Colors, Functions, Realizers, and Roles

Jonathan Cohen*

You may speak of a chain, or if you please, a net. An analogy is of little aid. Each cause brings about future events. Without each the future would not be the same. Each is proximate in the sense it is essential. But that is not what we mean by the word. Nor on the other hand do we mean sole cause. There is no such thing.
— *Palsgraf v. Long Island R.R. Co.*, 248 N.Y. 339, 351-52, 162 N.E. 99, 103 (1928) (Andrews, J., dissenting).

Who could forget the great functionalism debates in 1970s and 1980s philosophy of mind? Certainly not philosophers of perception, many of whom have recently proposed understanding colors in various functionalist terms. Perhaps inevitably, many of these theorists have followed an argumentative strategy used earlier in philosophy of mind applications of functionalism: they have urged that general considerations about causal efficacy can be used to decide in favor of realizer rather than role versions of functionalism.

Speaking for myself, I was not persuaded by that argumentative strategy in the philosophy of mind; nor do I find it any more convincing in its newer role in debates about color ontology. In this paper I want to say why, and to argue that role functionalism about color (*inter alia*) has more going for it than many have allowed.

I'll begin (§1) by reviewing the ways in which philosophers have applied functionalist ideas in discussions of color ontology. Then I'll present the causal arguments against role functionalism in the context of a dilemma that threatens functionalism more generally speaking, and show how it arises for the special case of functionalist theories of color (§2). Next, I'll consider several attempts to avoid the dilemma and conclude that these responses are unsatisfactory (§3). Finally, I'll attempt to resolve the dilemma directly by answering the causal objections against role functionalism (§4).

1 Color Functionalism

The resurgence in attention to matters of color ontology has provided a new locus for philosophical dispute that, it has seemed to many, is ripe for application of functionalist ideas.

---
*Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

The relevant part of the discussion about color ontology begins with the hope of identifying color properties with particular physical essences, and the discovery that this hope is dashed against the empirical finding (described in Campbell (1969), Hardin (1988), and Nassau (1980)) that color properties cross-classify the kinds of physical theory. In light of this finding, many have offered the usual functionalist stratagem of construing colors as multiply realizable functional properties unified by their functional roles. Moreover, it has seemed to many that the most promising functional roles in terms of which to understand colors involve the function of looking certain ways to certain kinds of subjects in certain kinds of circumstances.

Several variants on the basic color functionalist theme have emerged.

The most prominent is the classic dispositionalist view that *being red* is the disposition to look red to a normal subject in a normal perceptual circumstance.[1] This view should count as a functionalist view about color because it understands colors in terms of a functional role — namely, a functional role they play with respect to certain subjects/perceptual systems.

However, there are alternative (non-dispositionalist) ways of construing colors in terms of functional roles. Thus, there is a family of views defended in different forms by Jackson and Pargetter (1987), Jackson (1996), Jackson (1998), McLaughlin (2003b), McLaughlin (2003a) and Cohen (2003), according to which colors are (not the dispositions themselves, but) the properties that confer on their bearers the dispositions to look colored to certain visual systems in certain conditions. Moreover, there are functionalist views that differ from dispositionalism by appealing to very different sorts of functional roles. Thus, the "enactive" view of Thompson et al. (1992) and Thompson (1995) proposes to understand colors in terms of the varying (ecologically described) functions performed by the visual systems of different species. Or, again, Matthen (2005) defends the view that colors are constituted by their role in facilitating certain specialized epistemic and non-epistemic actions carried out by organisms possessing color vision.

Once again, all these positions should be counted as versions of color functionalism since they all understand colors in terms of the functional roles they play vis-a-vis certain kinds of subjects/perceptual systems. They differ in exactly which functional role vis-a-vis subjects/perceptual systems they make central, but agree in thinking that it is some such functional role, rather than, say, a particular sort of material constitution, that is essential to color properties (cf. Cohen (2003)).

---

[1]This formulation is schematic: it leaves open how we are to understand dispositions, what it is for something to look red to a subject in a circumstance, what a normal subject is, and what a normal circumstance is. Defenses of instances of this schema occur in McGinn (1983), Peacocke (1984), and Johnston (1992); some also find dispositionalist views in the writings of Galileo, Boyle, Newton, and Locke.

2   Functionalist Troubles

2.1   Troubles for Functionalists

Functionalism provides an attractive and well-understood theoretical framework for thinking about colors. However, it is unclear whether the benefits of the framework outweigh its costs. One reason for concern on this score comes from a classic anti-functionalist dilemma revolving around a distinction between two importantly different forms of the theory: realizer functionalism and role functionalism (on this contrast, see also Block (1980), 177–181).

Realizer functionalists think of the task of giving a theory of Fness in two stages (cf., Armstrong (1968), Armstrong (1970), Lewis (1972)). In the first, philosophical, stage, one identifies and spells out the F role for the particular F at issue. In the second, empirical stage, one looks to empirical investigation to discern what entities actually play the F role, and identifies Fness with these realizers of the F role. For example, consider pain (in the usual, overly simple and schematic way). At the first stage a realizer functionalist might identify the pain role as that of causing avoidance behavior and 'Ouch!'-utterances. At the second stage we might find that C-fiber firing in fact realizes the pain role. According to the realizer functionalist about pain, pain just is whatever entity is identified at the second stage as the realizer of the role adduced at the first stage — on the envisaged outcome, this would mean that pain just is C-fiber firing. Crucially, on this view, the pain role is merely an aid to the identification of what pain really is (viz., the realizer of the role); the role is not itself part of what pain really is, as it were. (Indeed, the formulation of realizer functionalism I've given specifically allows that the properties mentioned in the specification of the pain role can be contingently associated with the realizer.)

Role functionalism, in contrast, holds that Fness is the property of having some or other property (distinct from Fness) that plays the F role (cf., Putnam (1967), Block and Fodor (1972)). Fness is, according to role functionalism, a higher order property — one that involves quantification over properties that realize the F role. The F role is, therefore, no mere inessential identificatory aid for the role functionalist, as it was for the realizer functionalist. Quite the contrary: role functionalism earns its name by making the F role a central, essential element of Fness itself.

With this distinction in mind, we are now in a position to appreciate the promised anti-functionalist dilemma, which is helpfully brought out in Yablo (1995) in terms of a conflict between two platitudes about Fness. On the one hand, our account of Fness should satisfy the platitude of *commonality* — that Fness (whatever it turns out to be) "is shared by all and only Fs" (Yablo (1995), 482). On the other, our account of Fness should satisfy the platitude of *causality* — it should allow that Fness "has F-ish causal powers" (482). The problem, Yablo urges, is that neither realizer nor role versions of functionalism is well-placed to secure both of these desiderata — the thought is that realizer functionalism can't deliver commonality, while role functionalism can't deliver causality.

Consider realizer functionalism first. One of the things that initially motivated functionalism about Fness, recall, is the thought that, even if the Fs are physically heterogeneous, they are united by sharing the F role. But according to realizer functionalism, that shared role is a mere (and inessential) guide to the identification of Fness, rather than an essential element of Fness. Fness itself, on this view, is whatever plays the F role, and (as early functionalists were eager to point out) there is reason for thinking that many different things play the F role. In effect then, realizer functionalism dispenses with the functionalist's best hope for securing commonality by relegating the F role to the status of a mere identificatory aid. Thus, what the Fs share is not what the realizer functionalist identifies with Fness; and what the realizer functionalist identifies with Fness is not shared amongst all and only the Fs.

But if realizer functionalists have troubles with commonality, it can seem that role functionalists have even worse troubles with causality. According to role functionalism, Fness is the property of having some or other property that plays the Fish role, where the latter is given at least partly in terms of its causal effects. That is, Fness is the property of having some or other property (some realizer of Fness that is, of course, numerically distinct from Fness) that causes $G_1 \ldots G_n \ldots$. But does Fness, so construed, have the Fish causal powers? I.e., does the property of having some or other realizer property that causes $G_1 \ldots G_n$ — does *that* property cause $G_1 \ldots G_n$? Two related reasons have convinced philosophers that it cannot, and instead that what causes $G_1 \ldots G_n$ are the various realizers of Fness that role functionalists insist on distinguishing from Fness.

The first reason is that the Fish causal powers are intimately but trivially bound up with the role functionalist understanding of Fness — just as, in Molière's famous jape at the Aristotelians in *Le Malade Imaginaire*, the effect of inducing sleep is intimately but trivially bound up with the concept of a *virtus dormitiva*. The reason this gives for doubting the causal efficacy of Fness comes from the (Humean) thought that, whatever causal connections are, they should not be so trivial or definitional, lest the respectable empirical project of discovering causes should become no longer respectable. Here is Frank Jackson's version of this point, expressed as a reason to deny causal efficacy to *fragility* (where the latter is construed functionally/dispositionally):

> . . . to allow that fragility itself, as opposed to its categorical basis, causes breaking on dropping, would be to allow that there are properties that have causal powers essentially. If fragility does or would do the causing in one world, then it does or would do the causing in any world. In every world the fact of being fragile is intimately connected with the fact of breaking on dropping, and if that intimate connection counts as fragility causing or having the power to cause breaking in any world, it will so count in every world. Clearly, the same goes for functional properties in general — to allow them as causes is to violate a good Humean principle (Jackson (1996), 394; cf. Yablo (1995), 482).

Thus, Jackson invites us to conclude, it must be that it is the categorical basis of *fragility*, rather than *fragility* itself, that causes breaking on dropping.

The second reason for worrying that role functionalism makes its target causally impotent comes from concerns about causal competition between the role property and the realizer property pressed persistently by Jaegwon Kim (Kim (1989a), Kim (1989b), Kim (1993a), Kim (1998)). The thought here is that, if there is an extended causal process going on (viz., the causal process of bringing about the effects that constitute the F role) in which the exemplification of Fness plays a part, it seems reasonable to think of the lower level realizers of Fness as playing that part. But if this causal work is being done by F's realizers (on every occasion), then it looks to threaten causal overdetermination to attribute causal efficacy (with respect to the very same effects) to F, assuming F is numerically distinct from its realizers. So, on the reasonable assumption that we don't want to accept this sort of causal overdetermination — as Jackson calls it, this "curious and ontologically extravagant kind of overdetermination" (Jackson (1996), 202) — we seem compelled to give up the attribution of the causal powers at issue to F.

Might a role functionalist resist this line of argument by insisting that, while there is genuine zero-sum causal competition, it is the roles that win that competition for causal powers rather than the realizers? It is true that if the argument from causal competition is to be effective against the causal potency of roles, we need a reason for giving an initial leg up to the realizer; obviously, if we begin with the assumption that the role is efficacious, we have just as good or bad an argument against the causal potency of the realizer. However, for (role and realizer) functionalists, there is good reason for thinking that the particular causal potency at issue is non-negotiably something that belongs to the realizers. Namely, it is presumably required for thinking of the realizers as realizers that they causally bring about the effects in terms of which F is characterized.[2]

To repeat, then, the worry is that the (undisputed) causal power of the realizers to bring about the effects constitutive of the F role leaves no room for the roles to have that causal power. But since role functionalists identify Fness with those roles, this means they fail to make room for the causal power of Fness in bringing about the effects constitutive of the F role. And that is a clear violation of the causality platitude.

2.2 Troubles for Colors

Like other species of functionalism, color functionalism comes in role and realizer flavors. Of the works listed above, Jackson and Pargetter (1987), Jackson (1996), Jackson (1998), and McLaughlin (2003a) defend realizer functionalisms, while the views of McGinn (1983), Peacocke (1984), Johnston (1992), and Cohen (2003) are best construed as versions of role functionalism. Unsurprisingly, the dispute amongst these views has been carried out partly by appealing to concerns about causality and commonality.

---

[2]Another reason might come from the *virtus dormitiva* objection; this gives one way of understanding the two arguments as being related to one another.

For example, Jackson attacks (dispositionalist versions of) role functionalism by giving a version of the *virtus dormitiva* argument in (Jackson (1996), 203–204), and a version of the causal competition argument in Jackson and Pargetter (1987), Jackson (1996), and Jackson (1998), ch4 (cf. Shoemaker (1990), 116). Another version of the *virtus dormitiva* argument (similarly aimed, in the first instance, at dispositional rather than role functionalist theories of color), is given by McLaughlin:

> In implying the falsity of the claim that redness is the disposition to look red, our functional analysis is, I believe, faithful to our common conception of redness, for it is part of that conception that redness disposes its bearers to look red. The disposition to look red doesn't do that. Nor does the second-order property of being a basis for the disposition. Only a basis for the disposition does (McLaughlin (2003b), 480).

Correspondingly, considerations about commonality have been used to argue against realizer functionalism:

> When Armstrong says that the red-role is played by a physical property, he is only playing a scientific hunch; there *could*, he admits, be "*irreducibly* diverse causes in the physical surfaces bringing about identical colour-appearances for human observers." But where Armstrong puts this forward as a sort of doomsday scenario in which colors are reduced to the status of pseudo-qualities, nowadays it is thought to be more or less the situation.... Something similar is of course the standard line on suffering things — they too have nothing but a causal syndrome in common (Yablo (1995), 481; cf. Cohen (2003), 29–30).

Since (at least in the color literature) recent debate has seen more discussion of the causal objections against role functionalism, it is worth emphasizing the seriousness of the trouble Yablo is pointing to in the quoted passage. I suppose it would be bad enough if the relevant realizers turned out not to satisfy the commonality desideratum in non-actual worlds (on standard Kripkean assumptions about identity statements involving rigid designators being necessary if true). But, as Yablo insists, our best physical evidence is that the non-commonality of the realizers is actual, endemic, and severe (cf. Nassau (1980), Nassau (1983)).

To their credit, recent realizer functionalists about color have been admirably clear about the empirical bet they're making: they admit candidly that they would accept some sort of color eliminativism — an extremely revisionary alternative, I take it — if they lose this bet. Thus, Jackson (1998) admits that colors are (as a matter of empirical fact) disjunctive, but hopes that "the disjunction is not excessively disjunctive" (108), and allows that if things don't work out according to this hope, then "we would have to declare colour a pervasive illusion" (112). Likewise, McLaughlin (2003b) adds a commonality-ensuring clause to his functional characterization of color properties (479), and insists that, if it should

turn out that there is no common realizer of the relevant role (i.e., if nothing simultaneously satisfies his conditions of being a realizer of the functional role *and* commonality amongst the realizers), then nothing is colored, and color projectivism is true (481). But why, in the face of the empirical counter-evidence reported by Nassau, should anyone have any confidence at all in this bet?

The realizer functionalist might hope to answer this challenge by pointing to the idea that, even if looking red is implemented by a diverse set of physical structures, they might all be unified in respect of the way (as a matter of contingent fact) they happen to affect light (cf. Jackson (1998), 109ff, McLaughlin (2003b), 497–499). Unfortunately, this reply is unsatisfying. Recall that the commonality desideratum requires not merely that the *F*s turn out, *de facto*, to have something or other in common; it requires that *F*ness (as spelled out by a candidate account of that property) should be shared by all and only the *F*s. Here's an analogy. Suppose I desire to own every can-opener ever made in order to complete my collection. It goes little distance toward satisfying the commonality desideratum with respect to the property *being a can-opener* to cite the fact that all of its physically heterogeneous instances contingently have the property of being desired by me. What is wrong with that attempt to secure commonality with respect to *being a can-opener* is that what it cites as common between instances is a feature that is plainly not constitutive of *being a can-opener* (on anybody's account). But the present suggestion on behalf of the realizer functionalist fails in the same way. It doesn't show that the physically heterogeneous realizers share that which, according to the realizer functionalist, is constitutive of *being red*. Rather, it shows that the realizers share something else — a feature that is not only contingent, but (by the realizer functionalist's own lights) not constitutive of *being red*.

There's another curious point about the response now under consideration. On that response, the something else in terms of which the diverse realizers are said to be alike is a causal syndrome/functional role that involves their effect on light. But, come to think of it, we already knew that the diverse realizers share a causal syndrome/functional role: they are all, after all, realizers of the reddish role. For the realizer functionalist, that latter shared causal syndrome isn't enough to answer the commonality desideratum; it won't provide an answer unless we take that causal syndrome to be constitutive of *being red*, which would mean giving up on realizer functionalism in favor of role functionalism. But it seems that the proposal we're now considering is no better. For the shared causal syndrome (assuming it is, indeed, shared) that involves the effect distinct realizers have on light isn't enough to answer the commonality desideratum unless we take *that* causal role to be constitutive of being red, which, as before, would mean giving up on realizer functionalism in favor of role functionalism.[3]

In the wake of the failure of this response, there are two distinct threats to the realizer functionalist. First, her ability to meet the commonality desideratum is held hostage to the empirical fortune of how unified the class of realizers

---

[3]It would also mean giving up on the basic thought that colors are somehow constituted in terms of their effects on subjects/visual systems — a thought that, we had said, provides motivation for realizer and role functionalists alike.

of the relevant functional role turns out to be; that is worrisome in itself insofar as induction over philosophers' past bets about empirical outcomes is any guide. The second concern is the straightforward difficulty that current empirical evidence tells unambiguously against the commonality of the realizers; needless to say, current empirical evidence may be wrong about this, but it seems inadvisable to plump for a philosophical theory that depends on our best empirical accounts turning out wrong.

2.3 Summing Up

Commonality is a *prima facie* problem for the realizer functionalist, insofar as (on the empirical evidence) the realizers of the relevant functional roles look to be extremely heterogeneous except in respect of their causal effects; but, for the realizer functionalist, causal effects are inessential to the analysandum. Consequently, on her view, what is common to all and only the Fs is inessential to Fness, and what is essential to Fness is not common to all and only the Fs.

On the other hand, it is doubly unclear how the role functionalist can allow for the causal efficacy of Fness: it is not clear how roles could have the Fish causal powers the role functionalist wants them to have without thereby overdetermining their effects, and it is unclear how roles could have the Fish causal powers since the former are trivially connected to — indeed, constituted in terms of — the latter.

What's a functionalist to do?

3 Desperate Measures

One thought would be that we should escape the dilemma altogether by finding an alternative to both role and realizer versions of functionalism. In this section I'll review some suggestions about how we might do that, and find them wanting.

3.1 Color Terms and Semantic Diversity

One proposal for escaping the dilemma attempts to secure the virtues of both role and realizer functionalism by holding that 'red' in our mouths can vary in interpretation between the role and realizer readings. One simple version of this view would be to say that 'red' is lexically ambiguous in the way that 'bank' is — i.e., it is ambiguous between two (otherwise) unrelated readings, as 'bank' is ambiguous between the side-of-a-river reading and financial-institution reading. A more sophisticated version of the semantic diversity view (inspired by the account of pain in Lewis (1980)) would be that the lexical item 'red' makes a univocal but incomplete contribution to the semantic value of all its hosts, and that that contribution has parameters that get their values from features of the context to supply a complete, and intercontextually varying, interpretation. On this picture, a given utterance of the word could end up having different interpretations depending on how its parameters are contextually filled.

I think the first version, on which 'red' is lexically ambiguous in the way that 'bank' is — it has two, completely unrelated, meanings — is implausible. For one thing, if this view were correct, one would expect that the two readings of the putatively ambiguous term should fail to coincide in their lexical expression in other languages: for if the coincidence of lexical expression for the two readings in L is *merely* coincidental, then it would be unlikely that that coincidence should be sustained in languages other than L (Kripke (1977)). But this prediction about 'red' is, as far as I can see, not borne out by the data. A second standard test of lexical ambiguity that can be used against the current proposal concerns VP ellipsis. The thought is that expressions susceptible of distinct readings must be interpreted consistently across VP ellipsis, and that where they are not the result is anomalous. Thus, (1) can't be made true by John's first visiting a riverside and then a financial institution; likewise, (2) is anomalous.[4]

(1) **John visited one bank and then another.*

(2) **The tailor pressed one suit in his shop and one in the municipal court.*

If 'red' really were lexically ambiguous in the way that 'bank' is, then we should expect to find anomalous utterances of

(3) *This circle is red and this square is too.*

Since we do not, we have reason to reject the ambiguity proposal under consideration.

The more complex, Lewis-inspired thesis is a slightly harder case. It is not vulnerable to the argument that appeals to other languages because it treats the divergent interpretations of 'red' as systematically related to one another; if this were true, it would explain the (putative) fact that the relevant lexical coincidences are sustained across languages. That said, the proposal is susceptible to the argument from VP ellipsis. For the constraint on interpretation across VP ellipsis also applies to indexicals and other expressions whose semantic interpretation is constituted in part by contextually supplied values: the interpretation of context-sensitive expressions must be invariant across VP ellipsis.[5] Thus, (4) lacks an interpretation on which Mary and Sally kick different individuals, and (5) lacks an interpretation on which the place that Mary likes is distinct from the place that Sally likes.

(4) *Mary kicks me and Sally does too.*

(5) *Mary likes it here and Sally does too.*

In light of this constraint, then, consider the Lewis-inspired semantic diversity thesis according to which the semantics of 'red' is like an indexical or demonstrative in having a univocal but schematic semantic value that is completed

---

[4](2) is from Bach (1998); cf. Zwicky and Sadock (1975).

[5]Cf. Stanley (2003) for further discussion of the phenomenon and attempts to deploy it against the commitments taken on by contextualist treatments of the Sorites paradox.

by some feature or features of the context. If this view were correct, then we should expect that 'red', like other context-sensitive expressions, is semantically invariant under VP ellipsis. To test this prediction, let a Martian-tomato be a tomato that exemplifies the red role in virtue of having a new and heretofore unencountered realizer, and let a mad-tomato be a tomato that exemplifies the very property that in a standard ripe tomato is a realizer of the red role, but such that this object does not exemplify the red role. The Lewis-inspired view predicts, incorrectly, that (6) should be anomalous:

(6) *The Martian-tomato is red and the mad-tomato is too.*

Pending some reason for thinking that this standard test is defective in its application to the case, this gives us reason for rejecting the Lewis-inspired ambiguity proposal.

There are two other reasons for being dissatisfied with both semantic diversity proposals. One is that such proposals seem, in one clear sense, too easy to be trusted. After all, in literally any situation where a theory of Fness needs to satisfy n desiderata, it will be possible to construct a theory according to which our word for Fs is n-ways semantically diverse between a range of properties, each of which satisfies just one of the n desiderata. Because semantic diversity theories can be generated by such a general recipe, it seems advisable to demand independent motivation (i.e., motivation other than the possibility of accommodating theoretical desiderata) before accepting one.[6] Again, as far as I can see, no such independent motivations have been supplied in the case at hand.[7]

A second, and even more important, reason for dissatisfaction with semantic diversity theses is that they seem not to make much of an advance against the dilemma we have been considering. Our problem was that we wanted to respect both the platitudes of causality and commonality, and that neither role functionalism nor realizer functionalism seemed able to provide for both of them simultaneously. But the semantic diversity view (in either version) seems no better in answering this problem. This view (in either version) does not provide a single property that satisfies both of our desiderata — it provides two different

---

[6]Here I am agreeing with Kripke's famous remark that,

> It is very much the lazy man's approach in philosophy to posit ambiguities when in trouble. If we face a putative counterexample to our favorite philosophical thesis, it is always open to us to protest that some key term is being used in a special sense, different from its use in the thesis. We may be right, but the ease of the move should counsel a policy of caution: Do not posit an ambiguity unless you are really forced to, unless there are really compelling theoretical or intuitive grounds to suppose that an ambiguity really is present (Kripke (1977), 19).

Cf. Grice's Modified Occam's Razor: "Senses are not to be multiplied beyond necessity" (Grice (1978), 47).

[7]Lewis, in particular, is sensitive to this worry (cf. Lewis (1980), 217). He proposes to evade it by proposing independent motivations for the semantic diversity he claims to find in 'pain'; I suppose one's attitude toward the thesis he adopts will turn largely on one's sympathy toward the proposed motivations.

properties, each of which satisfies only one of them. And that is not what we wanted.

Of course, the proponent of semantic diversity has a ready reply: she will try to convince us that our initial desires were misplaced — that we were merely misled into thinking that one property needed to satisfy both by running together the diverse semantic values taken on by our word 'F'. But, come to think of it, if that explanation were satisfactory, then there would be no need for the semantic diversity theory to begin with. After all, the realizer functionalist can allow for the existence of role properties, and hold that, while the realizers/colors account for the causality requirement, the closely related but numerically distinct role properties account for the commonality requirement. Likewise, a role functionalist can allow for the existence of realizer properties, and hold that, while the roles/colors account for the commonality requirement, the closely related but numerically distinct realizer properties account for the causality requirement. On either view, the claim would be that the constraint not explained by the colors has been wrongly taken to be a constraint on colors themselves only because we confuse the colors with certain closely related but numerically distinct properties. My point is not that these explanations provide convincing ways out of the dilemma by themselves— my view is that they do not. It is only that they are no less convincing than what the semantic diversity theorist offers; indeed, I take these explanations to be preferable to what the semantic diversity theorist offers, insofar as they do not run afoul of standard tests for semantic diversity.

For these reasons, I think we would do best to search for other ways out of our puzzle.

3.2   Colors Are Unreal/*Sui Generis*

Some authors have concluded from the perceived failures of other accounts that colors are unreal projections on the world (cf. Hardin (1988), Boghossian and Velleman (1989), Maund (1995)). Given the complaints against role and realizer functionalisms now on the table, it might seem that I am just providing grist for the irrealist's mill. However (and I think this is something that even proponents of irrealism would allow), irrealism amounts to such a radical revision of our views about the world that we should consider it a position of last resort. In this particular instance, I think there is more work to be done in attempting to rescue some version of color functionalism from our objections before we give up and resign ourselves to irrealism about color.

I have much the same attitude toward the view that colors are *sui generis* or primitive — viz., properties that are real, have genuine instances, and satisfy our desiderata of commonality and causality, but are not analyzable in physical or functional terms (cf. Campbell (1993), McGinn (1996), Watkins (2002), Watkins (2005), and possibly Yablo (1995) and Stroud (2000); see Byrne and Hilbert (2006) for a helpful attempt to classify and evaluate critically various versions of primitivism). I see no contradiction in maintaining this conception of color properties, but find little positive reason (i.e., little reason for endorsing

the view on its own, as opposed to reason for rejecting the alternatives) for believing that colors just happen to satisfy all of our desiderata and yet admit of no informative analysis. Again, it's not that I have some knock-down reason for thinking colors couldn't be like that. But, just as in the case of the ambiguity proposals, primitivism is worrisome because it is too easy to wield as a *deus ex machina* for satisfying arbitrary conjunctions of desiderata (subject only to the constraint of mutual consistency); that it can be so wielded in any case should make us wary about accepting the view too hastily in every case. As it stands, I think it is fair to say that primitivism, like irrealism and the ambiguity views we have considered, is seriously undermotivated.

The undermotivation of the alternatives I have considered in this section will seem especially important if, as I shall argue in §4, there is a more straightforward way out of the dilemma that consists in defending role functionalism against the causal objections we have canvassed.

4   Role Functionalism and Causality

The worry faced by the role functionalist is that it is hard to see how functional roles could have the right causal powers. Namely, the concern is that functional roles can't have the particular sort of causal powers that we want our target properties (e.g., colors) to have because (i) those causal powers are already had by the realizers of the roles (the exclusion worry), or because (ii) the roles are trivially bound up with those particular causal powers in a way that is inappropriate for genuine bearers of the powers (the *virtus dormitiva* worry).[8] I want to show that functional roles can have the causal powers constitutive of the *F*ish role after all, and that the exclusion and *virtus dormitiva* worries are misplaced. I'll make my case for this conclusion in a series of steps, and then consider and reply to objections.

4.1   Step One: Exclusion and Causal Relevance

(The short version of this step would read: Learn the lessons of Yablo (1992), and then move on to step two. The long version that follows is intended to amplify on the short version and show how it applies to the case at hand.)

Suppose a bell is set up to ring any time an object over 1000kg in mass is placed on a certain scale. And suppose that when x, which happens to have a mass of 1037kg, is placed on the scale, the bell rings (as expected). What caused the bell to ring on that occasion? Answer number one: the event of x's exemplifying the property *being over 1000kg* caused the event of the bell's ringing. Answer number two: the event of x's exemplifying the property *being*

---

[8]It is worth noticing that neither objection depends on thinking that functional roles must always be entirely causally impotent; rather the idea is that, even if functional roles are allowed to have causal powers, nonetheless they won't have the particular causal powers that we want to grant to our target properties.

1037kg caused the event of the bell's ringing. Which of our two candidate
answers correctly answers our question?

The best, most pre-theoretically plausible response — the response that our
philosophical account should end up vindicating — is that *both* answers are
correct, which is to say that both of the mass properties mentioned can be cited
in true, felicitous answers to a request for a cause of the bell's ringing. Let's say
that a property is causally relevant to an effect in some explanatory context C
if it can be cited in a true and felicitous answer to the question asked in C of
what causes that effect. On this terminology, what we want to end up saying
is that the two properties mentioned can both be causally relevant in a single
context, despite being distinct.[9]

Unfortunately, our earlier causal objections against role functionalism
threaten to prevent us from saying what we want to say. For, one might in-
sist (hearkening back to an earlier formulation), if *being 1037kg* does the causal
work that leads to the bell's ringing, then there's no residual explanatory work
left to be done by *being over 1000kg*, so the causal efficacy of the former pre-
cludes the causal efficacy of the latter. If so, and assuming (plausibly) that
causally relevant properties are causally efficacious, it would seem that there
should be a similar kind of exclusion at the level of causal relevance — hence
that at least one of our two candidate answers must be rejected.

A useful diagnosis of the situation, due to Yablo (1992), runs along the
following lines. Our two mass properties, *being 1037kg* and *being over 1000kg*,
are closely related; in particular, exemplifying the former is one among many
ways of exemplifying the latter. Indeed, the former is a more specific, hence
determinate, version of the latter and less specific, or determinable. That is to
say, x's exemplifying the more specific/determinate property determines that x
exemplifies the less specific/determinable property; the occurrence of the former
event constitutes the occurrence of the latter. And now the thought is that
determinables and their determinates *do not* compete for causal relevance: since
the having of the one constitutes the having of the other, which of them is cited
in a causal explanation is a matter of choice of compatible descriptions of an
event (how general or specific do we want to be), rather than a choice between
incompatible events.

Indeed, Yablo points out, in many cases the determinable is *better* suited to
being cited in causal explanations than the determinate precisely because of the
greater generality of the former: it captures the thought that the bell would have
rung even in situations where x was not 1037kg. That is, an explanation in terms
of the determinable is often preferable because it adds what Campbell (1993)
calls "modal data" (263) that an explanation in terms of the determinate would
lack: the former says more — and more that is potentially of interest — about
how the effect would have ensued in nearby worlds where the determinate is not
exemplified but the determinable is.

Yablo's suggestion is that, far from competing against each other for causal

---

[9]I'll usually omit the relativization of causal relevance to explanatory contexts unless it
matters.

relevance, the determinates and their determinables work together to achieve a kind of non-competitive efficacy — the causal relevance of the one is not precluded by, but is *constituted by* the causal relevance of the other.

4.2   Step Two: Determination and P-Determination

If this strategy works for *being over 1000kg* and *being 1037kg*, one might hope to extend it to the case of functional roles and realizers. After all, having a certain realizer just is one among many ways of having the role property, just as being 1037kg just is one among many ways of being over 1000kg.

   On the other hand, there is a key disanalogy that threatens to block the extension of the explanation to role/realizer cases. What made the picture plausible for our mass properties was that the having of one of them *metaphysically determines* the having of the other. But, for many role/realizer pairs, the properties are not connected that intimately. For example, it's not metaphysically necessary (*de re*) that being a realizer of the can-opener role comes with exemplifying the can-opener role. Whether instances of this realizer property are instances of this role depend on all sorts of metaphysical contingencies — two that come to mind are the laws of nature and the way that cans are built. And, if anything, the case with the red role property and its realizer is even worse, insofar as the dependence of the latter on the former is mediated by not only the laws, but extremely complicated contingencies about how visual systems work.[10] So maybe the right thing to say is that the Yablo strategy for securing non-competitive causal relevance is applicable to a determinable/determinate pair only when the sense of 'determination' that connects the two is at least as strong as metaphysical necessity. (Plausibly this is part of what we mean by saying that the pair stand in the determinable/determinate relation in the first place.)

   But on the first hand, once again, it's not so clear why full metaphysical determination should be required to apply the Yablo strategy. What makes the strategy look plausible for *being 1037kg* and *being over 1000kg* is that, since having the former just is a way of having the latter, explanations of the effect that cite the former add nothing over explanations of the effect that cite the latter except greater specificity; in particular, they don't add a second event that would then be a competitor for the title of causal relevance.[11] However, if that's all that is going on in the application of the Yablo strategy, it begins

---

[10]Depending on exactly how the role is specified, its exemplification might also depend on the existence of perceivers. However, one can finesse this worry by specifying the role dispositionally; if the role is that of *being disposed* to cause relevant reactions in perceivers of the right sort, then something can exemplify it even in worlds without such perceivers.

[11]This point does not depend on taking sides about whether events should be individuated by a coarse- or a fine-grained standard. If you prefer a coarse-grained individuation (see, for example, the essays in Davidson (1980)), you can express the point by saying that the event cited in the first explanation is identical to the event cited in the second explanation. Alternatively, on a more fine-grained conception of events (cf. Kim (1976)), the claim would be that the event cited in the first explanation (i) overlaps spatiotemporally with the event cited in the second explanation, and (ii) (partially) constitutes the occurrence of the event cited in the second explanation.

to seem plausible once again that we could extend that strategy to role/realizer cases.

Consider again, the can-opener role and one of its particular realizers. As noted, it's not metaphysically necessary (*de re*) that being a realizer of the can-opener role comes with exemplifying the can-opener role, insofar as the realizer only plays the role given contingencies about the laws of nature and about cans. But this much seems true: *holding fixed the actual truths about laws and cans*, anything that exemplifies the realizer will exemplify the role. Moreover, this is so not because exemplifying the realizer will cause some other event of exemplifying the role. Rather, it is so because, again fixing the actual truths about laws and cans, exemplifying the realizer will just *be* one among many ways of exemplifying the role. Therefore, given the actual truths about laws and cans, explanations of a particular can-opening event *e* that cite the can-opener realizer add nothing over explanations of *e* that cite the can-opener role except greater specificity; in particular, they don't add a second event that would then be a competitor for the title of causal relevance.

It looks as if the most significant difference between the case of the mass properties and the case of the can-opener is that the first pair of properties is guaranteed to hang together given only what is metaphysically necessary, where the second pair of properties is guaranteed to hang together given what is metaphysically necessary plus what is (contingently) true about laws and cans. In other words, to get the can-opener role/realizer pair to hang together you need to assume much more. But in explanatory settings where you can make these further assumptions, you are guaranteed that the role/realizer pair will indeed hang together.[12] (Needless to say, this allows that there will be some explanatory settings where the additional needed assumptions are not in place; and in these settings, explanatory bets are off.)

To keep the distinction between the two cases in view, let's reserve 'metaphysical determination' for the relation between the mass properties, and let's describe the weaker determination-given-laws-and-further-metaphysically-contingent-facts relationship between our role/realizer pair as a kind *partial determination*, or *p-determination* for short. The idea is that the can-opener realizer property doesn't determine the can-opener role property all by itself; on the contrary, it takes the realizer property plus some further (metaphysically contingent) constraints to determine the role property. To capture this, we'll say that one property p-determines another relative to certain contingent constraints if *x*'s exemplification of the first, together with the fact that those contingent constraints are met, make it the case that *x* exemplifies the second.

And now, the important point for our purposes is that what's needed to extend the Yablo strategy to the role/realizer case is much weaker than (full, metaphysical) determination: it is only that the target pair should hang together in the sense that *x*'s having one makes it the case that *x* has the other (without creating further events that would then be competitors for causal rele-

---

[12]Importantly, one's guarantee that the mass properties will hang together is dependent on assumptions too — viz., assumptions about metaphysical necessities. It is just that, compared against the assumptions needed for the can-opener case, these assumptions are relatively weak.

vance). And for this, p-determination will be sufficient given that the additional constraints relative to which the first p-determines the second are presupposed in the explanatory context.

The conclusion we are after, then, is that we can have non-competitive causal relevance for a pair of properties P_1 and P_2 in a context C if P_1 p-determines P_2 relative to constraints that are presupposed in C. Luckily, role/realizer properties seem to fall comfortably under this heading in many explanatory contexts, so this gives us a way of seeing how properties of that sort can be compatibly causally relevant in those contexts. In particular, this gives us a way of seeing how the redness role property and redness realizer property can both be causally relevant. To wit: since the realizer property p-determines the role property relative to the natural laws and facts about visual systems, explanations in contexts presupposing those laws and facts that cite the realizer property won't block explanations that cite the role property. Hence, the causal relevance of the redness realizer property (relative to such contexts) won't be any obstacle to the causal relevance of the redness role property (relative to such contexts).

4.3 Step Three: Causation and Causal Relevance

So far I've argued that role/realizer pairs needn't compete for causal relevance, where causal relevance is understood in terms of admissibility into true and felicitous answers to questions about causes. But there's still some distance between that lesson and the desired conclusion that role properties have the Fish causal powers. That's because it's compatible with taking roles and their realizers to be non-competitively causally relevant (in the sense explained) to think that there is competition over what is the cause (of the Fish effects). It might be, for instance, that all sorts of properties are causally relevant to an effect without being causally efficacious of that effect.[13] If so, even if there is no competition (/exclusion) at the level of causal relevance, it could turn out that there is competition (/exclusion) at the level of causation.

Philosophers have come up with a number of strategies for responding to the charge that exclusion could operate at the level of causation even if it doesn't operate at the level of causal relevance. For example, one strategy is to reject the idea that causation is something more metaphysically basic than, something less rooted to our explanatory practice, than causal relevance (Baker (1993), Burge (1993)); if this is right, then non-competitiveness at the level of causal relevance would entail non-competitiveness at the level of causation. A further response builds on the idea that the exclusion problem seems to threaten the

---

[13]One way of making this thought plausible would be to claim that, whereas causal explanations are (like all other explanations) sensitive to contextual features of the communicative settings in which they occur, causation itself is not. Consequently, one might go on to say, causal relevance will be a much less discerning relation than causation — i.e., that '_ is causally relevant to e' will have satisfiers that are not also satisfiers of '_ causes e'. (Of course, this way of developing the point depends crucially on the controversial idea that causation is unlike causal relevance in not being context-sensitive; for dissension, see Schaffer (2006) and Northcott (2006) (and, on some readings, Lewis (1973)).)

causal efficacy of all sorts of properties that we normally take to be causally efficacious. In particular, it seems to generalize easily to other multiply realizable properties, including those that figure centrally in causal explanations in the special sciences; but if so, then the appropriately Moorean reaction would be to reject the metaphysics of causation responsible for the trouble rather than to conclude that all multiply realizable properties are epiphenomenal (cf. Fodor (1989), 138–141, Van Gulick (1992), 325, Baker (1993), 77, Burge (1993), 102; but see Kim (1997) and Kim (1998), 77–87, 112–120). While I'm sympathetic to these responses, I want to suggest that we can get by in the present context with less — less, anyway, than a theory of the metaphysics of causation that would sustain these responses. Namely, we can reply to the worry with merely the admission (reached in §4.2) that causal *relevance* is non-competitive between roles and realizers.

To see why, consider again the background in which the causality demand arose. The demand arose because we decided that Fness, whatever it turns out to be, should be required to have the Fish causal powers (viz., the powers in terms of which the Fish role is elaborated). And presumably we decided this because, having elaborated the Fish role, we wanted to say that Fs *qua* Fs cause the effects that comprise that role.

However, crucially, we *can* say that role properties *qua* role properties cause the effects that comprise their roles, so long as those role properties are causally relevant to those effects — that, after all, is exactly what causal relevance means. And since, as we have seen in §4.2, conceding the causal relevance of realizer properties is compatible with the causal relevance of role properties, such a concession in no way prevents us from saying about the role properties what we want to say about them. What I'm suggesting, then, is that the causality platitude with which role functionalists have been struggling is agnostic between requiring mere causal relevance and requiring causation. Thus, whether or not there is an important distinction between the two, the admission that role properties can be causally relevant is enough to meet the causality platitude.

Since, as I've said, my point goes through whether or not causation is a more selective relation than causal relevance, I've been deliberately neutral on that issue. But suppose, *arguendo*, that it is; in that case, causal relevance will presumably be a derivative relation constituted somehow in terms of causation. If so, then what I've shown is only that functional roles are causally related to their (Fish) effects in some derivative way. Does this not show that, after all, roles aren't *really* (i.e., non-derivatively) causes? Not unless the intuitions undergirding the causality platitude are that Fs must be non-derivative causes. But there is no reason to think that the intuitions come metaphysically marked in that way. On the contrary, our intuitions (as opposed to metaphysical theories we might hold) are agnostic between derivative causation and non-derivative causation. To say otherwise, it seems to me, is to confer on intuitions about causality much more power for discerning metaphysical distinctions than we have reason to think they deserve.

4.4   Step Four: *Virtus Dormitiva*

In the steps so far I've argued that functional roles escape the exclusion arguments against their causal efficacy. The final step in my series will be to show that the answer to the exclusion argument already outlined provides an answer to the *virtus dormitiva* objection as well.

Recall that the *virtus dormitiva* argument builds on the idea that functional roles are definitionally, hence necessarily, bound up with the Fish effects. And the thought is that, given Hume's thesis about the non-necessity of causal connections, this shows that role properties can't be causally relevant.

Consider, as a canonical example of this type, the sleeping potion's property of having a *virtus dormitiva*, and grant that this property is indeed necessarily connected to the effect of inducing sleep. On the view I am defending, the property is nonetheless causally relevant to that effect in an explanatory context C just in case it can be cited in a true and felicitous answer to a question about the cause of the potion's effect in C. And I contend that there are such contexts. For example, consider a context in which there is an alternative suggested explanation of the potion's effect according to which magicians cast sleeping spells on all and only those who have recently ingested the sleeping potion (which is itself otherwise not in any way sleep-inducing); in that sort of context, citing the *virtus dormitiva* of the potion excludes a live rival explanation, and thereby tells us something we might genuinely want to know. In this context, the sleeping potion's *virtus dormitiva* is indeed causally relevant.

Of course, once again, accepting the causal relevance of this property relative to the context at hand leaves open that other properties might be causally relevant as well relative to the same context; moreover, given that this explanatory context is remote from our own epistemic situation, we are liable to find this property particularly infelicitous as an answer to questions about the cause in our own context (relative to other possible answers). These facts may explain why many have thought that the property can't be causally relevant. But, given the current understanding of causal relevance, we can now see that that thought is mistaken (cf. the example of Brutus's property of being Caesar's killer in §4.5 for more on this theme).

What, then, of Hume's thesis? The motivation offered earlier for Hume's thesis was the claim that its denial threatens to drain the indisputably interesting and difficult empirical project of looking for causes of its interest: for, if it sufficed to causally explain the capacity to induce sleep by citing a *virtus dormitiva*, then it is hard to see why anyone would care as much as we manifestly do care about causal explanations. There is the feeling that explanations citing necessary causes would always fail to give us new information, and so would always pale in comparison to the interest of explanations in terms of non-necessary causes.

However, we are now in a position to see that this feeling is misplaced. For we have seen that according causal relevance/efficacy to a necessary cause (the role) is no obstacle to the causal relevance/efficacy of a non-necessary cause (its realizer). In particular, our earlier considerations show how the non-necessary

cause and the necessary cause compatibly enjoy causal efficacy because the former p-determines the latter relative to constraints presupposed in the explanatory context. In fact, allowing necessary causes that are role properties in particular positively *invites* us to pursue the empirical project of looking for other (non-necessary) causes: for citing a role property as a cause leads naturally to the question of what, as a matter of contingent fact, realizes the relevant role. The realizer property that is cited in answering this question will, on the view proposed, itself enjoy causal relevance efficacy by p-determining the role property relative to the constraints presupposed. Hence, contrary to the motivation offered above, denying Hume's thesis by allowing necessary causes need not drain the interest of the empirical project of looking for (other) causes. Thus, there is no longer any reason to insist on Hume's thesis, and consequently no longer any reason to take seriously the *virtus dormitiva* argument against the causal efficacy of roles.

4.5   Objections and Replies

Objection: The proposal on offer purports to show that distinct properties (roles and realizers) can be cited in causal explanations of one and the same effect. In particular, it argues from facts about our causal explanatory practices to the conclusion that, when the distinct properties in question stand in the relation of p-determination, they can both be causes. On the other hand, it doesn't provide anything like a theory of the causal relation itself — it doesn't tell us anything about *how* roles or realizers (or anything else, for that matter) manage to cause anything. But this misconstrues what needed to be explained. The pressing metaphysical question is not *whether* role properties can be causally efficacious, but *how* they can be — and what I've said has left this matter entirely untouched (cf. Kim (1997), 288; Kim (1998), 60–63).

Reply: I agree that no theory of causation has been given. Indeed, I have attempted to be as non-committal about the metaphysics of causation as possible. My point has been, rather, to argue that causal arguments against role functionalism (and in particular, against role functionalism about color) are unpersuasive. If you think that the causal efficacy of functional roles was not seriously in doubt all along, then I will not have advanced your understanding by what I have said — you, like me, will have been disposed favorably toward role functionalism from the beginning. On the other hand, and as noted, many theorists have not been so disposed: indeed, many (e.g., Jackson, Shoemaker, McLaughlin) have given up on role functionalism about color because, they say, causal arguments genuinely throw into doubt the causal efficacy of functional roles.[14] This paper is intended to allay the objections that have led those theorists away from role functionalism, and to clear the path to their acceptance of role functionalism once again.

---

[14]If Kim is right in contending that no one has been seriously moved by the exclusion arguments to reject the causal efficacy of higher level properties outright in philosophy of mind, then this marks a distinction between the discussion in philosophy of mind and that in the philosophy of color.

Objection: The view under consideration is too permissive. It allows that all sorts of properties should be causally relevant, and potentially causally efficacious, that are not. For example, this view "makes Brutus's property of being Caesar's killer causally relevant to Caesar's death, and Drano's property of being plumber-recommended causally relevant to the unclogging of my drain" (Yablo (1995), note 27).[15]

Reply: It is true that, on the present view, Brutus's property of being Caesar's killer is causally relevant to Caesar's death in some contexts — namely, those contexts in which that property can be cited in a true and felicitous answer to the question of what caused Caesar's death (e.g., someone might ask the question while still ignorant about whether Caesar's death was brought on by natural causes). I want to suggest that the intuition that it is not causally relevant reflects the thought that, at least in many explanatory settings, we would want to know more than that true and felicitous answer could tell us. Given what we already know about Caesar's death, that answer doesn't advance our understanding as much as, say, an answer that cites Brutus's property of having stabbed Caesar — a property that p-determines Brutus's property of killing Caesar. But if causal relevance between pairs of properties that stand in the relation of p-determination is non-competitive, as I have urged, then our preference (on some occasions) for an explanation citing some other causally relevant property is fully compatible with the claim that Brutus's property of being Caesar's killer is also causally relevant.

Objection: If one takes a p-determinable and its p-determinate both to be (non-competitively) causally relevant, it seems as if one would be committed to saying that the property formed by conjoining the two (call it 'the conjunctive property') is as well. But that seems deeply unsatisfactory:

> ...there are a host of perplexing questions that such a conjunctive explanation raises: for instance, how precisely could both the instantiation of [the p-determinable] and the instantiation of [the p-determinate] be responsible for [a single effect]? Is the instantiation of one property more responsible than another? If so, how? What is the relation between these two properties? Do we have reason to reject either [the p-determinable] or [the p-determinate] and therewith [the conjunctive property] (Lackey (2002), 388)?

The thought expressed here, I take it, is that the difficulty of these questions derails the causal relevance of the conjunctive property. For if, as many have held, explanatory goodness is assessed in terms of a resulting gain in unity or coherence of our system of beliefs, then the conjunctive property is explanatorily bad rather than good, so not causally relevant after all: accepting it yields a system of beliefs that is less, rather than more unified and coherent than its predecessor.

Reply: Given plausible constraints on explanatory goodness, the proponent of non-competitive causal relevance for p-determinables and p-determinates is

---

[15]Yablo's target here is not the view I have defended but the so-called program model of causal relevance of Jackson and Pettit (1990).

not committed to saying that the conjunctive property will be causally relevant. Recall that a property is causally relevant to an effect in a context in which it can be cited in a true and felicitous explanation of that effect. The truth requirement entails that the conjunctive property is only a candidate for being causally relevant in explanatory contexts in which both its conjuncts (the p-determinable and p-determinate properties) are both exemplified. But it's hard to see how an explanation citing the conjunctive property could be felicitous in any such context, given that the p-determinate and p-determinable properties are available as alternatives. For, compared against those citing the p-determinate property, explanations citing the conjunctive property add on needless verbiage (by conjoining the p-determinable) without bringing any compensatory gain in explanatory generality (/modal data) of the kind mentioned in §4.1 (there is nothing, given what is presupposed, that exemplifies the conjunctive property but fails to exemplify the p-determinate). Whereas, relative to those citing the p-determinable property, explanations citing the conjunctive property add on needless verbiage (by conjoining the p-determinate) and constitute a loss in generality (there may be things that exemplify the p-determinable but don't exemplify the p-determinate).

To be sure, I have not offered a theory of explanation or explanatory felicity. But I take it to be uncontroversial that explanatory felicity (like felicity for any utterance) requires conformity to something like a Gricean maxim of quantity (Grice (1975)). Thus, a quite general and well-supported constraint on rational communication — a constraint that is independent of any particular views about functionalism or causation — explains why a causal explanation will be infelicitous if a more economical and at least equally general alternative is available. This, I suggest, explains why the conjunctive property will not be causally relevant even if both its conjuncts are.[16]

---

[16]The foregoing assumes that we are in an explanatory context in which the usual pragmatic constraints are in place. Of course, this leaves it open that there could be remote explanatory contexts where these constraints are overridden for one reason or another, and therefore relative to which the conjunctive property is causally relevant. But don't Lackey's "perplexing questions" make this implausible? What makes this implausible, I suggest, is the admitted remoteness of the envisaged explanatory context. On the other hand, I think that Lackey's questions can be answered relatively straightforwardly:

- How precisely could both the instantiation of [the p-determinable] and the instantiation of [the p-determinate] be responsible for [a single effect]? See §§4.1–4.3.
- Is the instantiation of one property more responsible than another? No. If so, how? Not applicable.
- What is the relation between these two properties? The p-determination relation.
- Do we have reason to reject either [the p-determinable] or [the p-determinate] and therewith [the conjunctive property]? No.

(Thanks to James Messina for pressing me on this point.)

5 Conclusion

It's not hard to see why philosophers have been attracted to the charms of some sort of functionalist story about color: as in earlier applications, functionalism as applied to color offers its proponents the means to accept empirical findings about the physical heterogeneity of the analysandum without thereby giving in to irrealism or primitivism. However, role functionalism has received short shrift in recent discussions, even by those sympathetic to functionalism about color. It has been rejected in favor of realizer functionalism on the grounds that the latter theory better accommodates demands about the causal efficacy of colors. But this assessment both overestimates the force of those causal objections against role functionalism and underestimates that of commonality objections against realizer functionalism. Once these misconceptions have been corrected, role functionalism about color should seem more attractive than it has so far.[17]

References

Armstrong, D. (1968) *A Materialist Theory of the Mind*, Routledge, London.

Armstrong, D. M. (1970) "The Nature of Mind," in C. V. Borst, ed., *The Mind/Brain Identity Theory*, MacMillan, London. Reprinted in Block (1980), 191–199.

Bach, K. (1998) "Ambiguity," in E. Craig, ed., *Routledge Encyclopedia of Philosophy*, Routledge, London. http://www.rep.routledge.com/article/U001.

Baker, L. R. (1993) "Metaphysics and Mental Causation," in Heil and Mele (1993), 75–95.

Block, N., ed. (1980) *Readings in the Philosophy of Psychology, Volume 1*, Harvard University Press, Cambridge, Massachusetts.

Block, N., and J. A. Fodor (1972) "What Psychological States Are Not," *The Philosophical Review* 81:2, 159–182. Reprinted in Block (1980), 237–250.

Boghossian, P. A., and J. D. Velleman (1989) "Colour as a Secondary Quality," *Mind* 98, 81–103. Reprinted in Byrne and Hilbert (1997a), 81–103.

Burge, T. (1993) "Mind-Body Causation and Explanatory Practice," in Heil and Mele (1993), 97–120.

Byrne, A., and D. R. Hilbert, eds. (1997a) *Readings on Color, Volume 1: The Philosophy of Color*, MIT Press, Cambridge, Massachusetts.

---

[17]This paper began life as an email exchange between Brian McLaughlin and me over the relative merits of role and realizer functionalisms about color; I am deeply indebted to McLaughlin for (among many other things) helping me to get clearer about these issues, although of course he disagrees with my conclusions. I've also benefited from discussion of these matters with Chris Barker, Nancy Cartwright, James Messina, Ed Minar, Dana Nelkin, Adam Pautz, Agustín Rayo, Sam Rickless, Rob Rupert, Michael Watkins, and Stephen Yablo. Finally, I'm grateful for comments from members of the philosophy department at the University of California, San Diego, where I presented a version of this paper.

Byrne, A., and D. R. Hilbert (1997b) *Readings on Color, Volume 2: The Science of Color*, MIT Press, Cambridge, Massachusetts.

Byrne, A., and D. R. Hilbert (2006) "Color Primitivism," in R. Schumacher, ed., *Perception and Status of Secondary Qualities*, Kluwer, Dordrecht.

Campbell, J. (1993) "A Simple View of Color," in J. Haldane and C. Wright, eds., *Reality, Representation, and Projection*, Oxford University Press, New York. Reprinted in Byrne and Hilbert (1997a), 177–190.

Campbell, K. (1969) "Colours," in R. Brown and C. D. Rollins, eds., *Contemporary Philosophy in Australia*, Allen & Unwin, London.

Cohen, J. (2003) "Color: A Functionalist Proposal," *Philosophical Studies* 112:3, 1–42.

Davidson, D. (1980) *Essays on Action and Events*, Clarendon Press, Oxford.

Fodor, J. A. (1989) "Making Mind Matter More," *Philosophical Topics* 17. Reprinted in Fodor (1990), 137–159.

Fodor, J. A. (1990) *A Theory of Content and Other Essays*, MIT Press, Cambridge, Massachusetts.

Grice, H. P. (1975) "Logic and Conversation," in P. Cole and J. L. Morgan, eds., *Syntax and Semantics*, volume 3, Academic Press, New York, 41–58. Reprinted in Martinich (1990).

Grice, H. P. (1978) "Further Notes on Logic and Conversation," in P. Cole, ed., *Syntax and Semantics: Pragmatics*, volume 7, Academic Press, New York, 113–127. Reprinted in Grice (1989).

Grice, H. P., ed. (1989) *Studies in the Way of Words*, Harvard University Press, Cambridge, Massachusetts.

Hardin, C. L. (1988) *Color for Philosophers: Unweaving the Rainbow*, Hackett, Indianapolis.

Heil, J., and A. Mele, eds. (1993) *Mental Causation*, Oxford University Press, New York.

Jackson, F. (1996) "The Primary Quality View of Color," *Philosophical Perspectives* 10, 199–219.

Jackson, F. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford, New York. Originally given as the 1998 Locke Lectures.

Jackson, F., and R. Pargetter (1987) "An Objectivist's Guide to Subjectivism About Color," *Revue Internationale de Philosophie* 160, 127–141. Reprinted in Byrne and Hilbert (1997a), 67–79.

Jackson, F., and P. Pettit (1990) "Program Explanation: A General Perspective," *Analysis* 50, 107–117.

Johnston, M. (1992) "How to Speak of the Colors," *Philosophical Studies* 68, 221–263. Reprinted in Byrne and Hilbert (1997a), 137–176.

Kim, J. (1976) "Events as Property Exemplifications," in M. Brand and D. Walton, eds., *Action Theory*, Redel, Dordrecht, 159–177.

Kim, J. (1989a) "Mechanism, Purpose, and Explanatory Exclusion," in J. E. Tomberlin, ed., *Philosophical Perspectives 3, Philosophy of Mind and Action Theory*, Ridgeview Publishing, Atascadero, California, 77–108. Reprinted in Kim (1993b), 237–264.

Kim, J. (1989b) "The Myth of Nonreductive Materialism," *Proceedings and Addresses of the American Philosophical Association* 63, 31–47. Reprinted in Kim (1993b), 265–284.

Kim, J. (1993a) "The Nonreductivist's Troubles with Mental Causation," in J. Heil and A. Mele, eds., *Mental Causation*, Oxford University Press, Oxford, 189–210. Reprinted in Kim (1993b), 336–357.

Kim, J. (1993b) *Supervenience and Mind: Selected Philosophical Essays*, Cambridge University Press, New York.

Kim, J. (1997) "Does the Problem of Mental Causation Generalize?" *Proceedings of the Aristotelian Society* 97, 281–297.

Kim, J. (1998) *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, MIT Press, Cambridge, Massachusetts. Originally given as the 1996 Townsend Lectures.

Kripke, S. (1977) "Speaker Reference and Semantic Reference," in P. A. French, T. E. Uehling, and H. K. Wettstein, eds., *Contemporary Perspectives in the Philosophy of Language*, University of Minnesota Press, Minneapolis, 6–27. Reprinted in Martinich (1990).

Lackey, J. (2002) "Explanation and Mental Causation," *The Southern Journal of Philosophy* 40, 375–393.

Lewis, D. (1972) "Psychophysical and Theoretical Identitifications," *Australasian Journal of Philosophy* 50:3, 249–258. Reprinted in Block (1980), 207–215.

Lewis, D. (1973) "Causation," *The Journal of Philosophy* 70, 556–567. Reprinted in Lewis (1986), 159–172.

Lewis, D. (1980) "Mad Pain and Martian Pain," in Block (1980). Reprinted in Lewis (1983), 122–130.

Lewis, D. (1983) *Philosophical Papers, Volume I*, Oxford University Press, New York.

Lewis, D. (1986) *Philosophical Papers, Volume II*, Oxford University Press, New York.

Martinich, A. P. (1990) *The Philosophy of Language, second edition*, Oxford University Press, New York.

Matthen, M. (2005) *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception*, Oxford University Press, Oxford.

Maund, B. (1995) *Colours: Their Nature and Representation*, Cambridge University Press, New York.

McGinn, C. (1983) *The Subjective View: Secondary Qualities and Indexical Thoughts*, Oxford University Press, Oxford.

McGinn, C. (1996) "Another Look at Color," *The Journal of Philosophy* 93:11, 537–553.

McLaughlin, B. (2003a) "Color, Consciousness, and Color Consiousness," in Q. Smith and A. Jokic, eds., *Consciousness: New Philosophical Perspectives*, Oxford University Press, New York, 97–154.

McLaughlin, B. (2003b) "The Place of Color in Nature," in R. Mausfeld and D. Heyer, eds., *Colour Perception: Mind and the Physical World*, Oxford University Press, New York, 475–502.

Nassau, K. (1980) "The Causes of Color," *Scientific American* 243, 124–154. Reprinted in Byrne and Hilbert (1997b), 3–29.

Nassau, K. (1983) *The Physics and Chemistry of Color*, Wiley, New York.

Northcott, R. (2006) "Causation and Contrast Classes," under review.

Peacocke, C. (1984) "Colour Concepts and Colour Experiences," *Synthese* 58:3, 365–81. Reprinted in Rosenthal (1991), 408–16.

Putnam, H. (1967) "The Mental Life of Some Machines," in H.-N. Castaneda, ed., *Intentionality, Minds and Perception*, Wayne State University Press, Detroit. Reprinted in Putnam (1975), 408–428.

Putnam, H. (1975) *Mind, Language, and Reality: Collected Papers, Volume 2*, Cambridge University Press, New York.

Rosenthal, D. (1991) *The Nature of Mind*, Oxford University Press, New York.

Schaffer, J. (2006) "Contrastive Causation," *The Philosophical Review*. In press.

Shoemaker, S. (1990) "Qualities and Qualia: What's in the Mind?" *Philosophy and Phenomenological Research* 50, Supplement, 109–131. Reprinted in Shoemaker (1996), 97–120.

Shoemaker, S., ed. (1996) *The First-Person Perspective and Other Esays*, Cambridge University Press, Cambridge.

Stanley, J. (2003) "Context, Interest Relativity and the Sorites," *Analysis* 63:4, 269–280.

Stroud, B. (2000) *The Quest for Reality: Subjectivism and the Metaphysics of Colour*, Oxford University Press, New York.

Thompson, E. (1995) *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception*, Routledge, New York.

Thompson, E., A. Palacios, and F. Varela (1992) "Ways of Coloring: Comparative Color Vision as a Case Study for Cognitive Science," *Behavioral and Brain Sciences* 15, 1–74.

Van Gulick, R. (1992) "Three Bad Aruments for Intentional Property Epiphenomenalism," *Erkenntnis* 36.

Watkins, M. (2002) *Rediscovering Colors: A Study in Pollyanna Realism*, Dordrecht, Boston.

Watkins, M. (2005) "Seeing Red: The Metaphysics of Colour without the Physics," *Australasian Journal of Philosophy* 83.

Yablo, S. (1992) "Mental Causation," *Philosophical Review* 101, 245–280.

Yablo, S. (1995) "Singling Out Properties," *Philosophical Perspectives* 9, 477–502.

Zwicky, A., and J. Sadock (1975) "Ambiguity tests and how to fail them," in J. Kimball, ed., *Syntax and Semantics, volume 4*, Academic Press, New York.