Notice that this mirrors proposals that treat the contents of self-locating beliefs as sets of \textit{centered worlds}. (See e.g. Chalmers (1996, 2002).) Some people also include a specification of a \textit{time} in their centered worlds—we’ll get to this in a moment.}

There’s something very appealing about the treatment of self-locating beliefs as self-ascription of properties. But this strategy runs into trouble with \textit{temporally self-locating beliefs}.

Lewis can’t say that, when we have beliefs about what time it is, we’re self-ascribing some property. All but the most unfortunate people are temporally extended. So our attitudes toward properties (considered as sets of individuals) will single out sets of temporally extended things—four-dimensional spacetime worms. But singling out a worm doesn’t tell us what time it is. So (though they don’t say this in the Bond movie) sometimes—when, for example, we want to know what time it is—the worm is not enough.
If we say that properties are either the sets of their instances or functions from worlds to extensions, we can't treat my beliefs about what time it is as self-ascriptions of properties. No property of Egan—the continuant person—is going to make the relevant distinctions between different times. The most natural thing to say, probably, is that it's not people that have temporally self-locating beliefs, but their temporal parts. It's not me, but my present temporal part, that thinks that it's 5:30 (by thinking that it's one of the 5:30 person-stages). But this sounds strange. It's certainly surprising to learn that people never have beliefs about the time.\footnote{In fact, people probably don't have many beliefs at all, if we adopt a temporal parts based solution to the problem of change. Properties like believing that $P$ will almost always be temporary, and so they won't be properties of continuant people, but of their stages. There is, of course, footwork to be done here—there are ways of making it out that, even though people never instantiate belief-properties, people still have beliefs by having parts that instantiate belief-properties. We can do this, but it would be nice if we didn't have to.}

It would be nice not to have to say this, and we don't have to if we accept the second revision. If properties are functions from $<\text{world}, \text{time}>$ pairs to extensions, then people really can have temporally self-locating beliefs by self-ascribing properties.\footnote{Again, it's a bit tricky to say carefully what's going on. Just as we can treat functions from worlds to extensions as sets of ordered pairs of worlds and individuals, we can treat a function $f$ from $<\text{world}, \text{time}>$ pairs to extensions as the set of $<w, t, i>$ triples (where $w$ is a world, $t$ is a time, and $i$ is an object) such that $i$ is a member of $f(<w, t>)$. Then we can think of your self-locating beliefs (including your temporally self-locating beliefs) as attitudes toward sets of such triples, such that the belief is accurate iff $<\text{now}, \text{you}>$ is in the set. This mirrors accounts of centered-worlds contents in which centered worlds include a specification of the time. (Thanks to Juan Comesana for discussion here.)} So we get a nicer story about temporally self-locating beliefs if we accept the second revision: beliefs about what time it is can be treated as self-attributions of properties by people, not just by their stages.
**Conclusion**

Problems about accidental second-order predication motivate us—force us, I think—not to identify properties with the sets of their instances. If we identify them instead with functions from worlds to extensions, we get a theory of properties that is neutral with respect to disputes over counterpart theory, and we avoid a problem for Lewis’s theory of events.

Similar problems about temporary second-order predication motivate us—though this time I don’t think that they force us—to give up this theory as well, and to identify properties with functions from \(<\text{world, time}>\) pairs to extensions. Again, the replacement theory is neutral with respect to a metaphysical dispute that the old theory (seems to) force us to take a stand on—the dispute over whether objects have temporal parts. It also allows us to give a smoother semantics for predication, and to better accommodate our intuitions about which objects temporary properties are properties of, and to make temporally self-locating beliefs genuinely self-locating.
Secondary Qualities, Self-Locating Belief, and Self-Locating Assertion

Introduction

It’s common for philosophers to want to draw a distinction between those qualities that are fully objective, genuine qualities of things as they are in themselves, and those that are somehow subjective, less than fully real, qualities of things as they are for us. This distinction is elusive; while it’s quite compelling intuitively, it’s also extremely difficult to characterize. My goal in this paper will be to offer such a characterization. The distinction we’d like to draw is one that’s at least in the neighborhood of the traditional distinction between primary and secondary qualities—primary qualities are typically taken to be of the former kind, secondary qualities of the latter.

1. The Job Description

I’m going to call the elusive distinction between the fully real, perfectly objective qualities of things and the less-than-fully-real, subjective qualities the primary/secondary quality distinction.

I definitely don’t want to claim that the distinction I’m going to draw captures everything everybody’s ever wanted to say about any distinction that has ever gone by that name. I doubt that any distinction could do that—people have said an awful lot about the distinction (or more likely, some distinctions) between primary and secondary qualities, and it’s probably not all consistent. All that I’m hoping to do here is provide a distinction that captures a cluster of ideas (about objectivity, less-than-full reality, etc.)
that seem to be (a) central to at least some understandings of the traditional distinction, and (b) hard to cash out in a satisfying way.

There’s a lot of rhetoric that surrounds (at least very many versions of) the traditional distinction between primary and secondary qualities. The ways in which the distinction has been drawn don’t seem to justify the rhetoric. What I want to do here is provide a distinction (the only one, I think) that justifies the rhetoric.

I take the task of justifying the rhetoric to be more important than making sure that the distinction classifies particular qualities as primary and secondary along the lines that philosophers have traditionally wanted to divide them. In particular, I take it to be of only secondary importance that the traditional paradigm cases of secondary qualities—sensory qualities like colors, tastes, smells, etc.—turn out to be secondary qualities, when the distinction between primary and secondary qualities is cashed out in the way that I suggest. I think that, in fact, my distinction does carve quite close to the traditional one, but what I’m primarily concerned to do is to provide a distinction that justifies, for whichever qualities turn out to be properly classified as secondary, the rhetoric that’s been employed in discussing the qualities that have traditionally been classified as secondary. I don’t care so much about justifying the traditional classifications.

Following are some examples of the sorts of rhetoric that I’m concerned to justify. Some of the passages are about secondary qualities in general, or the distinction between primary and secondary qualities, while others make claims about particular qualities which are taken to be secondary. Notice that claims of the second kind can be separated into two parts: first, the claim that secondary qualities are thus-and-so, and second, the claim that some quality Q is a secondary quality. What I want to focus on from these
passages is just the first part—the characterization of what’s supposed to be distinctive of the secondary qualities, and how they’re supposed to be different from the primary qualities, rather than the claims about which qualities are primary and which are secondary.¹

St. Paul:

There is nothing unclean of itself: but to him that esteemeth any thing to be unclean, to him it is unclean...²

Galileo:

Hence I think that these tastes, odors, colors, etc….hold their residence solely in the sensitive body; so that if the animal were removed, every such quality would be abolished and annihilated.³

Bernard Williams:

In understanding, even sketchily, at a general and reflective level, why things appear variously coloured to various observers, we shall find that we have left behind any idea that, in some way which transcends those facts, they ‘really’ have one colour rather than another. In thinking of these explanations, we are in fact using a conception in which colour does not figure at all as a quality of the things.⁴

We can say, and indeed say truly, that grass before there was consciousness was green… But it is, nonetheless, relative, relating to human tastes and interests.⁵

¹ I suspect that, for example, colors probably aren’t secondary qualities, but the claims that are made about the colors below are still useful for characterizing the primary/secondary quality distinction.
² Romans 14:14 (thanks to Tyler Doggett for the reference.)
³ Galileo (1842: 333).
⁴ Williams (1978: 242).
⁵ Ibid: 243. Williams is actually talking about amusingness, not green, in the second part of the quote, but nothing hangs on this.
Secondary Qualities

Thomas Nagel:

The third step is to try to form a conception of that true nature [of the physical world] independent of its appearance either to us or to other types of observers. This means not only not thinking of the physical world from our own particular point of view, but not thinking of it from a more general human perceptual point of view either: not thinking of how it looks, feels, smells, tastes, or sounds. *These secondary qualities then drop out of our picture of the external world…*  

Colin McGinn:

Secondary qualities resemble properties like being poisonous or nourishing in this respect: plainly, these properties are relative to some implicit or explicit choice of creature as that with respect to which a substance is declared poisonous or nourishing. This relativity implies that there is no genuine disagreement between us and the Martians when they call an object green which we call red…  

I think it is an *a priori* truth that only the primary qualities correspond to how things are in themselves…  

What the scientifically informed view denies is [not that objects are coloured, but] just that objects are *objectively or intrinsically* coloured, i.e. that objects have colour in the way that they have shape; it denies that possession of colour is an observer-independent condition.  

There are some themes that emerge from the series of quotations above. Secondary qualities are supposed to be *observer-dependent*: the secondary qualities depend (in some non-trivial way) on the existence or the peculiarities of observers in a way that primary qualities don’t. They’re supposed to exhibit some sort of *relativity*:

---

8 *Ibid* p114.
Secondary qualities are unlike primary qualities in that an object can have a secondary quality relative to one observer that it lacks relative to another. Finally, they’re supposed to be somehow metaphysically second-class. They display a sort of unreality—there’s something less than fully real about the secondary qualities. It’s hard to say just what that’s supposed to amount to before giving more of a story about the distinction.

The primary/secondary quality distinction that I’m after is whichever one justifies (or does the best job of justifying) the sorts of rhetoric displayed above. Again, this may or may not map nicely onto any of the many distinctions of the same name that are to be found in the philosophical literature. I’m not terribly concerned about whether it does or not. What I care about is that it capture the thought which I think underlies (at least a great deal of) the interest in drawing any of the primary/secondary distinctions that have in fact been drawn: that there is an important difference between the fully real, observer-independent qualities of things, and the metaphysically second-class, observer-dependent qualities of things.

2. Trouble

One distinction that would at least get us the more real/less real part of the primary/secondary rhetoric is the distinction between properties that things actually have and properties that they don’t. But this obviously would be an extremely unsatisfying way to draw the distinction. For one thing, it would probably involve us in attributing an awful lot of systematic error. For another, it doesn’t really justify the rhetoric. Any distinction between primary and secondary qualities that justifies the traditional rhetoric will have to run a lot deeper. To justify, for example, the talk about relativity and

---

10 McGinn is particularly explicit about this.
observer-dependence, it will have to turn out that, among the qualities that we *correctly* attribute to things, some of them are more real, more objective, more a part of the absolute conception of the world, than others.¹¹

But it might be that some sort of projectivist error theory is the best that we can do. Gideon Rosen (1994) discusses the urge to draw a metaphysically substantive distinction between two kinds of *facts*, such that facts of one kind characterize the world as it is in itself, are objective, fully real, etc., while facts of the other kind characterize the world as it is for us, are subjective, less than fully real, etc. He makes a persuasive case that this can’t be done.¹²

If Rosen is right that there’s no interesting metaphysical distinction in point of objectivity and subjectivity between *facts*, then it’s hard to see how there could be one at the level of *qualities*. If we had qualities that were, in some metaphysically interesting sense, subjective, less than fully real, etc., then we could get metaphysically second-class facts by correctly attributing those qualities to things. So if there aren’t any interestingly subjective facts, there must not be any interestingly subjective qualities, either.

Rosen argues pretty convincingly that the usual ways in which people have tried to draw the distinction won’t work. I’ll briefly discuss just two examples here:

That some fact, quality, or entity is *mind-dependent* doesn’t seem to impugn its full, first-class reality. Certainly a fact’s being *causally* mind-dependent—in that its

¹¹ Not everyone says this. One could be a projectivist, and think that the secondary qualities are qualities of our experiences that we mistakenly attribute to things outside the mind. Then the more real/less real distinction would be the distinction between properties that are correctly/mistakenly attributed to the things we attribute them to. (See for example Boghossian and Velleman (1989).) This requires accepting an error theory, though, instead of just a theory about the nature of the qualities that we correctly attribute to things. And a lot of the primary/secondary rhetoric seems to suggest a difference in the kind of quality being correctly attributed (to things outside the mind), rather than a difference in the accuracy of the attribution.

obtaining was brought about by some mental activity—doesn’t make the fact
metaphysically second-class. Facts about the existence of artifacts, for example, are as
metaphysically respectable as facts get, and they’re causally mind-dependent—the
existence of my kitchen table was brought about, at least in part, by the thoughts, plans,
and intentions of some carpenter.

Response-dependence also seems not to do the trick. That a thing is disposed to
cause response R in a subject S in circumstances C is a perfectly objective fact about the
thing. Consider Locke’s tertiary qualities, of being, for example, disposed to melt wax in
ordinary circumstances, or Rosen’s example of being disposed to annoy fox terriers under
ordinary circumstances. These are perfectly objective features of whatever has them.
The fact that something is disposed to melt wax, or to annoy fox terriers, isn’t in any way
metaphysically second-class, subjective, or less than fully real. And if these aren’t
metaphysically second-class, then neither is, for example, being disposed to cause
sensations of kind K in humans.

Our almost obsessive interest in these latter kinds of facts—facts about which
things are disposed to cause which responses in humans—is subjective, parochial, and so
forth, but the facts themselves are perfectly objective. Facts about what’s disposed to
cause certain kinds of sensations in us are just as much a part of the world as it is in itself
as the facts about which things are disposed to melt wax or annoy fox terriers. So
response-dependence doesn’t seem to get us any metaphysically interesting distinction
between the objective, genuine facts and those that are subjective and therefore somehow
second-class.
The moral of Rosen’s story seems to be: if an object has a property, then it’s a perfectly objective matter of fact that it has that property. We can find properties such that our reasons for being interested in which things have them are subjective and parochial, but that doesn’t make them metaphysically second-class—it doesn’t make the fact or the property subjective in any metaphysically interesting sense. 13

Still, I think there are some features that we attribute to things that are subjective in a way that makes them metaphysically second-class.

3. Centering Features

There’s an attractive picture of mental and linguistic content on 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. 14

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.

Typically (always?) the relevant sets of worlds are picked out by attributing properties to things—I believe that Fido is furry, desire that my dissertation is complete,

13 Robert Stalnaker has suggested the distinction between natural and unnatural properties as another place to locate the primary/secondary quality distinction. This is a metaphysically interesting distinction between kinds of properties that we correctly attribute to things. But it’s probably not metaphysically interesting in the right way—it doesn’t license the sorts of relativity and observer-dependence rhetoric that surrounds the secondary qualities.

and assert that France is hexagonal. The content of my belief, desire, or assertion is the set of worlds in which the relevant things have the relevant properties.

There’s a distinctive role that the property plays in determining the possible-worlds content of my belief, desire, or assertion: given an object, the property determines a set of worlds—the worlds in which the object has the property. So given Fido, the property, \textit{being furry} determines a set of worlds: the worlds in which Fido is furry.

Equivalently, given a world \(w\), the property determines an extension—the set of things that have the property in \(w\). Given a world \(w\), \textit{being furry} determines an extension—all of the things that are furry in \(w\). Fido is in the extension of \textit{being furry} in \(w\) iff \(w\) is a member of the proposition expressed by “Fido is furry”.

So we can think of properties as functions from worlds to extensions (or, equivalently, as functions from objects to possible-worlds propositions).\(^{15}\) This illustrates the distinctive role that properties play in characterizing, and maybe in determining, the contents of representations—we can characterize the possible-worlds content of a representation by saying which properties it attributes to which things. Maybe it’s also true that the way the content gets determined is by bits of the thing doing the representing that are associated with the properties getting put together in the right way with bits that are associated with the objects. (That \textit{is} how it works with natural language—at least, for the simple, subject/predicate bits of it.\(^{16}\) Fodor & co. think that’s how it works with

\(^{15}\) Note that we need to have merely possible objects available to be elements of the extensions in the first case, and values of the functions in the second, for this to work. Also, this isn’t quite my official story—I think that properties are functions from \(<\text{world}, \text{time}>\) pairs to extensions, but nothing hangs on the difference here. (See “Second Order Predication and the Metaphysics of Properties”.)

\(^{16}\) Things are a bit more complicated for, e.g., universal and existential claims, where the properties aren’t being attributed to any particular object, but not, I think, in a way that makes any trouble.
mental representation, too, but that’s more controversial. So I’m going to be adamant about the description/characterization claim, tentative about the determination claim.)

Take “Kermit is a frog”. This sentence picks out a set of worlds—the one in which Kermit’s got the property, being a frog. A belief or assertion whose content is that set of worlds is one that represents Kermit as having a certain feature—namely, being a frog. Not-so-bold claim: this is a really central, indispensable way of characterizing the possible-worlds contents of mental states, assertions, etc.

There’s good reason to think that some contentful mental states—beliefs and desires, for example—don’t have (merely) possible-worlds content, but have centered-worlds content instead (or as well). (This is motivated by arguments and examples from Perry (1979) and Lewis (1979).)

A possible world is a way things might have been. For present purposes, we can either follow Lewis and take possible worlds to be big concrete universes—more things of the same kind as us and all of our surroundings—or we can follow Stalnaker, and take them to be maximally specific properties that the whole universe might have instantiated. A centered world is to a possible world what a map with a “you are here” arrow added is to an arrowless map. Centered worlds single out not just a way for the world to be, but also a location within the world. They’re probably best thought of as ordered pairs of a world and a center. There are different ways of picking out a center—the center could be, for example, a spacetime point or an individual within the world.

---

17 I should point out that I’m working with a pretty undemanding notion of what it takes to correctly attribute a property to something. I attribute a property F to an object a iff I believe some proposition p such that (i) p is the proposition that you get by applying F to a, and (ii) p is true. (In order to just attribute F to a, all I have to do is satisfy (i).) This is a pretty minimal standard for correct attribution—it doesn’t make any claims about the structure of the relevant mental representation, for example.

Not much hangs on this decision, but it will be convenient for present purposes to take centers to be \(<\text{individual, time}\>\) pairs.\(^{19}\)

If we think of representation in terms of selection among possibilities, then we can think of representations with \textit{possible-worlds contents} as selecting among possibilities \textit{for the world}, and those with centered-worlds as selecting among possibilities \textit{for the agent}. 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. The latter are more fine-grained than the former, since (at least in most cases), each world is going to contain a number of positions that some agent might occupy.

We ought to say that some mental representations have centered-worlds content. When I have beliefs about my location, or the time, my belief is not well represented as an attitude toward a possible-worlds proposition. Possible-worlds propositions don’t cut finely enough—knowledge of, and belief about, possible-worlds propositions can pin down which world I’m in, but they can’t pin down my \textit{location} (either spatial or temporal) within that world.\(^{20,21}\)

(A brief digression: Perry draws an importantly different conclusion from the cases. He wants to locate the extra structure not in the contents of belief, but in the belief-states. So, when each of us believes that our own pants are on fire, there’s no

\(^{19}\)This is in order to remain as neutral as possible about issues having to do with persistence and the nature of properties, while keeping the individuals-as-centers picture which makes thinking about centered worlds more natural and intuitive than the spacetime-points-as-centers picture.

\(^{20}\)See Lewis (1979) and Perry (1979) for the classic examples. Stalnaker (1981) also discusses a number of nice examples, though he offers an alternative account of the phenomenon there.

\(^{21}\)A complication: anything that determines a possible-worlds proposition determines a centered-worlds proposition. It just determines (to introduce a technical term) a \textit{boring} centered-worlds proposition, that includes, for each world \(w\), either all of the positions in \(w\) or none of them. So the claim is really that many contentful mental states have contents that determine \textit{interesting} (i.e., non-boring) centered worlds propositions.
common *content* to our beliefs, but we’re in the same kind of belief-state. If this kind of account is right, then we don’t need centered-worlds contents for propositional attitudes. I’m going to assume without argument that this kind of account *isn’t* right, and that propositional attitudes do, at least sometimes, have centered-worlds contents. (And that since centered-worlds contents are the contents of propositional attitudes, that it’s safe to call them propositions.) It’s worth noticing that if I’m wrong about the need for centered-worlds content, pretty much everything I say from now on will be false.)

We describe the *possible-worlds content* of a representation (largely) in terms of the *properties* that things are represented as having. By representing Kermit as having the property, *being green*, we pick out a class of worlds—all and only the worlds in which Kermit is green. When I believe that the world is a certain way, I represent some things as being green and others as being red, some things as being furry and others as being scaly, and so on. In this way I narrow down the range of worlds that I take to be candidates for actuality. (We can tell the same kind of story of other kinds of representational states and entities—for example desires, fears, and natural language sentences.)

When I have beliefs not just about what the world is like, but about my location within it—when I have *self-locating* beliefs—something very similar is going on. I represent some things as being *nearby* and others as *far away*, some events as *present* and others as *past or future*, and some objects as being *on my foot* and others as *in my ear*. In this way I narrow down the range of possible *predicaments*—possible locations within worlds—that I take to be candidates to be the one that I am in. (Again, the same goes for other representations with this kind of content.)
The ways in which we describe centered-worlds contents are very similar to the ways in which we describe possible-worlds contents. In the possible-worlds case, one very common way to single out a set of worlds is by attributing some property like *being green* to some object like Kermit. In the centered-worlds case, a very common way to single out a set of centered worlds is by attributing some “property” like *being nearby* to some object like Kermit. Just as *being green* is a function that, when we plug Kermit into it, delivers a set of worlds (the ones in which Kermit’s green), *being nearby* is a function that, when we plug Kermit into it, delivers a set of centered worlds—the ones where Kermit’s near the center. At least, this looks like a tempting thing to say. But maybe we shouldn’t be so quick to say that there are any such “properties” as *being nearby* or *being on my foot* for us to attribute to things.

It’s certainly true that if there are such “properties”, they deserve the scare quotes—they’re not *properties*. Properties are (or at least determine) functions from worlds to extensions. A “property” like *being nearby* won’t do that. Which things are nearby—which things are in the extension of the “property”, *being nearby*—depends not just on which world is actual, but also on *where you are* within the world. 22 You and I are worldmates, but lots of things are near me and far away from you. So *being nearby* doesn’t determine a function from worlds to extensions, and so it’s not a property. 23

---

22 It also depends on the currently active standards of nearness, but let’s ignore that for now and pretend that there’s only one standard of nearness that applies in all contexts.

23 Again, my official view is that properties are functions from <world, time> pairs to extensions, but that doesn’t matter for present purposes; *being nearby* doesn’t determine one of those, either. Why am I so sure that they’re not properties? Well, this is partly stipulative. Nothing bad happens if I allow that they’re properties, but insist that there’s an important distinction between two kinds of properties. One reason for insisting on the name, though, is that properties ought to be the sort of thing about which a nominalist/universalist/trope theorist debate wouldn’t just be crazy. There’s no plausibility at all to the idea that centering features are some screwy kind of universal, or sets of screwy kinds of tropes.
This shouldn’t be surprising. If it was a property, it wouldn’t be fit to play the role that it does in determining centered-world contents, because we wouldn’t be able to use it to distinguish between different positions within a given world. What it is is the analogue of a property for centered-worlds contents. That is, it’s a function (or something that determines a function) from centered worlds to extensions, in the same way that a property is (or determines) a function from worlds to extensions. So given a world and a center, is nearby will give us the set of things that are nearby if we’re at that location in that world. And given an object, being nearby will give us the set of centered worlds in which that object is nearby (that is, near the center). We can pick out a set of centered worlds by saying that they’re the ones in which Kermit is nearby—the ones that, when plugged into the being nearby function, deliver an extension that includes Kermit.

We need a name for these things. I’ll call them centering features. ‘Features’ to indicate that they’re property-analogues, not properties, and ‘centering’ because their descriptive role is to select not just a world, but a center (attributing a property to Kermit selects between worlds that might be the one that I inhabit, while attributing a centering feature to Kermit selects between predicaments that might be the one that I am in). 24,25

But whatever we call them, two things should be pretty clear: First, they’re useful for describing centered-worlds contents. In the same way that it’s useful to talk about the

---

24 Thanks to Alex Byrne for the name. 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. One difficulty with the primary/secondary quality distinction I’m trying to draw here is that it’s not obvious that there are any predicates that express centering features—more on this in section 5.

25 In the same way that there are interesting and boring centered-worlds propositions (in the sense defined in note 20), there are interesting and boring centering features—the boring ones are the ones that, when combined with objects, always determine boring centered-worlds propositions, and the interesting ones are the ones that sometimes determine interesting centered-worlds propositions.
properties that people attribute to things in describing their beliefs about which world is actual (try specifying the content of my belief that Kermit is green without appealing to properties), it’s useful to talk about the centering features that people attribute to things in describing their self-locating beliefs (beliefs about which predicament they’re in).

Second, there’s nothing mysterious about them. Properties are one kind of set theoretical object—functions from worlds to extensions—and centering features are set-theoretical objects of another, analogous kind—functions from centered worlds to extensions.26

Some terminology:

A quality or feature is something that determines a function from objects to contents, where the contents might be either sets of possible worlds or sets of centered worlds.

A property is a kind of quality—the kind that determines a possible-worlds content.

A centering feature is a quality that determines a centered-worlds content.

A representation attributes a feature to an object iff the representation has as its content the proposition that’s determined by applying the quality to the object.27

26 Or at least they stand in some very intimate relation to such set-theoretic objects. The identity claims are more controversial. But even if the identity claims are false, centering features don’t seem to be any more mysterious than properties. (Two further, somewhat sketchy points: The set-theoretic identity claims seem, if anything, more plausible for centering features than for properties, since centering features don’t have, as part of their job description, the objective-similarity-grounding and relevance-to-causal-powers roles that properties are often taken to have in theirs. Second, if there’s a split in the property role, between, say, the properties that play a role in describing and determining content and the ones that explain the similarities between and causal powers of objects, then the set-theoretic identity claims look more plausible for the content describing and determining properties, and centering features will be of the same kind as these sorts of properties.)

27 Better, if its content is something that entails the proposition that you get by applying the quality to the object. (This is to get it to turn out that ‘Kermit is green and Big Bird is yellow’ attributes being green to Kermit.)
representations that attribute properties to things have possible-worlds content, and representations that attribute centering features to things have centered-worlds content.)

4. Secondary Qualities as Centering Features

Centering features are well suited to play the role of secondary qualities, and the property/centering feature distinction is well suited to play the role of the primary/secondary quality distinction. That is, it’s well suited to justify the kinds of relativity, observer-dependence, and less-than-full-reality rhetoric that’s distinctive of talk about secondary qualities.

I said earlier that centered worlds are to possible worlds what maps with “you are here” arrows are to maps without such arrows. A helpful extension of that metaphor is to think of possible-worlds contents as distinguishing between classes of maps, and centered-worlds contents as distinguishing between classes of maps-with-arrows. This picture of the difference between centered-world and possible-world contents makes it easier to see how the distinction between properties and centering features parallels the

---

28 Note that these needn’t be exclusive—there might be representations with both kinds of content.
29 Let’s distinguish two questions: (i) Can we make sense of there being any features of things represented in experience that behave in the way secondary qualities are supposed to? (ii) Is it plausible that the particular features that people have taken to be secondary qualities (e.g. colors) behave that way?

The claim I’m mostly going to be concerned with is that the answer to the first question is “yes, and they’re centering features”. I’m not so sure about the second, though I’ll have a bit to say about it toward the end of the paper. The main thing that I want to do is open up a space for a primary/secondary quality distinction, or something very like it—to show that there can be a metaphysically significant distinction in point of objectivity, genuine reality, etc. between two classes of qualities, both of which we correctly attribute to things outside the mind. I’m less worried (at least for now) about saying which side of the line particular qualities like colors, smells, etc. fall on.

The second question is interesting, but it’s only interesting once we’ve got a serviceable answer to the first question—a serviceable characterization of what the distinction is. Questions about which sides of some line different qualities fall on are only interesting (or answerable) once we have some idea of where the line is. So before we can sort out whether or not, for example, colors are secondary qualities, we need to know just what (or at least roughly what) being a secondary quality amounts to.
distinction between primary and secondary qualities. (How to restate the same points without the metaphor should be clear enough.)

Remember the themes we drew from the series of quotations in section 1. Secondary qualities are supposed to be observer-dependent, relative, and less-than-fully real in a way that distinguishes them from the primary qualities.

One way that the less-than-full-reality talk is sometimes put is that the secondary qualities go missing from the “absolute conception” of the world. They’re features of things “as they are for us” rather than “as they are in themselves”.

When we characterize the distinction between primary and secondary qualities as the distinction between properties and centering features, we can get a plausible understanding of what “absolute conception” means, such that secondary don’t appear in the absolute conception of the world. The absolute conception of the world is supposed to be the one that everybody has got to accept, regardless of what their perceptual apparatus, etc. is like. That looks like the conception that picks out which world is actual, and no more. Any conception of the world that has a content that’s more fine-grained than a possible-worlds proposition is going to be non-mandatory—it needn’t be shared by all of the maximally well-informed inhabitants of a given world. So the features that appear in the absolute conception will be the ones that determine possible-worlds contents—the properties. The features that determine centered-worlds contents—the centering features—won’t show up in the absolute conception. They’ll only show up in our particular, located, parochial conception.

Here are two alternative statements of the same idea: (i) If you fail to believe some true possible-worlds proposition, then you’ve failed to completely characterize the
world. Failure to believe all of the true centered-worlds contents, though, is compatible with having completely characterized the world, though not your place within it. (ii) You don’t need centering features in order to draw the map right—all you need for that is properties. You only need centering features in order to put the “you are here” arrow in the right spot. And while all of the maximally well-informed inhabitants of a world have got to agree on what the map looks like, they don’t have to agree on where the “you are here” arrow points.

This difference between properties and centering features—that all of our maximally well-informed worldmates need to agree on which properties things have, though they needn’t agree on which centering features they have—promises to justify a great deal of the rhetoric of less-than-full-reality (as well as the rhetoric of relativity) that surrounds the secondary qualities.

Another bit of the rhetoric of less-than-full reality is that secondary qualities are supposed to be, “not part of the world as it is in itself, but of the world as it is for us”. We can see how this is a natural way to describe centering features, too. Representations with possible-worlds contents describe the world as it is in itself—they tell us (if they’re accurate) what the world is like. Representations with centered-worlds content describe what the world is like for us—the tell us (if they’re accurate) about our own individual situation in the world; our own individual predicament. We can change how things are represented as being for us without changing how the world is represented as being in itself, because we can represent ourselves as being in a different predicament without representing ourselves as being in a different world. (This is one of the main points of the examples in Lewis (1979).)
We also have a nice account of the *relativity* of secondary qualities. Colin McGinn (1983) says that colors are secondary qualities, and wants it to be possible that (i) the Martians attribute *being green* to the things we attribute *being red* to, and vice versa, and (ii) we’re both right. If colors are centering features, we can get things being green for us and red for Martians, and vice versa. In general, we can get incompatible features F and G such that one observer represents some object as F, another represents it as G, and they’re both right.\(^{30}\) We can also get the sort of change over time that McGinn wants—if we were all taste-permuted in the right way, things that used to be sweet would start being bitter. We’d still attribute the same features to things when we called them ‘sweet’ or ‘bitter’, but different things would have the features, because of the changes in us (assuming *being sweet*, for example, is something like the centering feature, *being disposed to cause S sensations in me now*).

It’s far from clear that McGinn’s right about how colors or tastes behave, but that’s not really the point. The point is to make sense of the *possibility* of there being some features or other that act the way McGinn wants colors and tastes to act.\(^{31}\)

We can also justify the sort of observer-dependence rhetoric that we encountered in section 1. (The first pass through this will be fairly metaphorical. The metaphors are cashed out in footnote 30.)

Take an ordinary possible world, without a center. Nothing has any centering features there, because there’s no center for them to bear the relevant relations to. Add the “you are here” arrow, and things suddenly take on a number of new qualities—

\(^{30}\) See “Appearance Properties?”

\(^{31}\) McGinn points out a number of suggestive analogies between indexicals and secondary qualities. One of the things I want to do is explain the similarities—there are all these close analogies between secondary qualities and indexicals because secondary qualities are centering features, which are close cousins of indexical terms.
qualities that they get not (or not entirely) in virtue of what the world is like in itself, but in virtue of where the arrow points. It’s quite natural to think of these features as being added to the world by the selection of a center. The selection of a center—the adoption of some subjective position within the world—provides the world with all of these features which aren’t present in the world considered on its own, without any center. 32

In the same way, centering features (unlike properties) disappear when you take away the “you are here” arrow. Nothing’s nearby or in my ear until you specify a center. This looks like a satisfying sort of observer-dependence.

Note that we don’t, even on the assumption that colors are centering features, get the truth of, “if there had been no observers, nothing would’ve been colored”. Lots of merely possible things have dispositions to cause various responses in me, even the ones in worlds where I don’t exist (or don’t have any counterparts). 33 The fact that the centering features don’t appear until we select a center seems like enough, though, to justify (at least most of) the kind of observer-dependence and less-than-full-reality talk

---

32 Less metaphorically: models that only contain centerless worlds won’t have the resources to provide extensions for centering features. To assign an extension to a centering feature, you need to have a model that picks out a privileged center. It’s much like what happens with sentences involving “actually”—you need a model that picks out a particular world as actual before you can assign them truth-values. Taking the “absolute” view of logical space, where we don’t single out a particular world as actual, we don’t get truth-values for anything that’s got a significant occurrence of any rigidifying machinery. (See e.g. Davies and Humberstone (1980).) In the same way, taking the “absolute” view of the world, where we don’t single out a particular location as mine, we don’t get extensions for any centering features. An objective model of the world—one that doesn’t privilege a particular center—doesn’t have the resources to deliver extensions for centering features. Add the center, though, and you get extensions for the centering features.

33 Does attribution of dispositional centering features get us involved in the potentially unpleasant business of trying to evaluate counterfactuals with centered-worlds propositions as antecedents and/or consequents? No. To see why not, it’s easiest to think about centered-worlds beliefs along Lewisian lines, as self-ascriptions of properties. When I attribute a centering feature like, being disposed to cause G experiences in me in C to Kermit, I’m self ascribing the property, being one of the things in which Kermit is disposed to cause G experiences in circumstances C. That is, I’m taking myself to be one of the things of which some open sentence along the lines of, “if x was in C, Kermit would cause G in x” is true. The only counterfactual that needs to get evaluated here is one with possible-worlds propositions as both antecedent and consequent.
that many philosophers go in for when discussing the secondary qualities.\textsuperscript{34} The secondary qualities “fall away” when you take the objective, observer-independent view—stop thinking from the perspective of some observer, and all of the secondary qualities disappear. (See note 30 if you’re worried about all the metaphors here.)

An example by way of summary: We need centered-worlds contents for belief because we have beliefs not just about what the world is like, but also about our position within it. When Sarah the transatlantic sailor consults her charts, sextant, and GPS, she’s not (primarily) deliberating and gathering evidence about how the world is—her main concern is with her position within a world that she already knows most of the relevant facts about.

If we have beliefs with centered-worlds contents, then we attribute centering features to things. When Sarah believes that Gibraltar is nearby, she believes the centered-worlds proposition that you get by applying the centering feature, being nearby, to Gibraltar. If Sarah’s belief is accurate, then she’s correctly attributed a centering feature to Gibraltar.

The centering feature that she attributes to Gibraltar is metaphysically second class compared to properties of Gibraltar like being built around a big rock. The property, unlike the centering feature, figures in the “absolute conception of the world”—the conception that has to be shared by all of Sarah’s maximally well-informed worldmates. Since Sarah correctly attributes both being nearby and being built around a big rock to Gibraltar, there’s an interesting, metaphysically significant distinction in point

\textsuperscript{34} Though Galileo, at least, actually seems to have been advocating an error theory, on which we’re mistaken when we attribute colors to things. A nice feature of the view under consideration is that we can make sense of the thought that seems to be motivating the error theory without being forced to actually be error theorists.
of objectivity between two kinds of features that Sarah correctly attributes to things in her environment.

5. Centering Features in Language

Since it’s very plausible that some of our propositional attitudes have centered-worlds content, it’s very plausible that we attribute centering features to objects in thought. What’s less clear is that we attribute centering features to objects in language.

It would be a shame for my account of the distinction between primary and secondary qualities if we didn’t have centering features being attributed to things in language. If the primary/secondary quality distinction only shows up at the level of features attributed in thought, it will be hard to justify the sorts of rhetoric that I’m concerned to justify. All of that rhetoric happens in natural language, and it’s supposed to be about the features that are attributed to things by natural language predicates. So it looks like a distinction that only appears at the level of thought, and not in language, won’t be able to do the work that I want it to do.

Another reason why it would be nice for me to have some natural language predicates that express centering features is that the question of whether, for example, colors, sounds, or tastes are secondary qualities is supposed to be one that admits of some interesting debate. But if we only attribute centering features to things in thought, never in language, the debate’s going to be very short. Colors can’t be centering features, because the colors are the qualities that we attribute to things in sentences like “Kermit is green” and “fire engines are red”.

- 62 -
One reason to be worried about whether or not the primary/secondary quality
distinction, as I’ve characterized it, does show up in natural language is that its doing so
has some surprising semantic consequences. If there are some predicates that express not
properties, but centering features, then there will (unless something very fancy is going
on) be some sentences of natural language that express centered-worlds propositions
rather than possible-worlds propositions. This is at odds with what standard semantic
theories tell us, and so it’s going to require some motivation.

Another reason to be worried is that it seems, at least at a first glance, to conflict
with the attractive picture of assertion presented in Stalnaker (1978). On Stalnaker’s
picture, “the essential effect of an assertion it to change the presuppositions of the
participants in the conversation by adding the content of what is asserted to what is
presupposed.” So when I make some assertive utterance, it has the effect (if my
conversation partners cooperate) of adding the content of my utterance to our stock of
conversational presuppositions. Call an assertion that’s heard and understood by all
parties to the conversation, doesn’t meet with any objections, etc., a successful assertion.

The following constraint on the possible contents of sentences falls out of the
Stalnakerian account of assertion:

35 Stalnaker (1981: 147) discusses essentially this problem, pointing out that if assertions, like beliefs, are
self-ascriptions of properties, then our account of assertion will have to be more complicated than seems
plausible. Either communication is less direct and more inference-laden than we might have thought, or the
contents of assertion fail to match up with the contents of the beliefs expressed. More on this sort of
problem in the next section.
ASSERTION CONSTRAINT: If C is the content of some sentence S, then successful assertions of S add C to the common presuppositions of the parties to the conversation.\(^{37}\)

The constraint seems plausible, and the picture of assertion that produces it is, as I’ve said, very attractive. But it’s trouble for centered-worlds propositions as contents of natural language sentences. (And so it’s trouble for centering features as semantic values of natural language predicates.) When I say “my pants are on fire”, I had better not be expressing the centered-worlds proposition that I believe when I believe that my pants are on fire. When I say to you, in a growing panic, “my pants are on fire” it wouldn’t serve my purposes at all if you came to believe the centered-worlds proposition. If you did, then you’d believe that your pants were on fire, and you’d start panicking, too, which would be no use to anybody.\(^{38}\)

All the evidence points to the conclusion that what I’m trying to get you to come to believe when I say “my pants are on fire” isn’t the centered-worlds proposition that I (and all the other well-informed people with burning pants) believe, but the possible-worlds proposition that Egan’s pants are on fire (which I probably also believe, along with all of the other well-informed witnesses at the barbecue). So given the assertion

\(^{37}\) As stated, this only says anything about sentences which have the same content in every context. But lots of sentences aren’t like that. When a sentence includes indexicals, the sentence doesn’t have any particular content until we specify a context of utterance. Still, the assertion constraint will be true (or at least plausible) for sentences that express the same proposition in every context, and that’s where the trouble is. Also, a similar thing is quite plausibly true for other sentences—if S has C in context T, then successful assertions of S in T add C to the presuppositions. Same idea, but a little ickier to state.

\(^{38}\) I’ll often talk about believing rather than presupposing in cases (like this one) where it’s pretty clear that the way in which a successful assertion of some sentence would get the parties to the conversation to presuppose the sentence’s content would be by getting them to believe it (and believe that everybody else believes it, etc.).
constraint, the content of (an utterance of) "my pants are on fire" had better be the 
possible-worlds proposition, and not the centered-worlds proposition.

It’s tempting to conclude from this that no assertion ever has the aim of adding a 
centered-worlds proposition to the presuppositions of the conversation. Because they’re 
the sort of thing that’s liable to be true evaluated relative to one party to the conversation 
and false evaluated relative to another, centered-worlds contents can look like lousy 
things to add to the conversational presuppositions, or to a stock of common beliefs. If 
that’s right, then, given the assertion constraint (and the plausible claim that every 
declarative sentence of natural language can be used to make assertions), no centered- 
worlds proposition would ever be the content of a natural language sentence, and so no 
sentence of natural language would attribute a centering feature to anything.

But this is too fast. There are some motivations for taking the contents of at least 
some natural language sentences to be centered-worlds propositions.

One motivation comes from the behavior of epistemic modals. There are some 
phenomena that are extremely surprising and difficult to explain if we take the semantic 
values of utterances of a sentence like “John might be in Cleveland” to be possible- 
worlds propositions, but which are nicely explained by taking the semantic values of such 
utterances to be centered-worlds propositions.39 In fact, many of the sorts of phenomena 
that motivate a centered-worlds semantics for epistemic “might” also seem to appear in 
the case of “tastes great”, “disgusting”, and other predicates that look like good 
candidates to be secondary-quality terms. I think that this is pretty good evidence that a 
centering-feature account of the primary/secondary quality distinction is on the right

39 See Egan, Hawthorne, and Weatherson (forthoming) for details.
track. However, a careful discussion of those features is a task for another paper, and so I
won’t pursue it here.

Another motivation for allowing centered worlds propositions as semantic values
of utterances in natural language—the one that I’ll emphasize here—parallels the
motivation for introducing centered-worlds propositions as the objects of propositional
attitudes.

Cases where people get lost in the library even though they know the floorplan
(and even where all of the people are) motivated us to say that the contents of belief were
centered-worlds propositions. When Carolina, lost in the Stanford library, deliberates
and gathers evidence about where she is, her deliberation and her evidence-gathering are
aimed at ruling out possibilities for her, not possibilities for the world. What she wants to
find out, first and foremost (and perhaps exclusively, if she’s knowledgeable enough) is
not which world she’s in, but which room she’s in—which location she occupies, within
the world that she may already know that she inhabits. And the best way to model the
sort of sifting and sorting of possibilities that Carolina is engaged in when she’s trying to
get oriented in the library is not in terms of sets of possible worlds, but sets of centered
worlds.

People sometimes get lost in groups. Groups of people can get lost in the library
together even when the floorplan, and the locations of all the groups of people in the
library, are common knowledge. The people in these groups can engage in collective
deliberation about where they are. Members of the group can, for example, make
assertions about where they (collectively) are, which other members of the group might
agree with or dissent from. It won’t do any good, in this kind of situation, to make
assertions about which world is actual—that could already be common knowledge. Asserting some possible-worlds proposition would be no help. Again, the best way to model the sort of sifting and sorting of possibilities that’s going on in the collective deliberations of lost groups is not in terms of sets of possible worlds, but sets of centered worlds. The claims that members of the group make, agree with, and dissent from, must have centered-worlds content, not possible-worlds content.40

So it looks like we need to have some assertions of centered-worlds propositions, and for that we probably need centered-worlds contents for the natural language utterances that are being used to make the assertions. This is enough, I think, to head off the objection that no natural language sentence ever has centered-worlds content, and therefore no natural language predicate expresses a centering feature. There’s good reason to think that not every utterance has (interesting) centered-worlds content (e.g. ‘Oakland is near San Francisco’), but the possibility of collective deliberation in lost-group cases gives us good reason to think that some of them do. And that’s enough to leave a space for centering features as semantic values of natural language predicates.

6. Self-Locating Assertion

To the extent that you’re worried about the prospects of a theory of communication that can accommodate self-locating assertion, it’s reasonable to be worried about allowing centered-worlds propositions as contents of natural language sentences. In order to assuage such worries, I’ll sketch a generalization of Stalnaker’s (1978) treatment of assertion that allows for assertions with centered-worlds contents.

40 There’s a question about whether the sorts of centered worlds we want here are the same as the sort that we want for paradigm cases of self-locating belief. Do we want worlds centered on individuals, or worlds centered on groups? More on this in the next section.
The underlying thought behind the Stalnaker approach is that in conversation, deliberation, and debate, people seek to distinguish between possibilities. The way in which they distinguish between possibilities for the world—possible ways for the world to be—has received a great deal of attention. The way in which people distinguish, in conversation, between possibilities for them—possible positions that they might occupy in the world—has received less.

The Stalnaker Framework

On the original framework, the sort of possibilities we’re distinguishing between are possibilities for the world. As the conversation progresses, if things go well, we collectively narrow down the range of live options for how the world might be. More propositions come to be presupposed by all parties to the conversation, and so the class of worlds that are compatible with the conversation’s presuppositions gets smaller.

A proposition is presupposed by a speaker iff the speaker takes it that all parties to the conversation accept the proposition, take each other to accept it, take each other to know that everyone else accepts it, etc.

A presupposition is shared iff it’s presupposed by all parties to the conversation.

A speaker’s context set is the set of worlds compatible with all of a speaker’s presuppositions.

A nondefective context is one in which all parties to the conversation have the same presuppositions, and therefore all share a common context set.

A defective context is one in which some party to the conversation presupposes some proposition that another party to the conversation does not.
The function of assertion—what Stalnaker calls the *essential effect* of assertion—is to add the content of the assertion to the conversation’s shared presuppositions, and reduce the common context set (assuming that the context is nondefective) by removing all of the worlds that aren’t compatible with what’s been asserted.

It’s natural to think about this in terms of speaker intentions—when I assert that p, I’m trying to get the other parties to the conversation to believe (or at least accept) that p. This isn’t quite right, though. We sometimes make assertions that we know won’t be accepted, with the intention of achieving some other effect. This doesn’t mean that we should revise our account of the essential effect of assertion. When we make assertions knowing that they won’t be accepted, we’re exploiting the usual role of assertion—the *essential* effect of assertion—in order to achieve some secondary effect.

Let’s look at a garden-variety example of assertion. Suppose Sarah and Liz are in the middle of their transatlantic voyage, and Liz sincerely asserts, “the Azores are 800 miles from Lisbon”. Liz is expressing a belief—a belief that the Azores are 800 miles from Lisbon—by uttering a sentence that expresses the proposition that she believes. This content is a set of possible worlds: all of the worlds in which the Azores are 800 miles from Lisbon. The effect of Liz’s assertion, if Sarah does not object (and let’s suppose that she does not) is to add the proposition that the Azores are 800 miles from Lisbon to the shared presuppositions (and, probably, to Sarah’s beliefs). After the assertion, both Liz and Sarah presuppose that the Azores are 800 miles from Lisbon. The context set is diminished by excluding all of the worlds in which the Azores are some other distance from Lisbon.
Secondary Qualities

Whether or not this helps them in their navigational tasks depends on whether they know anything about their own position relative to either Lisbon or the Azores. If they’re in a lost-in-the-library-type situation, where they both know all of the relevant geographical facts, and the locations of all of the boats with two-woman crews on the Atlantic, but are ignorant about which of those boats is theirs (and therefore which of the crews is them), then it won’t be any help at all. All that assertions about the distances between islands and cities can do is rule out some worlds as candidates for actuality. But Liz and Sarah already know (let us suppose) which world they’re in. What they need is information about where in the world they are—information that rules out not possible worlds, but possible predicaments.

Generalizing

The obvious way to generalize the framework is just to say exactly the same things, except replacing ‘possible world’ everywhere with ‘centered world’. (This counts as a generalization because what happens in the special case where all of the centered-worlds propositions under consideration are boring will be exactly what happens in the original framework.) Participants in a conversation are still collectively distinguishing between possibilities, but they’re distinguishing between possibilities for them, rather than possibilities for the world. Presuppositions will be centered-worlds propositions, and the context set will be a set of centered worlds.41

41 Note: So far as everything I’ve said so far goes, it’s still an open question which sentences have centered-worlds contents (or interesting ones). Admitting centered-worlds contents doesn’t force us to say that, for example, all sentences involving indexicals express (interesting) centered-worlds contents. All I’m committed to saying is that sentences with predicates that express (interesting) centering features have (interesting) centered-worlds contents (usually, barring fanciness about embedding, etc.).
Sarah consults her navigational equipment and sincerely asserts, “the Azores are nearby”. Suppose that “nearby” expresses a centering feature, and so the sentence expresses a centered-worlds proposition.\textsuperscript{42} Since her utterance is sincere, Sarah expresses one of her beliefs by uttering a sentence with the same content. The content of both Sarah’s belief and her assertion is the centered-worlds proposition that includes all and only those \(<i, t>, w>\) pairs such that \(i\) is near the Azores at \(t\) in \(w\). The effect of Sarah’s assertion, if Liz does not object (and let’s suppose that she does not) is to add the centered-worlds proposition that the Azores are nearby to the conversation’s shared presuppositions (and probably to Liz’s beliefs). After the assertion, both Liz and Sarah presuppose that the Azores are nearby. The context set is diminished by eliminating all of the \(<i, t>, w>\) triples such that \(i\) is not near the Azores at \(t\) in \(w\).

This gets the effect that we want—Sarah is able to express her self-locating belief by making an assertion with the same content as her belief; when Liz accepts Sarah’s assertion, she comes to share Sarah’s belief. This fits nicely with the intuitive model of communication according to which people express their beliefs in assertions which, if accepted, bring it about that the belief is shared by the other parties to the conversation. (I’m abstracting away from complications about forms of acceptance other than belief here, but I think nothing hangs on this.)

Allowing self-locating assertion requires us to recognize a new way for an assertion to be defective. Since centered-worlds propositions are liable to be true relative to some people in a world but not relative to other people in the same world, there’s a danger of asserting something that’s true (plausible, etc.) relative to some, but not all,

\textsuperscript{42} It’s almost certain that “nearby” doesn’t \textit{always} express a centering feature (at least, it doesn’t always express an interesting one). This is a potential source of trouble for a compositional semantics that delivers interesting centered worlds contents for all and only the sentences that ought to have them.
parties to the conversation. So we might have grounds to reject some self-locating assertion—to resist adding it to our stock of conversational presuppositions—not because it’s false relative to our own predicament, but because we have good reason to think that it’s false relative to the predicament of some party to the conversation, and as such a bad thing to accept, take everyone else to accept, take everyone else to take everyone else to accept, etc. There are some wrinkles here that need exploring. But for now let’s pass on to other things.

7. Two Worries

There are two worries that I’d like to address briefly before closing.

First, there’s a concern about whether the communication of centered-worlds propositions has to be *semantic*, rather than pragmatic. And if communication of self-locating content is a pragmatic phenomenon, we don’t have any reason to think that there are predicates that express centering features, and we’re left with a not-very-satisfying account of the distinction between primary and secondary qualities.

I have three responses to this worry: First response: centered-worlds semantics seems to be independently motivated. (See Egan, Hawthorne, and Weatherson (2004) for such a motivation.) Second response: even if we relegate the conveyance of centered-worlds contents to pragmatics, we still need the fancier story about

---

43 Note in particular that “tastes great”, for example, seems to give rise to the same puzzling phenomena that motivate a semantics that delivers centered-worlds propositions as the semantic values of many sentences that include epistemic “might”s.

44 Comesaño (2002) gives us another reason to want a semantics that assigns centered-worlds contents to some natural language utterances—if he’s right, “justified” expresses, at least on one of its uses, a centering feature. So, sentences like “my belief that I have a hand is justified” will express centered-worlds propositions. Comesaño’s account is very attractive—it does a lot of work in resolving some difficult puzzles about justification—and so it provides some extra motivation for a semantics that allows for utterances with centered-worlds contents.
communication. Just because the self-locating content is being communicated
*pragmatically* doesn’t mean we don’t need a story about how the communication works.
So it’s not clear how big the payoff really is for relegating the communication of self-
locating content to pragmatics. Finally, there doesn’t seem to be, in lost-group cases, the
sort of indirection or sense of moving to a repair strategy that’s suggestive of
communicating something other than the semantic content of the utterance that’s used for
the communication.

A second worry: in claiming that the communication of centered-worlds
propositions *is* a semantic phenomenon, I’ve committed myself to the existence of some
compositional semantics that delivers the right centered-worlds propositions as the
semantic values of the right sentences of natural language. And perhaps we should be
worried about the prospects for such a compositional semantics.

This is, I think, a reasonable thing to be concerned about, but there doesn’t seem
to be any principled reason to be skeptical about the prospects of such a theory. Of
course, the details will, eventually, have to be spelled out. But at least at the beginning of
the project, there doesn’t seem to be any particular grounds for pessimism.

*Conclusion*

Centering features display a lot of the behavior that’s supposed to be distinctive of
the secondary qualities—behavior that’s hard to find in any kind of *property*. So the
distinction between properties and centering features seems to be a good place to locate
the primary/secondary quality distinction. If we’re going to locate the distinction there,
and still have an interesting question about whether, for example, the colors are primary
or secondary qualities, we’ll need to say that some natural language sentences express (interesting) centered-worlds propositions. This requires a revision in our account of assertion, and to our semantic theories, which might seem to be a high cost. But the revisions are independently motivated (by the possibility of collective self-locating discussion, debate, etc.), and there seem to be some promising ways to undertake them.
Introduction

Sydney Shoemaker introduces appearance properties in order to reconcile intentionalism with the possibility of spectrum inversion without misrepresentation. In introducing appearance properties, Shoemaker does two things: he proposes a theoretical role for some family of properties to play, and he suggests a candidate family of properties to play that role. I’ll argue that his proposed candidates do not play the role as well as we would like, and suggest some new candidates. The reason for the question mark in the title is that, if I’m right, it turns out that the best candidates to play the appearance-property role aren’t properties.

1. Intentionalism and Spectrum Inversion

Intentionalists hold that the phenomenal character of an intentional mental state supervenes on its representational content. So, if intentionalism is true, there can’t be two visual experiences that are alike in their representational content, but differ in their phenomenal character.

Suppose Ernie and Vert are spectrum inverted with respect to each other. Ernie’s visual experiences when he looks at ripe tomatoes, fire engines, and cooked lobsters are phenomenologically just like ours. Vert’s visual experiences are phenomenally inverted. His experiences when he looks at ripe tomatoes, fire engines, and cooked lobsters are

---

1 Shoemaker (1994, 2000, 2001). In these papers he calls them ‘phenomenal properties’. He’s recently changed his terminology (in Shoemaker (forthcoming)), and I follow him in the new terminology here.

2 Let’s assume for simplicity that there isn’t any actual spectrum inversion, so that our visual experiences really are the same.
phenomenally like our (and Ernie’s) visual experiences when we look at unripe tomatoes, limes, and uncooked lobsters. We can put this in terms of the qualitative properties of Ernie’s and Vert’s experiences by saying that Ernie’s visual experience of Kermit is G, while Vert’s is R, where a G experience is one with the phenomenal character of our experiences of unripe tomatoes, etc., and an R experience is one with the phenomenal character of our experiences of ripe tomatoes, etc.

Suppose also that Ernie and Vert are (in relevant respects) historically and behaviorally indistinguishable. They both use color terms in the same way, they both put things into the same piles when asked to sort them by color, and so on. For example, when Ernie and Vert look at Kermit, they both say “he’s green”, and they both stack Kermit in the pile with the unripe tomatoes and uncooked lobsters when asked to sort things by color. A natural interpretation of this case is that it’s one in which Ernie and Vert are phenomenally spectrum inverted, but their visual experiences agree on the colors of things. Their visual experiences both represent Kermit as being green, even though their phenomenal character is different.³

Let’s suppose that this kind of case—in which two observers are phenomenally inverted, though their visual experiences have the same color content—is possible.⁴ Then we seem to have a counterexample to intentionalism.

However, an important part of the argument was suppressed in the previous paragraphs. Why should we agree that Ernie’s and Vert’s visual experiences have the same (overall) representational content? Here is a plausible argument: if Ernie’s and

---

³ At least, it’s natural if you’ve got certain theoretical commitments. In particular, it’s natural if you’re a sufficiently thoroughgoing externalist.

⁴ For some arguments that it’s not, see e.g. Byrne and Hilbert (1997b), Harman (1990), Hilbert and Kalderon (2000), and Stalnaker (2000). Note that if Stalnaker is right, then it’s not just that we can’t have cases like this—it’s that it doesn’t even make sense to talk about cases like this.
Vert’s experiences don’t differ with respect to *color* content, they don’t differ with respect to representational content *at all* (or at least, not in any way that’s potentially relevant). The only available place to locate a representational difference between Ernie’s and Vert’s experiences is in which colors their visual experiences represent Kermit as having. So since Ernie’s and Vert’s experiences have the same color content (both represent Kermit as being green), they have the same representational content *simpliciter*. This is the step that Shoemaker’s proposal (which will be discussed in what follows) is designed to undermine.

2. *A Bunch of Distinctions, a Problem, and Another Distinction*

Before moving on to Shoemaker’s proposal, we should pause to head off an alternative line of response to the conflict between intentionalism and the possibility of spectrum inversion without error.

Contentful experiences have *representational properties*. One kind of representational property is the property of having a certain content. There are a lot of competing ideas about what contents are, but all parties to the debate must concede that representations have *at least* the following sort of content: they make a distinction between the possibilities in which things are as they’re represented to be, and the possibilities in which things are otherwise. Maybe there are other, more fine-grained kinds of contents, too, but if there are, those kinds of finer-grained contents will each at least *determine* a coarse-grained possibility-carving content of the sort described above.

So one (fairly coarse-grained) sort of representational property is the class of possibility-carving properties. The paradigm cases of these are the properties of having a
certain possible-worlds content. Call these possibility-carving properties the *pure*
representational properties. (I will, from now on, stipulatively use ‘content’ to mean
‘possibility-carving content’.
5 I don’t mean to commit myself to any view about which
notions of content are viable, theoretically useful, etc. by making this stipulation—I just
need a less cumbersome expression than ‘possibility-carving content’, and I won’t be
employing any of the competing notions in what follows.)

When Grover believes, Oscar asserts, Miss Piggy fears, and Ernie and Vert see
that Kermit is a frog, Grover’s belief, Oscar’s assertion, Miss Piggy’s fear, and Ernie’s
and Vert’s visual experiences all share a pure representational property. They all have
contents that separate the worlds in which Kermit is a frog from the rest.

There are other kinds of representational properties. For example, *visually*
representing that Kermit is a frog, or representing Kermit, under mode of presentation K,
as a frog. Call these the *impure* representational properties. We can think of the pure
representational properties as very general properties of representing content C somehow-
or-other, and the impure representational properties as properties of representing content
C this way.
6

The intentionalist slogan is, “same representational properties, same phenomenal
character”. The distinction between pure and impure representational properties lets us
draw a distinction between two sorts of intentionalism. *Pure intentionalism* says that the
phenomenal character of experience supervenes on the pure representational properties of

———
5 For now, it does no harm to read ‘possibility-carving content’ as ‘possible-worlds content’. I’m hedging
now, though, because the difference will matter later on.
6 This more or less follows Chalmers’ distinction between pure and impure representational properties in
Chalmers (forthcoming). The notion of content I’ve employed is more restrictive than his, though. It’s
actually a bit more restrictive than I need for the points I’m about to make, but it’s difficult to state the
weaker notion in a happy way.
experience. *Impure intentionalism* says only that phenomenal character supervenes on impure representational properties.

Let’s also distinguish between *maximal* and *submaximal* representational properties of both kinds. Maximal representational properties specify the content (or the content plus the manner of representation) of an experience exhaustively, while submaximal representational properties specify only part of the content of the experience.\(^7\) So pure intentionalism says that maximal pure representational properties fix phenomenal character, while impure intentionalism makes the weaker claim that maximal *impure* representational properties fix phenomenal character.

If it’s possible to have both an occurrent thought (or some other sort of cognitive experience) and a visual experience with the same maximal pure representational properties, then pure intentionalism is false. (Since while the visual experience will have some distinctive phenomenal character, having the thought either won’t have any phenomenal character at all, or will have a very different one than that of the visual experience with the same content). Since this is almost certainly possible, pure intentionalism is almost certainly false. Call this the *problem of common content*. So it looks as if the intentionalist will have to appeal to impure representational properties in order to make her thesis plausible, given that there are some maximal pure representational properties that can be shared by experiences in different modalities (visual, cognitive, auditory, etc.).\(^8\)

---

\(^7\) More carefully, and more generally: for any kind of property \(K\), a property \(F\) is a maximal \(K\) property iff, for any \(K\) property \(G\), something’s having \(F\) either entails that it has \(G\) or entails that it lacks \(G\). That is, \(F\) is a maximal \(K\) property iff something’s having \(F\) fixes *all* of its \(K\) properties.

\(^8\) Alex Byrne (2001) discusses this sort of problem, and makes what I take to be the same distinction between varieties of intentionalism (though under a slightly different mode of presentation). He also offers some reasons for resisting this argument against pure intentionalism, so things are probably not so clear as I’ve suggested. (For one thing, while it’s clear that experiences in different modalities can share very many
So why be so worried about the sort of spectrum inversion case described above? What this case is a clear counterexample to is pure intentionalism, and we already knew that that was false, because of the problem of common content. For all that’s been said so far, it could be that Ernie’s and Vert’s experiences have the same possibility-carving content, but they each represent it in different ways. They might, for example, each represent the property being green under a different mode of presentation.⁹ So Ernie’s and Vert’s experiences would share all of their pure representational properties, while differing in their impure representational properties. Since the intentionalist already needs to appeal to impure representational properties to handle the problem of common content, where’s the harm in appealing to them here as well?

Notice a distinction between two kinds of impure representational properties. First, there are impure representational properties like visually representing that Fozzie is on stage, auditorily representing that there’s a rooster nearby, and representing in belief that no armadillo weighs more than any aircraft carrier. These are impure because they specify the sensory or cognitive modality in which the content is being represented. Call these representational properties modality-impure. Everybody has to concede that our experiences have modality-impure representational properties—it’s just a plain fact that there are a number of different representational modalities. (Not everybody has to concede that these properties are theoretically interesting, but everybody does have to concede that our experiences have them.)

---

⁹ See Chalmers (forthcoming) for such an account.
Compare these to impure representational properties like, *represents that Fozzie is on stage* (where Fozzie is presented under mode C), *represents that Fozzie is on stage* (where Fozzie is presented under mode S), and *represents that Kermit is green* (where being green is presented under mode G). Call these representational properties, which are impure because they specify modes of presentation for some of the things represented, *mode-of-presentation impure*, or *MOP-impure*. Not everyone will want to concede that our experiences have MOP-impure representational properties. It’s a controversial bit of theory that there are different modes of presentation under which we represent things (and properties). So while appealing to modality-impure representational properties to rescue intentionalism from the problem of common content is relatively theoretically innocent, appealing to MOP-impure representational properties requires some substantial theoretical commitments.

An intentionalism that allows for differences in phenomenal character due to differences in *modality-impure* representational properties, but not due to *MOP-impure* representational properties, is a close relative of pure intentionalism. It’s a theory in which pure intentionalism is true *within each modality*, though not across modalities. Call such a theory a *modality-specific pure intentionalist* theory. Within a given representational modality, phenomenal character supervenes on *pure* representational properties—no two experiences in the same modality can differ in their phenomenal character without differing in their pure representational properties.\(^{10}\) This is the sort of intentionalist theory that I’ll be concerned to square with the possibility of the

---

\(^{10}\) This distinction is the same as Byrne’s (2001) distinction between *intermodal* and *intramodal* intentionalism.
problematic cases of spectrum inversion. (I think that it’s the sort of intentionalist theory that Shoemaker is concerned with, as well.)

Modality-specific pure intentionalism is an attractive view because it seems to be the version of intentionalism with the fewest possible theoretical commitments—it will put fewer constraints on what we’re allowed to say in the rest of our theorizing than a version that, for example, appeals to MOP-impure representational properties as well.\footnote{Though one thing that might happen is that (a) it turns out that an intentionalism that appeals to MOP-impure representational properties is much less messy than one that doesn’t, and (b) we’re unavoidably committed to modes of presentation anyway, for independent reasons. If that’s how things shake out, then we’ll probably (assuming that we want to be intentionalists of some stripe) want to go with the MOP-impure version of intentionalism. But it’s not at all clear that that \textit{is} how things will shake out, and anyway, there’s an interesting chunk of theoretical space to get mapped out here—what’s the best story for a modality-impure intentionalist to tell about spectrum inversion?}

Cases of spectrum inversion without misrepresentation are \textit{prima facie} counterexamples to such a theory, since Ernie and Vert are both having visual experiences with the same (possibility-carving) content. Shoemaker’s introduction of appearance properties is a promising line of response to these apparent counterexamples.

\section{3. Appearance Properties}

Shoemaker’s strategy is to deny the claim that Ernie’s and Vert’s experiences differ with respect to representational content \textit{only} if they differ with respect to color content. Intentionalism is to be saved by supposing that, while Ernie’s and Vert’s experiences both represent Kermit as being \textit{green}, there is a representational difference with respect to the \textit{other} properties that they attribute to Kermit.

The theoretical motivation for this move is pretty clear: adding an extra layer of content to visual experience allows us to agree that Ernie’s and Vert’s experiences are the same with respect to color content, while denying that they’re the same with respect to
representational content *simpliciter*. So we can be intentionalists (modality-impure intentionalists, that is—this qualification is hereafter omitted) while allowing for the possibility of spectrum inversion without misrepresentation.

The addition of an extra layer of content also has some appeal at a less theoretical level. The intuitive idea behind the proposal is that, while Ernie’s and Vert’s experiences agree about what *color* Kermit is, they still, in *some* sense or other, represent him as *looking different*. So there should be some range of properties that aren’t colors, but still mark differences in how things look, with respect to which Ernie’s visual experience is different from Vert’s.\(^{12}\)

This identifies a role that we’d like to have some properties to play—marking a difference in how things (representationally) look to Ernie and Vert that isn’t a difference in which colors their visual experience attributes to things. Call whichever properties (if any) play this role the *appearance properties*, and call the one that Ernie represents Kermit as having ‘PG’ (for ‘phenomenal green’) and the one Vert represents him as having ‘PR’ (for ‘phenomenal red’).\(^{13}\) Now the proponent of appearance properties has an important question to answer—which properties are the appearance properties, exactly? For our purposes, we can treat this as the question, *which properties are PG and...*

---

\(^{12}\) This isn’t, by itself, a sufficient motivation for positing some further sort of content with respect to which Ernie’s and Vert’s experiences differ. Given that there is (at least in philosophical parlance) a distinction between *intentional* and *phenomenal* senses of ‘seems’ (one that Shoemaker (1982) either helpfully points out or perniciously introduces, depending on your view of the distinction), we might well want to say that there’s a similar distinction between intentional and phenomenal senses of ‘looks’, and that things look different to them only in the phenomenal sense.

\(^{13}\) These names are due to Shoemaker’s old terminology, on which the appearance properties were called *phenomenal properties*. 
PR? We should be skeptical about the prospects of the appeal to appearance properties until we’ve been given some idea of what these properties *are*, exactly.\(^{14}\)

Let’s look at the role that the appearance properties are supposed to play in some more detail. There are some constraints on what the appearance properties can be like, if they’re going to allow us to reconcile intentionalism with the possibility of spectrum inversion without misrepresentation. For ease of presentation, I’ll state them as constraints on what our two particular appearance properties—PG and PR—can be like. The generalizations should be clear enough.

DIFERENCE: Ernie’s and Vert’s experiences have to represent Kermit as having

*different* appearance properties. So it has to turn out that PG \(\neq\) PR.

It’s pretty clear why we need this. What we need in order to reconcile the phenomenal difference between Ernie’s and Vert’s experiences with the intentionalist’s supervenience claim is an accompanying representational difference. Adding more content to their visual experiences won’t help unless the content we add to Ernie’s experience is *different* from the content we add to Vert’s.

---

\(^{14}\) Is it so clear that we need to be given an informative account of which properties are the appearance properties? Yablo (1995) has argued that sometimes we ought to resist these sorts of demands for alternative specifications of which property F is—sometimes the specification that the objector wants an alternative to is the best one that’s available. This seems pretty plausible for already well-entrenched properties, but not so good for newly introduced, theoretically motivated properties. Plausibly, we’re not obliged to say which property *being red* is in any more informative terms. If, on the other hand, I introduce some property as *whichever property plays fancy theoretical role R*, then I think that I am under some obligation to show that there really is some antecedently respectable property that can play role R. In particular, in the case where I’m introducing a property as the one represented in a certain kind of experience, I’ll need to say something at least a little bit informative about how things are represented as being in those experiences, and this will require saying something about the nature of the represented property.
CORRECTNESS: Ernie’s and Vert’s visual experiences should both be representing Kermit *correctly* when they represent him as being, respectively, PG and PR.

There are two motivations for this constraint. First, we’re trying to find a way to reconcile intentionalism with the possibility of spectrum inversion *without* misrepresentation, and it would be strange to do this by simply relocating the misrepresentation—it would be strange to avoid having to attribute misrepresentation of color properties by replacing it with an attribution of misrepresentation of appearance properties. (Though maybe this is not as bad—maybe it’s better to have systematic error about weird, esoteric properties than to have systematic error about ordinary, everyday properties like colors.)

Second, and more persuasively, denying CORRECTNESS will force us to say that either Ernie’s or Vert’s visual experience, or both, systematically misrepresents the world (since one or both will represent things as having appearance properties that they don’t have). Picking one or the other to be getting things right would be objectionably arbitrary. Saying that they’re both getting things wrong requires us to attribute a lot of systematic misrepresentation (not just to Ernie and Vert, but in our own case, as well) and this ought to be avoided if possible.15

It’s worth noticing a requirement that follows from CORRECTNESS and the constraint that Ernie’s and Vert’s visual experiences agree on their color content:

---

15 Though see Boghossian & Velleman (1989). Thau (2001) leaves it open that visual experience might systematically misrepresent things as having a special range of properties that they don’t actually have.
NOVELTY: The appearance properties aren’t the colors. In particular, PG isn’t being green and PR isn’t being red.

(This is worth noticing because some of Shoemaker’s candidates wind up looking suspiciously color-like.)

The above requirements are non-negotiable. No properties that fail to satisfy them can be the appearance properties, because they can’t play the right role in reconciling intentionalism with the possibility of spectrum inversion, which is the whole reason for wanting to have appearance properties around in the first place.

There are three more features that it would be nice to have in the appearance properties, because they’re intuitively well motivated, but which aren’t absolutely required in order to do the work appearance properties were introduced to do.

SAMENESS: The appearance property that Ernie’s visual experience attributes to Kermit is the same as the appearance property that Vert’s visual experience attributes to, for example, cooked lobsters and ripe tomatoes.¹⁶

The idea here is that phenomenal spectrum inversion should go hand in hand with representational appearance property inversion. The reason why Ernie’s and Vert’s

¹⁶ Don’t be thrown off by the presence of both a DIFFERENCE and a SAMENESS requirement—I’m not imposing contradictory demands on the appearance properties. (At least, not yet.) DIFFERENCE requires that the appearance properties that Ernie’s visual experience attributes to Kermit be different from the one Vert’s experience attributes to Kermit. SAMENESS requires that the appearance property that Ernie’s visual experience attributes to Kermit be the same as the one that Vert’s experience attributes to ripe tomatoes.
visual experiences are phenomenally inverted is that they are representationally inverted with respect to certain appearance properties.  

So if Ernie’s visual experience represents paradigm green things as being PG and Vert’s represents them as PR, Ernie’s visual experience should represent paradigm red things as PR and Vert’s should represent them as PG. The appearance property, representation of which underlies the phenomenal character of Ernie’s experience of green things (like Kermit) should be the same as the appearance property, representation of which underlies the phenomenal character of Vert’s experience of red things like ripe tomatoes.

The intuitive idea behind appearance properties is that they’re the properties, representation of which underlies, explains, or maybe just is, the phenomenal character of our experiences. If we want to hang on to that claim, then we’ll want to have SAMENESS. Then we can say that the reason why Ernie’s and Vert’s phenomenology is inverted is because the appearance property content of their visual experiences is inverted. The reason why Ernie’s experiences of ripe tomatoes are like Vert’s experiences of unripe ones is because the (salient, phenomenal-character-determining) property that Ernie’s visual experience attributes to ripe tomatoes is the same one that Vert’s experience attributes to unripe ones. Without SAMENESS, we don’t have this explanation available.

17 Another way to put the requirement: it’s not just that for having a certain maximal intentional property necessitates that an experience has a certain maximal phenomenal property, but that some submaximal intentional properties (like representing something as PG), are such that any experience that has them also has some (submaximal) phenomenal property.

18 Shoemaker commits himself to this when he says that an acceptable account of the relation between phenomenal character and representational content will hold that “it is of the essence of any given quale that its instantiation by an experience makes a certain determinate contribution to that experience’s phenomenal character” (Shoemaker 1994: 28).
CONTRARINESS: PG and PR should be *contraries*—Ernie should be able to learn,
when he sees that Kermit is PG, that Kermit is *not* PR.

When Ernie learns that Kermit looks *that* way, he should learn that Kermit
*doesn’t* look (in the relevant respect) the way a ripe tomato looks, or the way Kermit
looks to Vert. Mutatis mutandis for Vert.

There’s an obvious problem with CONTRARINESS—it’s inconsistent with
CORRECTNESS. If Ernie’s experience represents Kermit as PG, and Vert’s represents
him as PR, and they’re both correct, then Kermit *is* both PG and PR, so PG and PR can’t
be incompatible in the way that CONTRARINESS seems to require. So it seems a bit
odd to look for a proposal that gives us CONTRARINESS, given that it’s inconsistent
with one of our non-negotiable desiderata.

Here are three things to say in response to this concern about whether or not we
ought to be at all concerned about trying to satisfy CONTRARINESS: First, it turns out
that the requirements of CONTRARINESS and CORRECTNESS *can* be reconciled, if
we adopt the right view about what kinds of feature are attributed to things in experience.
(I’ll explain how to effect the reconciliation in sections 5-7.) Second, Shoemaker seems
to be right about the intuitive attractiveness of CONTRARINESS, and so a proposal that
preserved it would be better than one that gave it up. Finally, given SAMENESS, there’s
a principled theoretical motivation for wanting CONTRARINESS.

There are some incompatibilities at the level of phenomenal character. The
distinctive phenomenal feature of experience that goes along with representing something
as PG is incompatible with the distinctive phenomenal feature of experience that goes
along with representing that thing as PR.\textsuperscript{19} So given SAMENESS, it looks like it’s impossible to have a visual experience that, for example, represents Kermit as both PG and PR—there’s a certain sort of representational content that’s not available to us (or at least, not available in visual experience). If PG and PR are contraries—if nothing can be both PG and PR—then it’s not very surprising that we can’t represent anything as both PG and PR. If PG and PR aren’t contraries—if things can be both PG and PR, but we can’t represent them that way—then we’ve got a brute restriction on which of the perfectly possible states of affairs (of a kind that visual experience is in the business of representing) we’re capable of representing in visual experience. This seems to be exactly the sort of consideration that makes us so sure that being red and being green are contraries. It would be nice to have CONTRARINESS in order to avoid this sort of brute restriction on our representational capacities.\textsuperscript{20}

Another desideratum that Shoemaker (1994) proposes is

CONSTANCY: The appearance properties should be features that are had by things even when unobserved.

\textsuperscript{19} The incompatibility isn’t between G and R, but between some very much more specific, extremely difficult to describe phenomenal properties. There may be a horrible problem lurking in this neighborhood, but if there is, it’s one that I won’t be able to discuss here. So I pass over it in near-silence, hoping that no one will make a fuss.

\textsuperscript{20} It’s reasonable to worry about how much we ought to trust the inference from our inability to represent a situation in which \( P \) in visual experience to the impossibility of \( P \). It’s pretty clearly not reliable across the board. We can avoid a lot of the apparent counterexamples by restricting our attention to the sorts of situations that visual experience is \textit{in the business of} representing (the colors and shapes of medium-sized physical objects, for example, but not much of anything about the very small or the abstract, and not the densities, origins, or futures even of the medium-sizers). We also ought to make sure that we don’t claim an entailment between inability to represent and impossibility.
This is certainly the least well motivated of the desiderata. For one thing, it’s not so clear how much of a pretheoretical, intuitive picture of the nature of appearance properties we really have, and we need a pretty detailed one to motivate this desideratum. For another, it’s not clear how much any picture we do have ought to constrain what we say about properties that are so theoretically motivated and distant from ordinary talk.

In fact, in more recent work, Shoemaker appeals to two kinds of appearance properties, only one of which satisfies CONSTANCY. But still, CONSTANCY does sound pretty plausible on its face—I think that the intuition Shoemaker appeals to is there, even if tentative and of uncertain authority.

4. Finding the Appearance Properties

To make the sort of move Shoemaker is proposing work, we need to find some candidates to be the appearance properties. The successful candidate will be the one (if any) that does the best job of satisfying all of the desiderata. It will have to satisfy all of the minimal desiderata—DIFFERENCE (PG and PR have to be distinct), NOVELTY (the appearance properties can’t be the colors), and CORRECTNESS (neither Ernie nor Vert is misrepresenting Kermit). It will also have to do a better job than the other candidates of satisfying the ‘bonus’ desiderata: SAMENESS (the property that Ernie attributes to Kermit is the one that Vert attributes to cooked lobsters), CONSTANCY (things can have appearance properties while unobserved) and CONTRARINESS (when Ernie learns that Kermit is PG, he learns that Kermit is not PR).

Shoemaker considers two candidates to be the appearance properties. One of them satisfies the minimal desiderata and SAMENESS, but satisfies neither
CONSTANCY nor CONTRARINESS. The other appears, at a first pass, to satisfy the minimal desiderata plus CONSTANCY. This seemingly better candidate, however, is actually ambiguous between three different proposals. On closer examination, two of these fail to meet the minimal desiderata, and the third delivers an implausible result about the appearance properties of non-actual things (and also fails to respect the intuition that motivates CONSTANCY). None satisfy CONTRARINESS, and all have some difficulty with NOVELTY.

The first of Shoemaker’s candidates is the family of properties like currently producing a G experience in some observer. (Recall that a G experience is one with the phenomenal character of our, and Ernie’s, experiences of unripe tomatoes, limes, and uncooked lobsters, and an R experience is one with the phenomenal character of our experiences of ripe tomatoes, fire trucks, and cooked lobsters.)

This satisfies DIFFERENCE, CORRECTNESS, NOVELTY, and SAMENESS, but not CONSTANCY or CONTRARINESS. It satisfies DIFFERENCE because something can be causing a G experience but not an R experience, and vice versa, so something can be PG without being PR (and vice versa). It satisfies CORRECTNESS because Kermit is causing a G experience in Ernie and an R experience in Vert, and so he really is both PG and PR. NOVELTY is satisfied, since it’s clear that the colors aren’t these kinds of properties—the colors of things don’t depend on their being presently perceived. Finally, it satisfies SAMENESS because it’s clear that the property Ernie’s

---

Note that this won’t give us a reductive intentionalism, since the representational properties are characterized in terms of phenomenal properties of experience. A nonreductive intentionalism is still worth pursuing, though. The question of the relation between representational content and phenomenal character is interesting independent of its connections to the mind-body debate, and the view that the connections are very, very close is quite attractive, independent of its consequences elsewhere.
visual experience attributes to Kermit is the one that Vert’s attributes to ripe tomatoes, and vice versa.

It doesn’t satisfy CONSTANCY because Kermit won’t be producing any G or R experiences in observers when he’s not being observed, and it doesn’t satisfy CONTRARINESS because Kermit is both PG and PR, which shows that they’re not contraries. Shoemaker (2000) dismisses these properties in favor of the next candidates because of their failure to satisfy CONSTANCY. In more recent work, he has given these properties a role in his theory, calling them *occurrent* appearance properties.

The second candidate is the family of properties like *being disposed to cause G experiences in some kind of observer*. These dispositional properties are the ones that Shoemaker endorses. Like the first proposal, this seems to satisfy DIFFERENCE, CORRECTNESS, and SAMENESS. Unlike the first proposal, it also satisfies CONSTANCY, since the dispositions are still present even when Kermit is not being observed. There’s a bit of a worry about NOVELTY: these properties look a lot like the ones that dispositional theorists of color want to identify with the colors. More on this later.

Unsurprisingly, it doesn’t deliver CONTRARINESS; PG and PR still aren’t contraries, since things can be disposed to cause G experiences in one kind of perceiver and R experiences in another. But this isn’t a fatal problem. Perhaps we ought to give up on CONTRARINESS, since it seems to be incompatible with the more important requirement of CORRECTNESS. Still, if there were a proposal that would make CORRECTNESS and CONTRARINESS compatible, that would be preferable.

---

22 Shoemaker (forthcoming).
There is another problem with Shoemaker’s dispositional candidates. Take the property PG, which is, on Shoemaker’s (2000) account, the property, *being disposed to cause G experiences in some kind of perceiver*. There are two readings of this: *being disposed to cause G experiences in some possible kind of perceiver*, and *being disposed to cause G experiences in some actual kind of perceiver*.24

There is trouble if Shoemaker opts for the first reading. Almost *everything* will be disposed to cause G experiences in *some* kind of perceiver or other. All it will take to be PG is to be potentially causally efficacious. But it would be nice if the property that Ernie’s visual experience (and not Vert’s) represents Kermit as having did more than rule out the possibility that Kermit is necessarily epiphenomenal. Kermit is fond of telling us that it’s not easy being green. Surely it’s not *that* much easier to be phenomenal green.

There is also a more principled worry in the neighborhood. It’s not just too easy to be PG. It’s also too easy to be PR. In fact, it’s *exactly* as easy to be PR as it is to be PG. Not only will all and only the potentially causally efficacious things be PG, but all and only the potentially causally efficacious things will be PR, as well. This gives us the unfortunate result that, necessarily, something is PG if and only if it’s PR. On the plausible view that necessarily coextensive properties are identical, it will turn out that PG *is* PR. But then the whole strategy collapses. If PG is PR, then we lose

DIFFERENCE—Ernie and Vert *aren’t* representing Kermit as having different

24 Terminological point: it’s plausible that kinds are properties, and so the *kinds* might exist necessarily. So it’s best to read ‘actual kind’ and ‘possible kind’ as ‘actually instantiated kind’ and ‘possibly instantiated kind’. But this is awkward to write, so I’ll continue to write ‘actual kind’ and ‘possible kind’ in the text, with the understanding that these may be read in the more careful way described above.
properties, and so we haven’t found a representational difference to ground the
phenomenal difference in their visual experiences.25

So Shoemaker should opt for the reading that says that PG is the property, being
disposed to cause G experiences in some actual kind of observer.26 This reading, too, is
ambiguous. This time the ambiguity comes from two available readings of ‘actual’. On
the first reading, something is PG at a world w iff it’s disposed to cause G experiences in
some kind of observer that exists in @, where @ is the actual world. Call the properties
we get on this reading the dispositional@ properties, and the proposal the dispositional@
proposal. On the second disambiguation, something is PG at a world w iff it’s disposed
to cause G experiences in some kind of observer that exists in w. Call these properties
the dispositionalw properties, and the proposal the dispositionalw proposal.

The difference between the two disambiguations is in the force of the word
‘actual’. On the first reading—the dispositional@ proposal—the force of ‘actual’ is to
restrict the relevant kinds of observers to ones that exist in our world. On the second
reading—the dispositionalw proposal—the force of ‘actual’ is to restrict the relevant kinds
to ones that exist in the world where the object is. (The reader may have noticed a
connection to discussions of two-dimensionalism here. The dispositional@ proposal

25 The view that necessarily coextensive properties are identical is plausible, but it’s not mandatory. Still,
two points: First, a difference only with respect to which finer-grained properties are represented still won’t
deliver a difference in pure representational properties. Second, it doesn’t seem as if we want this kind of
consideration to motivate our choice of a theory of properties. It would be better to have an account that
was neutral between theories of properties.

26 Another possible response: there’s a difference in which modes of presentation of PG (that is, PR) they
employ. But if you’re going to move to modes of presentation here, you might as well appeal to them right
away, and explain the intentional difference between Ernie and Vert as a difference in which modes of
presentation of being green they employ. (See Chalmers (forthcoming) for this kind of proposal.) The
move to appearance properties doesn’t seem to have any work to do once we’ve got modes of presentation
in the picture.
corresponds to the horizontal reading of ‘actual’, while the dispositional proposal corresponds to the diagonal reading.\textsuperscript{27})

There are two reasons why the dispositional proposal cannot be right. First, even if this move \textit{does} keep PG and PR from being necessarily coextensive, it’s just by luck. If there are enough kinds (or just the wrong kinds) of observers in the actual world, then it will still turn out that PG and PR are necessarily coextensive. In some worlds there are enough different kinds of observers that everything that’s potentially causally efficacious is disposed to cause G experiences in some kind of observer that exists there, and also to cause R experiences in some kind of observer that exists there. If the actual world is one of these, then PG and PR will be necessarily coextensive.

In other worlds, there are only observers like Ernie and observers like Vert. If the actual world is one of these, then again PG and PR will be necessarily coextensive. An object is disposed to cause G experiences in Ernie iff it’s disposed to cause R experiences in Vert, and vice versa. So if the actual world has only two kinds of observers, and they’re spectrum inverted relative to each other, an object will be disposed to cause G experiences in some actual kind of observer iff it’s also disposed to cause R experiences in some actual kind of observer. And so PG and PR will be necessarily coextensive, and so very plausibly identical, and so we won’t have a representational difference.

We should not adopt a proposal that only works if the contingent facts about what kinds of perceivers there actually are turn out the right way. A solution to the conflict between intentionalism and the possibility of spectrum inversion ought to be \textit{general}—

whether or not it’s viable ought not to hinge on these kinds of contingent facts about which kinds of perceivers there happen to be.\textsuperscript{28}

The second reason why this reading fails is that it delivers the wrong results for\textit{merely counterfactual} spectrum inversion. Suppose that (as I imagine is true) the case of Ernie and Vert is counterfactual, and that there isn’t any\textit{actual} spectrum inversion—there aren’t any actual observers that have R experiences when they look at green things. Then Vert is of a kind that doesn’t exist in \@. Kermit will be PG, because he’s disposed to cause G experiences in\textit{us} and we’re (obviously) of an actually existing kind. But Kermit\textit{won’t} be PR, because while he’s disposed to cause R experiences in Vert, Vert is not of any actually existing kind. So we haven’t satisfied CORRECTNESS. Vert’s visual experience (in \textit{w}) represents Kermit as being PR, but this is a\textit{mis}representation, since Kermit’s not disposed to cause R experiences in any kind of observer that exists in \@. So on this reading, we can accommodate\textit{actual} cases of spectrum inversion, but not counterfactual cases. But spectrum inversion cases don’t need to be actual to demand proper treatment. We need a theory that gets the right results for both actual and merely possible cases.\textsuperscript{29}

What about the dispositional\textit{w} reading—the one that says that something is PG in\textit{w} iff it’s disposed to cause G experiences in observers of a kind that has instances in\textit{w}?

A problem with this reading is that it gives the wrong results about the appearance properties of things in counterfactual situations. Take some uncooked lobster in a

\textsuperscript{28} Again, an appeal to modes of presentation might look promising here. See note 26 for why I think this isn’t a good idea.  
\textsuperscript{29} Another problem: we also lose the plausibility of SAMENESS for me and my counterfactual inverted counterpart. It’s extremely implausible that inverted-Egan, looking at Kermit’s local (presumably still green) counterpart in some counterfactual world \textit{w} and having an R experience, is representing things as being disposed to cause R experiences in some kind of observer that exists in \@. That is, it’s extremely implausible that his visual experience represents Kermit as having the appearance property that, according to the dispositional\textit{@} proposal, my experience represents ripe tomatoes as having.
counterfactual world $w$. Call him Lenny. Lenny is green, and we’re not spectrum inverted, so he’s disposed to cause G experiences in us. So he’s PG. Or at least, he ought to be PG. If something’s PG in $w$ iff it’s disposed to cause G experiences in some kind of observer that exists in $w$, then Lenny won’t be PG in worlds where there aren’t any observers like us. And this is implausible. If something’s disposed to cause G experiences in us—if, were we to look at it, we’d have a G experience—that ought to be sufficient for it’s being PG. The second reading doesn’t deliver this result.

This is a lot less bad than the problems for the dispositional reading—if we’re going to go with a dispositional story, we should definitely go with this one. It delivers a somewhat strange result about the appearance properties of merely possible things, but it’s not clear how big of a problem this really is. The more serious problem is that it doesn’t deliver CONTRARINESS. On the face of it, this is also not such a big deal, because no proposal is going to give us CONTRARINESS—it’s in direct conflict with CORRECTNESS, which is non-negotiable. It turns out, though, that we can get both CONTRARINESS and CORRECTNESS.

5. Self-Locating Content and Centering Features

Recall the discussion of centering features in the preceding chapter. Centering features are the analogues of properties at the level of centered-worlds content. While properties are functions from (at least) worlds to extensions, centering features are functions from centered worlds to extensions. This is a natural way to describe such features as, for example, being nearby.

---

30 Let’s suppose that lobsters’ visual phenomenology, if any, is such that other (uncooked) lobsters aren’t disposed to cause G experiences in them.
Two things about centering features should be pretty clear: First, they’re useful things to have around for describing centered-worlds contents. In the same way that it’s useful to talk about the properties that people attribute to things in describing their beliefs about which world is actual (try specifying the content of my belief that Kermit is green without appealing to properties), it’s useful to talk about the centering features that people attribute to things in describing their self-locating beliefs (beliefs about which predicament they’re in). Second, there’s nothing mysterious about them. Properties are one kind of set theoretical object—functions from worlds to extensions—and centering features are set-theoretical objects of another, analogous kind—functions from centered worlds to extensions.31

If we grant that we should take the contents of at least some kinds of representation to be sets of centered worlds rather than sets of worlds (as seems incredibly plausible), then a third point becomes clear: A difference in the attribution of centering features is a representational difference. If I my beliefs represent Oakland as nearby and yours represent it as far away, then there’s a difference in the representational content of our beliefs. This suggests a new set of candidates to be the appearance properties: perhaps Ernie and Vert differ in which centering features they attribute to Kermit. I will develop this proposal in the next section.

31 Or at least they stand in some very intimate relation to such set-theoretic objects. The identity claims are more controversial. But even if the identity claims are false, centering features don’t seem to be any more mysterious than properties.
6. Appearance Properties as Centering features

Here is a first pass at a proposal for what PG is: it’s the centering feature, being disposed to cause G experiences in me (or possibly, being disposed to cause G experiences in perceivers of my kind).  

It is important not to ‘read through’ the indexical. It’s tempting to understand the claim that Ernie’s visual experience attributes the centering feature, being disposed to cause G experiences in me to Kermit as the claim that Ernie’s visual experience attributes the property, being disposed to cause G experiences in Ernie to Kermit. But this would be a mistake. It’s important that the centering feature, being disposed to cause G experiences in me, is something that combines with an object to determine a centered-worlds proposition, not a possible-worlds proposition. It’s also important that it’s something that Ernie’s visual experience represents Kermit, limes, and uncooked lobsters as having, and that Vert’s visual experience represents ripe tomatoes, fire engines, and cooked lobsters as having. The property that we get by ‘reading through’ the indexical in a given case doesn’t have either of these features.

This meets all of the desiderata and avoids all of the problems of the other accounts. Let’s look at each desideratum in turn:

DIFFERENCE: PG is plainly not the same centering feature as PR. The set of possible predicaments in which Kermit is disposed to cause G experiences in me (or in

---

32 It’s a first pass because, for reasons we’ll get to in a moment, it can’t be quite right as it stands.
33 Of course we also get the occurrent versions—is causing a G experience in me now. So if (as Shoemaker does in more recent work) we want to appeal to both occurrent and dispositional appearance properties, we can still do that.
34 That is, it’s important that it combine with an object to determine an interesting centered-worlds proposition, to go back to the terminology of note n again.
35 As a consequence, these properties are also bad candidates to be the appearance properties. Their prominent defect is that they don’t satisfy SAMENESS.
observers of my kind) is clearly not the same as the set in which he is disposed to cause R experiences in me (or observers of my kind). So DIFFERENCE is satisfied.36

CORRECTNESS: Neither Ernie nor Vert is misrepresenting Kermit—Kermit really is disposed to cause G experiences in Ernie, and he really is disposed to cause R experiences in Vert—so CORRECTNESS is satisfied.

NOVELTY: If we agree that Ernie’s and Vert’s experiences agree on their color content, then these centering features can’t be the colors. Also, it seems quite plausible (though it’s not universally agreed to) that the colors of things are perfectly objective properties of them, and that which colors things have is not observer-relative.

If we thought that the colors of things were observer-relative, then we shouldn’t have been worried in the first place—we should have resisted the claim that it was possible to have spectrum inverted subjects whose visual experiences agreed on the colors of things.

SAMENESS: The centering feature that Ernie’s visual experience attributes to Kermit (and to unripe tomatoes) is the same one that Vert’s attributes to, e.g., cooked lobsters and ripe tomatoes.

CONTRARINESS: When Ernie learns that Kermit is PG, he learns that he’s not PR (just as when he learns that San Francisco is far away, he learns that it’s not nearby). So CONTRARINESS is satisfied, as well.

This is compatible with Vert, when he learns that Kermit is PR, learning that he’s not PG (just as Vert could, if he lived in Oakland, learn that San Francisco is not far away by learning that it’s nearby). In the case of centering features (like nearby and far away),

36 We probably also need an ‘in present circumstances’ qualification.
unlike in the case of properties (like green and red), CONTRARINESS and CORRECTNESS are compatible.

CONSTANCY: They’re dispositions, so there’s no problem about Kermit’s having them even when he’s not being observed.

I conclude that centering features are the best candidates to play the appearance-property role. If we are going to go in for a Shoemaker-style solution to the conflict between intentionalism and the possibility of spectrum inversion, we should say that the representational difference that grounds the phenomenal difference between inverted subjects is a difference with respect to which centering features (of the kind described above) are represented by their visual experiences.

Unfortunately there is a lingering problem about deviant dispositions (pointed out to me by Alex Byrne). Suppose that Kermit was, in addition to having his usual effects on Ernie’s perceptual system, was also disposed to cause Ernie to hallucinate a ripe tomato hovering in the air three feet to Kermit’s left. This would be strange, but it’s certainly not impossible. In this case, two bad things would happen. First, Ernie would correct in representing Kermit as PR, since Kermit really would be disposed to cause R experiences in Ernie. This is bad, because only red things ought to be PR to non-inverted observers. Second, Ernie would be correct in representing Kermit as both PR and PG, (since he really is disposed to cause both G and R experiences in Ernie), which shows that the first pass proposal doesn’t really give us CONTRARINESS.

Two strategies for revising the first pass proposal to avoid this problem suggest themselves. First, we could say that something is PG iff it’s disposed to cause G experiences in me in a non-deviant way, and start looking for some halfway satisfying
way of cashing out 'non-deviant'. Second, we could say something quite a bit fancier about what it is that PG things are supposed to do—if we’re fans of visual fields, the thing to say is something like, *an object is PG iff it’s disposed to cause the bit of the visual field responsible for representing it to go green*. I suspect that it is possible to make good on this sort of proposal without actually committing oneself to green’ patches on visual fields (by talking about very specific phenomenal properties of experiences and the relations of similarity between them), but I won’t argue for that here.

### 7. Another Constraint, and a Big Problem

Centering features are the best candidates to play the appearance-property role, but are they good enough candidates? There’s another non-negotiable constraint on what the appearance properties can be like, which I’ve suppressed so far because it generates a pretty serious problem—maybe nothing can play the role appearance properties were introduced to play.\(^{37}\)

**NECESSITY:** The connection between appearance property content and phenomenal character must be *necessary*.

To use Byrne and Hilbert’s (1997b) phrases, the connection between a visual experiences being ‘PG-feeling’ and its being ‘PG-representing’ must be necessary. Otherwise—if the connection between an experience’s representing something as being PG and its having the phenomenal character distinctive of our experiences of green things

\(^{37}\) Actually, the constraint is entailed by SAMENESS, but it’s also independently motivated, and while the stronger claim of SAMENESS isn’t quite non-negotiable, this one is.
Appearance Properties?

(call this phenomenal character G) were contingent—there could be another, higher-order spectrum inversion problem. If the connection between appearance property content and phenomenal character were contingent, we could, for example, have two subjects, both of whose visual experiences represent Kermit as being PG, but whose visual experiences have different phenomenal character. If this kind of case is possible, then we’ve made no progress—we’ve again got Ernie and Vert having experiences with the same content, but different phenomenal character.

Given a plausible externalism, it could happen that some perceiver auditorily represents things as being PG. (This might happen if, for example, all and only the things that are disposed to cause G experiences... make a certain noise when struck.) In this case, representation of PG wouldn’t go along with phenomenal character G. We can avoid this problem by moving, as we probably already need to do, from pure intentionalism to a modality-specific pure intentionalism. It’s more plausible that visually representing something as PG always goes along with G. But it’s not certain—maybe all and only PG things cause R experiences in certain nonstandard viewing conditions, for example.

This is potentially a very serious problem for any proposal that attempts to dissolve the tension between intentionalism and the possibility of spectrum inversion by adding a further layer of content. If the connection between representing the properties in the extra layer and the phenomenal character of experience is contingent, then we won’t have made any progress. But it’s hard to see how it could be necessary—we have to rule out the possibility of any phenomenally nonstandard ways of representing the appearance properties. And while, given the nature of appearance properties (or appearance features)
like PG, it’s quite plausible that having a visual experience with a certain phenomenal character is sufficient for representing something as having PG, it’s harder to see why it’s necessary. And it had better be necessary, or else we’ll have PG-representing experiences that don’t have G, and we’ll still have a counterexample to intentionalism.

I point out this problem only to set it aside—important as they are to the viability of the whole Shoemakerian project, I won’t address general questions about the satisfiability of NECESSITY here. All that I will say is that the candidates that I’ve proposed have at least as good a claim to satisfy it as Shoemaker’s do.\(^{38}\) Discussion of just how good any of those claims are will need to be left for another occasion. So my conclusion is conditional: if you’re going to go in for a Shoemaker-type strategy, then you should say that the extra features represented in Ernie’s and Vert’s visual experience are centering features. For these reasons, though, I’m hesitant about claiming that a Shoemaker-type strategy is the best way to go.

Conclusion

Appearance properties were introduced to play a certain role. They’re supposed to mark a difference in how spectrum inverted subjects represent colored objects like Kermit. I’ve argued that the best candidates to play this role aren’t properties, but centering features. So while it’s not clear whether or not there are any appearance properties—since while there are things that play the role, those things aren’t properties—the intentionalist should be happy. There is a representational difference between Ernie and Vert: their visual experiences attribute different centering features to

\(^{38}\) Actually, I suspect they have a better claim, since it seems more plausible to me that representation of centering features is intrinsic than that representation of properties. But I won’t argue for that here.
Appearance Properties?

Kermit. And as far as reconciling intentionalism with the possibility of spectrum inversion without misrepresentation goes, it’s just the presence of the representational difference, and not the details of what kind of difference it is, that matters.
References:


Locke, J. (1690) *An Essay Concerning Human Understanding*
Shoemaker, S. (2001) Introspection and Phenomenal Character, *Philosophical Topics*
Shoemaker, S. (forthcoming) 2002 Whitehead lectures