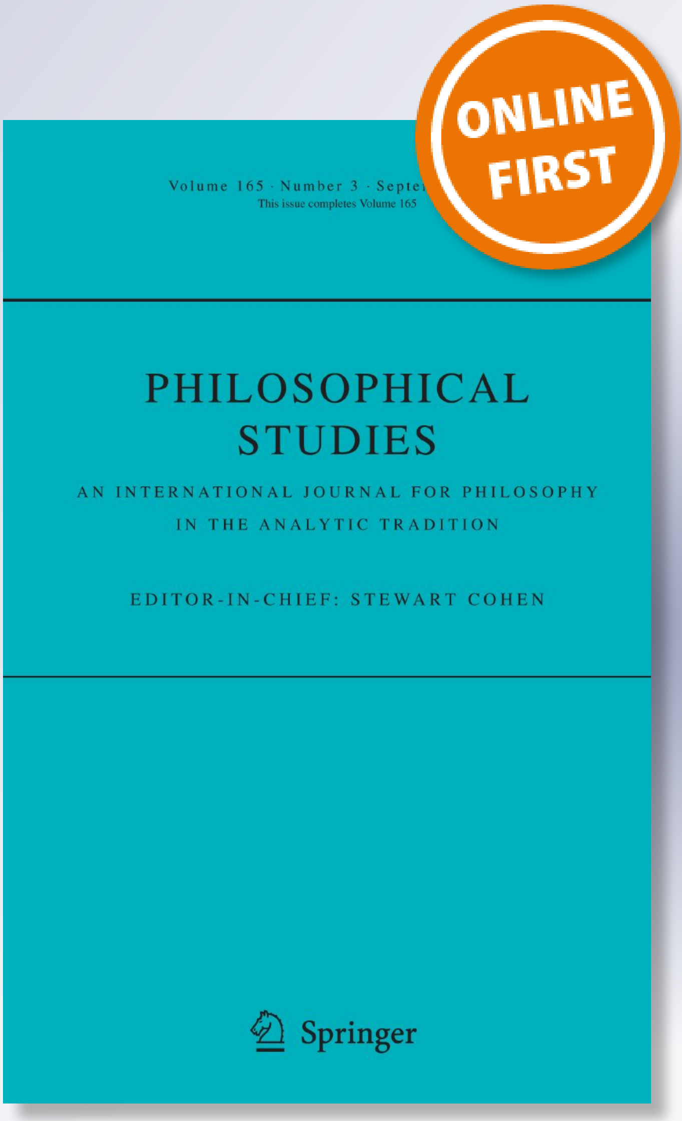
The collapse argument

Joseph Gottlieb

Philosophical Studies
An International Journal for Philosophy
in the Analytic Tradition

ISSN 0031-8116

Philos Stud
DOI 10.1007/s11098-017-1003-5



Springer

Philos Stud
DOI 10.1007/s11098-017-1003-5

The collapse argument

Joseph Gottlieb[1]

Abstract We can divide philosophical theories of consciousness into two main camps: First-Order theories and Higher-Order theories. Like all Higher-Order theories, many First-Order theories are mentalistic theories of consciousness: they attempt to reduce a mental state's being consciousness using mental (but non-phenomenal) terms, such as being available to certain cognitive centers. I argue that mentalistic First-Order theories, once fully cashed out, collapse into some form of Higher-Order theory. I contend that not only are there general considerations in favor of this conclusion, but that the four most prominent mentalistic 'First-Order' theories are, in fact, Higher-Order theories in disguise. Given a strong assumption in favor of some form of mentalism, if this is right, what emerges is a powerful argument for the Higher-Order theory of consciousness.

Keywords Consciousness · The transitivity principle · Higher-Order theories of consciousness · First-Order theories of consciousness · Attention · Availability · Working memory

1 Mentalistic theories of consciousness

Some forms of panpsychism notwithstanding, rocks and nematodes are presumably not phenomenally conscious, or conscious at all. We are. And we want to know why. We know (or at least standardly assume) that we are phenomenally conscious *because* we are in phenomenally conscious mental states, and nematodes (and the like) are not phenomenally conscious because they are not in such states (e.g. Kriegel 2009). But that doesn't help much. So what we want to know—what any theory of phenomenal

✉ Joseph Gottlieb
joseph.gottlieb@gmail.com

[1] Department of Philosophy, Texas Tech University, Box 43092, Lubbock, TX 79409-3092, USA

consciousness ought to tell us—is *what it is* to be in a conscious mental state. If it doesn't do that, the 'theory' on the table is not a theory at all.[1]

Let a *Higher-Order* (HO) theory of consciousness (e.g. Lycan 1996; Rosenthal 2005; Hellie 2007; Kriegel 2009; Gennaro 2012) be any theory of consciousness that accepts the following claim:

Transitivity Conscious mental states are mental states we are aware of in some way.[2]

Let a *First-Order* (FO) of consciousness (e.g. Dretske 1995; Tye 2000; Block 2009) be any theory of consciousness that *rejects* Transitivity. Let a *mentalistic* FO theory (e.g. Dretske 1995; Tye 2000) be any FO theory that attempts to reduce a mental state's being conscious to some mental but non-phenomenal notion, such as availability to belief-desire formation systems.[3] More specifically:

FO Mentalism Conscious mental states are mental states with a certain dispositional property F, where a mental state M's being F consists in M's being available to right sort of cognitive centers

Most FO mentalists (with many HO theorists) will add that conscious states are representational states. For present purposes that doesn't matter. What does matter is just this: the FO theorist's dispositional account of what is for a mental state to be conscious is supposed to *preclude* the need for (occurrent) higher-order awareness in an account of existence conditions for conscious states.

And that is precisely the claim I want to challenge. I'll argue that once fully cashed out, mentalistic FO theories collapse into HO theories. I call this *the Collapse Argument* for HO theories:

[1] Except where it is needed to enhance clarity, going forward I will drop the'phenomenal' qualifier. Hence it should be assumed, unless explicitly noted otherwise, that when I talk about a mental state being conscious, I am concerned with its being phenomenally conscious, where a mental state M of a subject S is phenomenally conscious if and only if there is something it is like for S to be in M. A theory of consciousness in the relevant sense, then, will be an account of the conditions under which there is something it is like for S to be in M.

[2] HO theories can be individuated in terms of how they cash out the 'in some way' Transitivity—e.g. it in terms of a higher-order thought (Rosenthal 2005), higher-order perception (Lycan 1996), or higher-order acquaintance (Hellie 2007). Also, some HO theorists prefer a slightly differently formulation of Transitivity:

Transitivity* Conscious mental states are mental states we are aware of ourselves as being in.

While I'll stick with the more classic Transitivity here, nothing substantive turns on this choice. Some who favor Transitivity* (e.g. Berger 2014) seem to think that consciousness is actually a property of subjects, and not a property literally instantiated by mental states. But on such a view there are still 'conscious states' in another sense: a state M is conscious when there is something that it is like for S to be in M. It is S that instantiates the property of being concious, but the nature of that property amounts to S instantiating the right kind of higher-order mental state (cf. Brown 2015). In what follows I bracket this sort of HO theory.

[3] Mentalistic reduction is not at odds with neurophysiological reduction. It is only at odds with direct neurophysiological reduction of consciousness. For a non-mentalistic FO theory, see Block (2009). The expression "mentalistic reduction" is borrowed from Gennaro (2012: 16).

1. If mentalism is true, either FO mentalism is true or HO theory is true.
2. If FO mentalism is true, conscious mental states are mental states with a certain dispositional property F.
3. If conscious mental states are mental states with a certain dispositional property F, then on any reasonable account of F, Transitivity is true.
4. So, if FO mentalism is true, Transitivity is true. [2–4]
5. Mentalism is true.
6. So, Transitivity is true. [1, 4, 5]
7. If Transitivity is true, some form of HO theory is true.
8. Therefore, some form of HO theory is true. [6–7]

The argument is valid. So the question is whether (1), (2), (3), (5) and (7) are true. Premises (1) and (7) are true by definition. I assume (5). So strictly speaking my thesis is this: if mentalism is true, some HO theory is true. I'm inclined to think mentalism is true, but I won't press that case here.[4]

The plan for what follows is thus to take up (2) and (3). Like (1) and (7), (2) is true by definition. Yet unlike (1) and (7), (2) is only true by stipulative definition. Since all that is built into the FO theoretic picture is a claim about what conscious states are not, it is possible for a FO theorist to reject to the consequent of (2). But as it happens the most well-developed forms of FO mentalism—Fred Dretske's, Michael Tye's, and Jesse Prinz's—embrace the consequent of (2). So while (2) is true by stipulation, it's consequent tracks what FO mentalists actually say. Still, it will be useful to examine *how* FO mentalists arrive at the consequent of (2). Doing so will give us two things. First, it will yield indirect motivation for (2). Second, it will yield some terminological distinctions that will be critical in our defense of (3)—the 'collapse premise'. Thus I'll look at (2) next in Sect. 2, and then proceed to (3) in Sect. 3. Section 4 concludes.

Two final notes before we proceed. First, a HO theory is simply an implementation of Transitivity. To implement Transitivity is to articulate *what sort* of higher-order awareness is constitutive of a mental state's being conscious. That awareness could be conscious (Kriegel 2009) or it could be unconscious (Rosenthal 2005). It is also important to realize that, while the most well-known versions of HO theory construe this awareness representationally, there is nothing about implementing Transitivity that requires this. HO awareness may very well be a species of representation, but it need not be. Benj Hellie (2007) has developed a HO theory where Transitivity is implemented via HO *acquaintance*. This is a non-intentional relation.[5] And crucially, even this much is denied by the FO theorist. Their rejection of Transitivity is not premised on the assumption that HO awareness is construed representationally. They reject Transitivity *full stop* (e.g. Dretske 1993; Tye 2009: 5). So to defend (3) I do not need to show that on any

[4] There is good reason to think that a successful reductive theory of consciousness will be mentalistic. As Carruthers (2005: 6) points out, to attempt to reduce consciousness directly in physical terms is to "jump over too many explanatory levels at once." See also footnote 3.

[5] David Chalmers (2013: 348, fn. 4) also favors a view like this, with the proviso that it appeals specifically to a Russellian instance-acquaintance that is irreducibly phenomenal.

reasonable account of F conscious states are represented states. I only need to show that on any reasonable account of F conscious states are states we are aware of in some way.

Second, central to what follows is the role of attention in FO forms of mentalism—in particular, in the FO theories of Tye and Prinz. Tye is not terribly explicit about the role attention plays in his theory, but Prinz is. Yet in both cases, it is the appeal to attention that ultimately gives way to the collapse.

2 Awareness, consciousness, and availability to cognition

After advancing an extensive argument against HO theory, FO theorist Fred Dretske offers the following positive proposal:

> An experience of x is conscious, not because one is aware of the experience, or aware that one is having it, but because, being a certain sort of representation, it makes one aware of the properties (of x) and objects (x itself) of which it is a (sensory) representation…Experiences and beliefs are conscious, not because you are conscious of them, but because, so to speak, you are conscious with them. (1993: 280–281)

To make the contrast with Transitivity stark, let's substitute 'conscious state' for 'conscious experience'.[6] The proposal is then:

Aware With Conscious mental states are mental states we are aware with

Let S be a subject and M be a mental state of S. So Aware With says: if S is in M, and M is conscious, then S will be aware of something in virtue of being M. What 'something'? Not M—or at least not necessarily M. Rather, the idea is that S will be aware of some x, where 'x' picks out something that is, or at least appears to be, instantiated in S's ambient environment.[7] Here's the point: if Aware With is a viable alternative FO theoretic proposal, (2) is false.

Now technically, as a threat to (2), Aware With is idle. Indeed, as we will see shortly, there is a sense in which Dretske himself rejects Aware With, and nor do I know of any other FO theorist who accepts it in the intended sense. But it will nonetheless be instructive to see why it is rejected, since doing so will bring us to the consequent of (2).

So here is an argument against Aware With, due originally to Rosenthal (2008: 837):

2a. If conscious mental states are mental states we are aware with, then every mental state is conscious.

[6] I do not mean to imply that this substitution is harmless, but it is one that FO and HO theorist make alike. I won't try to render this substitution harmless here, but if we think of mental states as properties, and experiences as events, then something like Jaegwon Kim's (1991) property-exemplification theory of events might be helpful. Cf. Bourget (2010).

[7] 'Environment' is to be construed broadly so as to include S's body.

2b. But not every mental state is conscious.
2c. Thus, conscious mental states are not mental states we are aware with.

Both sides agree that there are unconscious mental states, so (2b) is not up for grabs.[8] This leaves (2a). Is it true that all mental states, even unconscious ones, are states subjects are aware with? If the FO theorist can answer in the negative, she might be able to accept the antecedent of (2a), but still insist that not every mental state is conscious, since not every mental state is a state we are aware with. Only the conscious ones are.

To make this concrete, let's consider two examples. First: *masked priming*. A standard trial consists of three events in sequential order: a forward mask (usually a set of # symbols) presented for 500 ms, a prime presented briefly (about 50 ms), followed by the target to which a response is required—a lexical decision, for instance. Despite the brief period presentation, a prime related to the target facilitates a response when compared to an unrelated control.

Second: *blindsight*. The cause of blindsight is neural damage in the primary visual cortex. It results in large blind spots (or 'scotoma') in the visual field. Subjects with blindsight claim that they cannot see anything in their blind field. Yet when forced to choose they can accurately report on stimuli presented in the blind field with an impressive degree of reliability. While information about stimuli in the blind field is represented and processed in the visual system—evidently enabling the reliable accuracy of these reports—the information doesn't surface in conscious experience.

Rosenthal's (2a)–(2c) argument seems to rest on the following assumptions: (i) that (things like) masked priming experiments and blindsight is evidence of unconscious perceptual mental states, and that (ii) subjects, in virtue of being in these perceptual states, are aware of the stimuli, albeit unconsciously. This suggests a distinction between *conscious awareness* and *unconscious awareness*. Blindsight subjects are aware of the stimuli in their scotoma. And upon reflection, now (2a) doesn't seem unreasonable; after all, given the effects on their subsequent cognitive behavior, they certainly *act* like they are aware. Still, there is nothing it is like for such subjects to be aware of these stimuli, because the mental state in virtue of which they are aware of the stimuli is itself unconscious. So they are aware of the stimuli, but only unconsciously. They have unconscious awareness, but they don't have *phenomenal* awareness.[9] What's missing, or so the HO theorist will say, is the subject's being aware *of her awareness*. Only then is there something it is like for

[8] Since Freud, it is a common assumption that beliefs can occur unconsciously. And although there are some doubters (e.g. Phillips 2016), masked priming and blindsight are frequently taken to provide robust evidence of unconscious or subliminal perception. See, e.g. Kouider and Dehaene (2007) and Weiskrantz (2009). More on the perceptual cases momentarily.

[9] Rosenthal (1997: 743) gives us a more mundane example of unconscious awareness with the 'cocktail party effect,' which "occurs when one screens out the sounds of conversation other than one's own. Still, if one's name is mentioned in a conversation one had screened out, one's attention immediately shifts to that conversation, showing that one must have had an auditory consciousness of what was being said." Rosenthal takes this to show that a subject S can be conscious of x without the state in virtue of which the S is conscious of x itself being conscious, and that unconscious awareness is rather ubiquitous.

the subject to be aware of the stimulus, and so only then is the subject's state of awareness conscious in the relevant sense. The upshot is that articulating existence conditions for conscious mental states becomes articulating the conditions under which we have conscious, or phenomenal, awareness.

What can the FO theorist say? First off, it is worth noting that at least some FO theorists already seem to concede Rosenthal's reasoning, even if only implicitly. Dretske (2006) himself has an (aptly named) paper entirely devoted to finding the conditions under which perceiving x yields awareness of x. On the face of it this undercuts my point. But a closer look suggests otherwise. For example, in a section titled 'Conscious Awareness,' Dretske (*ibid*: 153) talks about finding an operational definition for 'conscious awareness.' And then later (*ibid*: 171), when giving his final considered view of the conditions under which a perceptual state is conscious, speaks of 'S being consciously aware of x.'[10]

This aside, I see two moves available for the FO theorist. We already hinted at the first: concede Rosenthal's (i), but deny his (ii), deny, that is, that unconscious perceptual states are states we are aware with. The second stops the train at (i). Perception is necessarily conscious. Blindsight subjects (for instance) might receive 'unconscious perceptual information' but they do not perceive with respect to their scotoma.

First, a basic point: it is not (or need not be) baked into the HO theoretic apparatus that unconscious perceptual states make their subjects aware of things. The most basic question is just what makes a given mental state M conscious; there is no stipulation that M, if a perceptual state, must be a state of awareness, or that M, if perceptual, is not necessarily conscious. All that is required is that mental states are not necessarily conscious.

The real issue, however, is this. Distinguish two readings of Transitivity. On an *extensional* reading, all conscious states are mental states we are aware of in some way. The HO theorist, of course, says that the extensional reading of Transitivity is true; the FO theorist says that it is false. However, the HO theorist's intended reading of Transitivity is actually a *constitutive* one; it is a claim about what M's being conscious consists in.[11]

Now take Aware With. Here too there is a both an extensional reading and a constitutive reading. The extensional reading is fine; the HO theorist will agree (or at least can agree) that Aware With is extensionally adequate. Yet the FO theorist cannot intend a constitutive reading. Two points draw this out. First, the FO theorist is free to reject the distinction between conscious and unconscious awareness and say that whenever S is aware of x there will be something it is like for S to be aware of x. So she will say that *whenever* S is aware of x, S will be in some mental state M in virtue of which she is aware of x, and M will be conscious. The FO theorist is also free to say that perception is necessarily conscious, and so that *whenever* S

[10] Rosenthal concurs with this reading: "A perception is conscious, then, because it makes one aware of the stimulus. But, Dretske recognizes that this will not do as it stands; subliminally perceiving something makes one aware of it, just not consciously aware of it" (2012: 1425).

[11] I recognize that the term 'constitutive' is imperfect here. But I trust that it is useful at least insofar as it provides a clear contrast with 'extensional'.

perceives x, there is something it is like for S to be aware of x, and so there will be some mental (perceptual) state M in virtue of there is something it is like for S to be aware of x. Either way, our bedrock question remains: what is it for S to be in a conscious mental state? The only difference is that rejecting the first distinction between conscious and unconscious awareness makes it such that no unconscious states are states of awareness, and rejecting the second distinction makes it such that no unconscious state is a perceptual state.

Second: recall that it is an assumption, shared by HO and FO theorists alike, that it is M's *being conscious* that is supposed to explain why there is something it is like for S to be aware of x, not the other way around.

Putting these two points together, the problem is clear. For when read constitutively, Aware With gets things backwards; there being something it is like for S to be aware of x is now what explains why M is conscious. That's false, and that it is false is (at least) one point on which both the HO and FO theorist will agree. And read extensionally, Aware With tells us what we already knew; we already knew that conscious mental states are those states we are (phenomenally) aware with. What we want to know is what conscious mental states *are* such that this is true.

The FO theorist is thus in a bit of a jam. Transitivity was out from the get-go. And although Aware With is extensionally adequate, it is inadequate when read constitutively. It is for this reason that I think FO theorists have gravitated toward the consequent of (2). Take blindsight again. If forced to choose, such subjects can more often than not correctly 'say' whether x is vertical, blue, or moving (thereby exhibiting 'perception' or 'awareness' of x). But as Dretske (2006) notes, what seems to be missing is an ability to exploit the incoming sensory information so as to initiate planning and fixate genuine belief. This suggest that consciousness is somehow bound up with cognitive processing, such that what makes a mental state conscious is its (or the information it carries) *availability* to the right sort of cognitive centers, viz. those that play a role in fixing belief and rational planning and object-selection (Dretske, *ibid*).[12] What explains why there is something it is like for S to be aware of x is that S is in some conscious state M. But what explains why M is conscious is that it has this specific dispositional profile. And as it turns out, this is precisely what extant FO theorists claim.

Such is all I will say about (2). The plan for the remainder is to argue for (3). That is, I'll argue that the FO theorist's attempt to understand a mental state's being conscious in terms of its being apt to produce non-phenomenal states like beliefs and desires will inevitability require HO machinery. This is the collapse.

[12] The appeal to rationality and reasons should not be understood so as to require deep conceptual sophistication. As we will see shortly, in Dretske's case what's relevant is reasons *why* not reasons *for*. Also, the appeal to rational processes is not always so obvert amongst FO theorists. This is the case on Prinz's (2012) 'AIR' theory, also to be discussed below. But Prinz does appeal to availability to working memory, which is the 'storehouse' for information used in such controlled cognitive processes. And that's all that matters for our purposes.

3 The collapse

We are now ready to defend (3). Recall:

(3) If conscious mental states are mental states with a certain dispositional property F, then on any reasonable account of F, TRANSITIVITY is true.

We are granting for sake of argument the antecedent of (3). The goal is to defend the consequent, i.e. to show how on any reasonable explication of what it is for M to be F, M's being F entails that S is aware of M.

My defense of (3) will have two prongs. First, I will offer a general argument for (3). The point of this argument is not to give a conclusive case for (3)—one that will convince the FO theorist, say—but just to grease the wheels for seeing why one might find (3) intuitive to begin with. Second, I will look at several of the most prominent (purportedly) FO mentalist theories: Baars' (1988) Global Workspace Theory (in its psychological form); Dretske's (unnamed) FO theory; Tye's PANIC theory; and Prinz's AIR Theory, with a focus on the latter three. I will argue that on each of these theories' account of F the consequent of (3) follows. Hence all of these theories collapse into some form of HO theory. Taken together both moves provide a robust case for (3).[13]

On the general case for (3): recall that F is a dispositional property. This (I would suggest) means that F is in need of some categorical base.[14] Suppose M's being F consists in M's being apt to produce beliefs and desires, or it's being available for free use in reasoning. We might then wonder: what is the categorical property in virtue of M is so apt or so available? Well, say M is my state of seeing the traffic (or my experience of the traffic). And let's add that while being in M, I am deciding whether to cross the street. I am not keen on getting hit by a car, so I 'use' M to reason whether it is safe to proceed to the other side. In virtue of what is M available for such use? The following strikes me as *prima facie* plausible: it is in virtue of my being aware of M that I can use M to decide to cross the street.[15] Absent another candidate categorical base—and it is important to note that most FO theorists do not provide one—it seems reasonable to take higher-order awareness (i.e. awareness of M) as the categorical basis of M's being F.[16]

[13] Recall that the FO theorist rejects both the extensional and constitutive readings of TRANSITIVITY. So even if any reasonable account of M's being F only entails TRANSITIVITY in the extensional sense, this will still be bad news for the FO theorist, since she will still inevitably need to appeal to HO machinery in her account of conscious states. But I will be discussing (3) as it pertains to the constitutive reading of TRANSITIVITY.

[14] The FO theorist might deny that dispositions require categorical bases, but I presume that's a road she will not want to pin the success of her theory to. In any event, I'll assume that all dispositions have categorical bases, or at least (*pace* Shoemaker 1979) bottom out into some categorical base.

[15] We sometimes use 'available' is an *objective* sense, as when my prescription is available for pick-up from the pharmacy even though I do not know this, and so it's being available for me doesn't entail that I am aware that it is. But we also use 'available' in a *subjective* sense, and it's hard to see how that cannot be what's intended in the present case; if I didn't know that I saw the traffic, in what sense would my seeing the traffic be available for me to act appropriately? If I do not know that the prescription is ready, I likely won't do anything.

[16] Notice too that even if FO theorists have some neural property N in mind as the categorical basis of F, it still might be the case higher-order awareness is reducible or at least correlated with N. See Kriegel

That's the (very) general case. Onward now to the particular theories. In (3) I mention "reasonable account[s]" of what it is for M to be F. What counts as reasonable is a tricky question, and I don't think it will be fruitful to get too precise on the matter. So I have opted to let "reasonable account[s]" just pick out those accounts that are well-developed and have gained (at least some) traction in the literature. This leaves open the possibility that some new FO mentalism will come around that avoids the collapse. Nevertheless, even if only all extant explications of M's being F entail that S is in some way aware of F, we will have still moved the ball considerably in favor of HO theories.

3.1 The global workspace theory

According to the Bernard Baars' Global Workspace Theory (GWT), a mental state M is conscious when and only when M is in the 'global workspace'—a functionally defined component of working memory, viz. one that makes M available to a wide-range of cognitive sub-systems (Baars 1988, 2002).[17]

Technically, GWT is only a potential form of FO mentalism. I say 'potential' because it's far from clear that Baar's intended target was phenomenal consciousness, our interest here.[18] But since it might be possible to incorporate something like GWT in a FO-theoretic approach to phenomenal consciousness, it's at least worth discussing if only to explain why doing so won't, in the end, be viable.

GWT says that conscious mental states are those that are *in* the global workspace—that is, not merely accessible to, but actually *encoded* in the workspace. This is crucial, since it is only by actually being in the workspace that the representations are in *active* contact with the relevant consuming mechanisms, and thus accessible to those mechanisms without the need for further-processing (cf. Block 2007). One such mechanism is the mechanism that underlies reporting. So a corollary of GWT is this: if a mental state M is conscious, M must be spontaneously reportable, be it verbally or by some other method (Block 2007: 491; Baars and Franklin 2007: 956).

But herein lies the problem. Reportability plausibly requires awareness; if I can report, at a time t, that my great uncle played in Frank Sinatra's band, then I am aware, at t, that my great uncle played in Frank Sinatra's band. If I wasn't aware of this I couldn't report it. Saying otherwise is unintelligible. By extension, if a mental state M of a subject S is conscious only if M is spontaneously reportable, then S must be aware of M. The reports of conscious states reflect our HO awareness of

Footnote 16 continued
(2009: 39) for this point. Kriegel also advances an argument similar to what I am calling the 'general argument', but in the context of trying to vindicate certain practices in the scientific study of consciousness.

[17] The theory has since been cashed out in neurobiological and computational terms (e.g. Dehaene et al 2006). However, so as to stick with our discussion of mentalistic theories, I am only interested in GWT in its psychological form.

[18] Block (2007) construes GWT as a theory of cognitive *access*, not as a theory according to which cognitive *accessibility* is constitutive of phenomenal consciousness.

those states. When I report that I saw a certain stimulus, I must have been aware of seeing that stimulus at the time I saw the stimulus.[19] And being aware of seeing the stimulus is just being aware of a (first-order) mental state. Consequently, it follows that if GWT is true, so is TRANSITIVITY.[20]

It might be objected that any FO mentalist who wanted to make use of GWT would push for a weaker corollary, something like: only conscious states of reportage-*capable* subjects are spontaneously reportable.

I'd imagine this weaker corollary might be preferable since many think phenomenal consciousness is present in subjects who are plausibly not capable of report—some non-humans and pre-linguistic infants, say. In any event, I don't see this as a problem. As Kriegel (2009: 306) points out, even if we only grant this weaker claim, we can still run the above argument. Since the relation between reportage-capability and consciousness is accidental, we could take the entailment of a *restricted* form of TRANSITIVITY by the weaker corollary as evidence that reportage-*incapable* subjects are also aware of their conscious states—i.e. as evidence that TRANSITIVITY is true *simplicter*. In any event, we have our first collapse.

3.2 Dretske's mentalism: justifying reasons

Next up is Fred Dretske's FO mentalism. According to Dretske (2006), perception is conscious only if the subject can cite the information provided by the perceptual state as a justifying reason for doing something. So we get:

DRETSKE'S MENTALISM For any mental state M of a subject, M is conscious only if S can cite the information provided my M as a justifying reason for φ'ing

However, as Rosenthal (2012: 1425) has noted, one cannot cite something one is unaware of. So, we might wonder: can we cite the information provided by the perceptual state as a justifying reason for doing something without being in some way aware of that perceptual state? A *yes* answer strikes me as implausible.

To see the implausibility, take note of what a justifying reason is. Using Dretske's (2006: 168) own example, when Jane grabs the umbrella, and we ask her why, she can give as a justifying reason for her action that it is raining. This qualifies as justifying reason even if it is not in fact raining; all that need be the case is that Jane *takes* it to be the case that it is raining—that she represents this as being so. Now suppose Jane, walking through a crowded rainy street, suddenly makes a short hop across a patch of the pavement. When asked why, she says, pointing to the ground: 'Well, I saw a big puddle there, and I didn't want to get my shoes wet.' By Dretske's lights, having given such a response, we can conclude that Jane's

[19] This argument, or something very much like it, has been pushed for years by Rosenthal. For a nice summary, from which the current presentation borrows, see Kriegel (2009: 305–306).

[20] I am not claiming that HO theories necessarily require that conscious states are reportable. I am only claiming that if a mental state must be reportable to be conscious, some form of HO theory follows. For more, see Sect. 3.4.

perceiving the puddle was conscious. She was consciously aware of the puddle. But can we conceive of her having said this, and then (truthfully) *denying* that she was aware of her perceiving the puddle, or aware of being aware of the puddle? I doubt it, and Dretske does not give us any reason to think otherwise.

We can press the point. Dretske denies (*ibid*: 171) that Jane (or any other subject capable of being in conscious states) need have the explicit belief that she perceived the puddle in order for her perceiving the puddle to be conscious, and he denies that a subject's having a reason for action must consist in that subject being in a state that employs some concept (Jane need not have the concept PUDDLE). But he does insist (*ibid*: 168–9) that Jane's justifying reasons be represented somehow in thought or experience. Given that, we might compare two scenarios:

Scenario One Jane's having a reason consists (at least in part) in her being in a state with the content: [puddle]

Scenario Two Jane's having a reason consists (at least in part) in her being in a state with the content: [saw the puddle]

Now suppose again that Jane makes a short hop across a patch of the pavement. Would the reason described in Scenario One, if ascribed to Jane, count as a *justifying* reason for her hopping across the puddle? Or would only the reason described in Scenario Two count? It's hard to see how the former would cut it. That I represent a puddle wouldn't make me jump, unless I believed (or otherwise represented) that I saw the puddle. After all, the puddle itself is irrelevant to Jane's having a reason to jump absent the fact that she sees it.

If it is true that only the reasons ascribed to Jane in Scenario Two counts as justifying reasons, then Jane representing's (the fact) that she saw the puddle is a constitutive feature of her consciously perceiving the puddle. And notice what follows. According to Dretske, justifying reasons are (putative) facts of which we are necessarily conscious (*ibid*: 171). Moreover, and by Dretske's (*ibid*) lights yet again, awareness of facts about x entails awareness of x. Here x is just Jane's state of seeing the puddle. Her state of state of seeing the puddle just is a perceptual mental state. Thus, so long as it is true that only the reasons ascribed to Jane in Scenario Two count as justifying reasons, Jane must be aware of her perceptual state for her perceptual state to be conscious. Since there is nothing special about Jane, it follows that TRANSITIVITY is true.

So, by Dretske's own lights the existence conditions for conscious states requires at least some HO machinery. Consequently, Dretske's FO theory is, in fact, a HO theory in disguise. Hence we have our second collapse.

3.3 Tye's mentalism: PANIC

Tye's case is trickier. Tye does not appeal to justifying reasons. Instead, Tye's explication of M's being F is this:

TYE'S MENTALISM For any mental state M, M is conscious only if M is poised, where M's being poised consists in its "being ready and available to make a direct impact" on beliefs and desires—i.e.

> beliefs and desires about the perceptible qualities represented by M (2000: 62).[21]

I have framed the property of being poised as a property of M, but it's not obvious that this is what Tye intends. He often talks of 'poised content'— indeed this is built into the acronym 'PANIC'. But he also talks of poised representational states (e.g. *ibid*: 62, 63, 1995: 172), and of experiences, not experience's contents, bearing the right functional roles (e.g. *ibid*: 62). That strikes me as better. As Uriah Kriegel (2002: 60) has previously noted, it is arguably a category mistake to attribute this sort of F to a content *c*: it is the mental state M which carries *c* that makes a direct impact on our beliefs and desires. We can of course speak loosely of a given content *c*'s being poised if by this we mean that *c* is carried by a state that is poised. I'll vacillate somewhat between these two ways of speaking, but it should be understood that going forward, only the latter is true in a strict sense.[22]

Tye does not spell out the "ready and available" clause in any meaningful detail. But—following Daniel Stoljar (forthcoming)—it seems plausible to think that description actually undersells the matter. Surely Tye wants something stronger than M's merely being available to cognitive processes that mediate between perception and behavior.[23] Something can be available and accessible and still be hard to actually access. That's not so with conscious states, however. When something (phenomenally) looks red to me, it is not as if I have to make some concerted, willful effort to form the belief that something looks red to me; I just do form the belief, and I do so with no effort at all. And if I am looking for something red, I just grab what looks red.

Thus to be poised is, as Block puts it, to "be on the verge of attacking" (1995: 245). For a mental state to be poised is for it to be on the verge of impacting our beliefs and desires. How are we to understand this stronger requirement? A natural suggestion is made by Stoljar: appeal to attention. Stoljar draws our attention to several features of attention (!), but only three are germane here. Most important is that attention is a secondary state. By this Stoljar means that S can attend to x only if S already bears some psychological relation to x that is not attention (*ibid*). In particular, S can attend to x only if S is aware of x.[24] In this sense, we might say that attention is a secondary state as its occurrence entails the existence of some primary psychological state—viz. a state of awareness.

Second, Stoljar (*ibid*) notes that attention enhances. Suppose I see an apple on my desk. But suppose I also attend to it. Now I am more likely to grab for it, toss it from

[21] Tye, like Dretske, is a representationalist, so (minimally) M will be a representational state, and if M and M* differ in phenomenal character, M and M* will differ in terms of representational content. While Tye's views on the purely representationalist component of his theory has changed over the years (2009, 2014), his commitment to the property of being poised being the 'difference maker' between conscious and unconscious states has remained the same.

[22] It is worth noting that Block, from whom Tye largely borrows this notion, also treats being poised as a property of M (1995: 231). So does Stoljar (forthcoming).

[23] Stoljar is primarily concerned with Block's account of access-consciousness, and his appeal to poise, but Stoljar's discussion carries over cleanly to Tye.

[24] Though it need not be conscious awareness. See below.

hand to hand, or take a bite then if I just saw it; in other words, if, in virtue of seeing the apple I am disposed to such-and-such, then I am normally more likely to do such-and-such if I also attend to it. In this sense, attention has an enhancing effect.

Finally, though my attending to the apple is something I can freely choose to do, it need not be. Attention is not always under voluntary control.[25]

Putting these points together, what emerges is a more complete picture of the conditions under which a mental state is poised. In fact, we can distinguish two pictures, only the first of which appears in Stoljar's discussion:

Poise# For any subject S and any mental state M of S, M is poised if and only if S attends to the content of M to a sufficient degree.

Poise@ For any subject S and any mental state M of S, M is poised if and only if S attends to M to a sufficient degree.[26]

If a mental state's being poised really is a matter of its being ready and available to engage our cognitive systems—as Tye says it is—and not simply something that must be dredged up with great effort, then it should not be surprising that attention plays such a role.

In fact, Tye would be friendly to the general suggestion that being poised is connected to attention, as it tallies well with his prior commitments, viz. his commitment to the view that being able to form *de re* thoughts about x (e.g. 'What's that?') is necessary to have phenomenal awareness of x. Tye (2010) thinks that the ability to have *de re* thoughts about x is contingent on the ability to attend to x; so, for Tye, S's being phenomenally aware of x entails that S can attend to x.[27]

Yet it also clear that Tye would not accept Poise@. For as noted, attention is a secondary state a pre-condition of which is plausibly (some form of) awareness. So,

[25] This feature of attention is important because of cases like Block's (1995) super-blindsighter. The super-blindsighter is supposed to be a case of access-consciousness without phenomenal consciousness. But as Tye (1995: 143) points out, because the super-blindsighter needs to prompt himself to guess, his mental states are not actually poised for direct control of thought and action. A mental state cannot be poised for direct control of thought and action by an act of willing—or, in the present case, a deliberate act of attention. I won't expound on the claim that attention is not always a matter of voluntary control here, but I don't need to. As we will see, Tye thinks attention is necessary for (phenomenal) consciousness. If that's right, it follows that attention is not always a matter of voluntary control, since it is not up to me whether I am conscious of x, for any x.

[26] By 'attend to the content of M' I follow Stoljar: to attend to the content of M is just to attend to 'what is seen', 'what is heard' etc. This translates into our earlier expression 'aware of x' where x is just what is seen, heard, etc. So if M is a visual state (or a state of seeming to see), then for S to attend to the content of M is to attend to whatever x S sees (or seems to see) in virtue of being in M. Also, I have slightly altered the first definition from Stoljar's original presentation to mesh better with the present discussion. Aside from some trivial schematic changes, the only difference is that Stoljar is actually defining access-consciousness in terms of poise. However, my treatment of poise (Poise#) is the same as his.

[27] Critically, Tye (2010) rejects the converse entailment; he does not think that phenomenal awareness is necessary for attention. More on this momentarily.

if S can attend to M only if S is aware of M, then Poise@ entails Transitivity. And if that's right, Tye's 'FO theory' is not really a FO theory after all. It's a HO theory.[28]

So Tye will certainly want to go in for Poise# instead. But can he? I'll argue now that he cannot. And since Poise# and Poise@ are almost certainly jointly exhaustive—what else might S attend to such that M is poised?—if I am right, Tye's PANIC theory collapses.

Revisit our earlier question: either there are perceptual states that make us aware of our environment that are unconscious—i.e. there are states of unconscious awareness—or there are not. Feeding both options into Poise#, we get:

Poise#$_{CA}$ For any subject S and any mental state M of S, M is poised in virtue of S's attending to the content of M to a sufficient degree, and S's attending to the content of M metaphysically necessitates that S is consciously aware of the content of M

Poise@$_{UA}$ For any subject S and any mental state M of S, M is poised in virtue of S's attending to the content of M to a sufficient degree, and S's attending to the content of M metaphysically necessitates (only) that S is unconsciously aware of the content of M.

On Poise#$_{CA}$, S's attending to the content of M metaphysically necessitates that there is something it is like for S to be aware of the content of M, i.e. that S is consciously aware of the content of M. (Hence the 'CA' subscript.) On Poise#$_{UA}$, S's attending to the content of M metaphysical necessitates only that S is unconsciously aware of the content of M, not that there is something it is like for S to be aware of the content of M. (Hence the 'UA' subscript.)

Now take Poise#$_{CA}$. This won't do. For recall: it supposed to be M's being conscious that explains why there is something it is like for S to be aware of the content of M. Yet if the FO theorist employs Poise#$_{CA}$ in an account of what it is for a mental state to be conscious, she'll get things the other way around: there being something it is like for S to be aware of the content of M will now explain why M is conscious.[29] This is so because on Poise#$_{CA}$ M is poised in virtue of S attending to the content of M, and S attends to the content of M only if S is consciously aware of the content of M. Consequently, if M is conscious in virtue of being poised, M will be conscious in virtue of S being consciously aware of the content of M—i.e. there being something it is like for S to be aware of M.

Tye will not have *this* problem if he appeals to Poise#$_{UA}$. For Poise#$_{UA}$ says that S can be aware of the content of M (or aware of x) without there being anything it is like for S be aware of the content of M (or aware of x). He will have at least two other problems, however. First off: Poise#$_{UA}$ is counterintuive. Suppose again I see an apple on my desk. Even if we allow that it is possible for me to see, and so be

[28] Though see Prinz (2012), to be discussed next in Sect. 3.4.

[29] It is worth noting that Block's appeal to poise does not suffer from this problem. For Block (1995), a mental state's being poised (or access-conscious) and it's being phenomenally conscious doubly dissociate. So Block would not appeal to Poise#$_{CA}$ (or Poise#$_{UA}$) in an account of what it is for a mental state to be phenomenally conscious.

aware, of my apple without my seeing being conscious, it makes little sense to then say that I could still attend to the apple. On a standard Jamesian view, attention, at least in the ordinary case, does not seem like the sort of psychological phenomenon that could occur unconsciously.[30]

More importantly though, Tye himself denies that M's being Poise#$_{UA}$ would make M conscious!

> [B]lindsight subjects can attend to things in the blind hemifield without being visually conscious of those things. They cite as evidence the fact that their blindsight subject was more likely to guess correctly whether a circle had been presented in a given location L in the blind field if the location was cued beforehand by bars just above and below L (bars that were themselves entirely within the blind field). A plausible view is that the bars attracted the blindsight subject's attention to the spatial region occupied immediately afterwards by the circle and thereby to the circle. On this view, he did not see the circle but he did attend to it. So, it is possible to attend to something in the visual case and yet not see it (2010: 424).

Tye (*ibid*: 413) is explicit that his use of 'see' here is the phenomenal sense. So what he is saying is that blindsight subjects can attend to the content of their mental states even though they are not—in present terminology—consciously aware of that content. Fine—let's grant that. Now Tye's full view is that conscious mental states are those mental states with abstract, non-conceptual content, that are also poised. But he also thinks that blindsight subjects have no phenomenal consciousness with respect to their scotoma, and not because they their states lack abstract, non-conceptual content, but because they are not suitably poised (2000: 63). So if Tye *did* say that M's being Poise#$_{UA}$ is sufficient for M's being conscious, he would have to conclude that blindsight subjects are phenomenally conscious with respect to their scotoma.

Thus on pain of inconsistency Tye cannot ascribe to Poise#$_{UA}$. Nor, as we saw, can he ascribe to Poise#$_{CA}$. Since these options jointly exhaust Poise#, Tye cannot ascribe to Poise#. This leaves Poise@. But Poise@ entails Transitivity. And as Poise# and Poise@ are arguably the most plausible ways to understand what it is for a mental state to be poised, it follows that Tye's theory entails Transitivity. As such, we have our third collapse.

3.4 Prinz's mentalism: the AIR theory

Finally, we have Jesse Prinz's (2012) AIR theory. Prinz's explication of M's being F is this:

[30] I don't mean to imply that this claim is entirely uncontroversial—Tye (2010), for instance, denies it. See also Kentridge 2011. But my sense is that it better tracks the default view; see, e.g. Mole (2008). See also our discussion of Prinz (2012) next in Sect. 3.4.

Prinz's Mentalism For any mental state M, M is conscious if and only if (i) M is an intermediate-level representation and (ii) M is available for encoding in working memory (2012: 97)

Clause (i) is irrelevant for our purposes. Clause (ii) does the work, since not all intermediate-level representations—vantage-point specific, bounded representations—are conscious. However, unlike GWT, (ii) does not require actual encoding. It only requires availability for encoding. Yet availability to working memory, like *being poised*, is a dispositional notion and thus in need of a categorical base. Prinz supplies one: attention. Thus, we have the 'AIR' theory: conscious states are attended intermediate-level representations. So unlike Tye, Prinz makes an explicit appeal to attention.

As the name suggests, on AIR theory M's being conscious consists in M's being attended to; AIR theory does not say that M's being conscious is a matter of attending to what M represents. This point comes out most clearly when Prinz is discussing why attention is missing in ordinary cases of subliminal perception:

> We might attend to a region of space before a stimulus is presented there, but even in this case, the stimulus presentation must cause a stimulus representation to be formed before that representation can itself be an object of attention. Or, put more accurately, we attend to an external stimulus only by attentionally modulating a representation of that stimulus, not by merely attending to the location in which it is presented. (*ibid:* 88)

This is in line with one of our abiding themes: that *we* are phenomenally conscious *in virtue of* being in phenomenally conscious states. On the AIR theory, a subject S is phenomenally conscious of a stimulus only if S attends to that stimulus. But what makes this the case is that S attends to the mental state (M) that represents that stimulus.

This point is crucial. In our discussion of Tye's PANIC theory is what claimed that, standardly conceived, attention is a secondary state: necessarily, if S attends to x then S bears some psychological relation *R* to x, where *R* is not itself attention. It was also remarked that a plausible candidate for *R* is awareness, albeit not necessarily conscious awareness. So the question becomes this: does Prinz's theory give us a reason to think that attending to M is not a way of being aware of M? I think the answer is *no*—or least that that's the most plausible way of reading Prinz. If that's right, the AIR theory collapses into a form of HO theory.[31]

Now Prinz denies that attention must be defined in terms of "phenomenal qualities" (*ibid*: 90). Again, this doesn't matter. For HO theorists like Rosenthal are happy to implement Transitivity in terms of unconscious awareness; HOTs are only themselves conscious when they are introspected—that is when they are the object of a second, distinct HOT. The comparison with HOT theory is particularly

[31] The argument to follow has been inspired by a discussion between Pete Mandik and Richard Brown, found on the latter's online blog. See https://onemorebrown.com/2007/06/21/is-there-such-a-thing-as-a-nurophilosophical-theory-of-consciousness/ and https://onemorebrown.com/2010/03/10/hot-damn/. Interestingly, according to Brown, Prinz is actually amenable to this argument, suggesting that he might see himself as a HO theorist, at least in some sense.

instructive. Here, roughly, is how that picture unfolds. In virtue of being in a perceptual state M, I am aware of what M represents. Call it x. Unless M is itself conscious, my awareness of x is unconscious—I am not consciously aware of x. For M to be (phenomenally) conscious, I need to aware of M. That's Transitivity. On HOT theory, this is a matter of representing myself as being in M via a suitable higher-order thought M*. When that occurs, I am consciously aware of x: I am, in other words, phenomenally conscious of x. But M* is not conscious unless I introspect. Hence, I am only unconsciously aware of M.

I think Prinz's AIR theory replicates this basic story. According to Prinz, attention is the process by which perceptual representations become available to working memory (*ibid*: 97). Prinz (of course) denies that this is a matter of *representing* those representations—there is no HO *content* involved (*ibid*: 90). So that's a departure from standard HO theories. But, once more, a HO theory is simply an implementation of Transitivity. Appeal to representation is strictly speaking dispensable. All we need is awareness. This leaves us with two options:

AIR# Conscious states are attended states; when we attend to a mental state, we are made conscious of what those states represent, but we remain *wholly unaware* of the state itself

AIR@ Conscious states are attended states; when we attend to a mental state, we are made aware (though not consciously aware) of that state, and thereby conscious of what that state represents

AIR@ is, in essence, exactly what happens on HOT theory, minus the obvert appeal to assertoric thoughts, instead adding attention as the relevant form of awareness. So Prinz must intend AIR#. The problem is that it's not clear what entitles Prinz to AIR#. Prinz contends that attention is not just necessary for consciousness, but sufficient (*ibid*: 113). This means that while attending is enough to be conscious, the thing we attend to (an intermediate-level representation) so *that* we are conscious in not something we are in any way conscious or aware of. That's peculiar. That peculiarity is compounded by the fact that when M is conscious, we are attentively conscious of what M represents, but only in virtue of attending to M first (*ibid*: 88). But all of this is done with M itself wholly in the dark all the while.

AIR@, by contrast, avoids this peculiarity, but at the cost of collapsing into a HO theory. This might not ultimately be what Prinz had wanted, but the resulting version of HO is novel, and does no obvious violence to anything else Prinz says. Three points draw this out, all of which are interrelated.

First: reportability. Recall that on GWT, all conscious states are reportable. That's (predictably) not so on AIR theory. Instead, conscious states are only *poised* to be reportable (*ibid*: 28). Prinz cites the well-known Sperling (1960) experiments as evidence for this claim. Subjects are presented with a 3×4 array of letters, flashed briefly on a screen. Typically, subjects could only report on average 4.3 letters (*ibid*: 1). But when using a tone after the array ceased to be visually present to indicate which line of the array subjects were to report, however, subjects did far better, suggesting—at least to Prinz—that all conscious states *can become* reportable with the right cue. They are just not reportable merely by being conscious.

I don't know if Prinz's interpretation is right. But there is nothing here that tells against HO theory, for not all HO theorists claim that conscious states are reportable. Gennaro's (2012) 'Wide Intrinsically View' (WIV) is an illustrative example. On the WIV, things like inferential role, report, and language-use are associated with *conscious* HO awareness (*ibid*: 250–251). But AIR@ does not posit conscious HO awareness. It posits unconscious HO awareness. This suggests a pair of categorical bases: one for reportability (conscious, perhaps conceptually sophisticated, HO awareness) and one for being poised to be reportable (unconscious HO awareness).

Second: overflow. Prinz agrees with Block (2007) that phenomenology 'overflows' working memory, although he insists that it does not overflow accessibility to working memory. This too can be accommodated by a HO theory. As Richard Brown (2014) points out, HO theorists don't need to claim that we are aware of our conscious states in *all their respects*. I might be aware that I am seeing letters (a fairly generic property) but not the identity of every letter I am seeing. In other words, I can be unaware of some of the information represented by the first-order state without the state itself being unconscious (*ibid*). What happens, then, is: I am phenomenally conscious of the entire 3×4 array, with representations of the identities of all the letters available prior to cuing. But only a small number (usually around four) ever get through, accessed by working memory. That's overflow, and perfectly consistent with HO theory.

Third: neural realization. For reasons connected to the first and second points, Prinz contends that the neural level realization of conscious states should be 'non-central' and not too 'high' (*ibid*: 32). In practice, this rules out the pre-frontal cortex (PFC).[32] Now it is true that HO theories sometimes implicate the PFC as essential to consciousness (e.g. Lau and Rosenthal 2011), but this too is contested amongst HO theorists. Once again Gennaro's WIV is instructive. According to Gennaro (2012: 278), only very sophisticated conscious forms of metacognition such as introspection require PFC activity. On the WIV, the PFC is inessential. More likely to be essential is posterior areas, such as the inferior parietal cortex (*ibid*: 281). And that's exactly what Prinz says (2012: 98). What's true is that any HO theory will require *some* level of global neural distribution for (first-order) conscious states. But that doesn't mean there is no leeway with how much, depending on what type of HO theory we are talking about. The AIR theory just takes a middle-of-the-road approach.

Therefore, I conclude that the best way to read Prinz's AIR theory is as HO theory. Hence we have our forth collapse.

[32] Although not part of the frontal lobe, the anterior cingulate cortex (ACC) is occasionally lumped with the PFC. By 'PFC' I am excluding the ACC; my point is that (i) while some HO theorists say that the PFC *proper* is essential to consciousness, (ii) not all HO theorists say this.

4 Conclusion

I have argued that all major extant versions of FO mentalism collapse into HO theory. The core observation running through these arguments has been that all attempts to cash out a mental state's being conscious in terms of some dispositional property (e.g. being citable as a justifying reason, being poised, being available to working memory) inevitably requires an appeal to HO awareness as that disposition's categorical base. This might be taken as evidence that *any* plausible FO mentalism will collapse into HO theory. And assuming that we have—as I think we do—a strong presumptive reason in favor of mentalism *simpliciter*, what emerges is a powerful case for HO theories of consciousness.

Acknowledgements Many thanks to an anonymous referee, Jessica Gottlieb, Daniel Moss, Saja Parvizian, and David Rosenthal for helpful comments and discussion.

References

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge, MA: Cambridge University Press.

Baars, B. J., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global workspace theory and IDA. *Neural Networks, 20,* 955–961.

Berger, J. (2014). Consciousness is not a property of states: A reply to Wilberg. *Philosophical Psychology, 27*(6), 829–842.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18*(2), 227–247.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*(5), 481–548.

Block, N. (2009). Comparing the major theories of consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences IV* (pp. 1111–1123).

Bourget, D. (2010). *The representational theory of consciousness*, Dissertation. Australian National University.

Brown, R. (2014). Consciousness doesn't overflow cognition. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.01399.

Brown, R. (2015). The HOROR theory of consciousness. *Philosophical Studies, 172*(7), 1783–1794.

Carruthers, P. (2005). *Consciousness: Essays from a higher-order perspective*. Oxford: Oxford University Press.

Chalmers, D. (2013). The contents of consciousness: Reply to Hellie, Peacocke and Siegel. *Analysis, 73*(2), 345–368.

Dehaene, S., Changeux, J. P., Naccache, L., Sackura, J., & Sergenta, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Science, 10*(5), 204–211.

Dretske, F. (1993). Conscious experience. *Mind, 102*(406), 263–283.

Dretske, F. (1995). *Naturalizing the mind*. Cambridge: MIT Press.

Dretske, F. (2006). Perception without awareness. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 147–180). Oxford: Oxford University Press.

Gennaro, R. (2012). *The consciousness paradox: Consciousness, concepts, and higher-order thoughts*. Cambridge: MIT Press.

Hellie, B. (2007). Higher-order intentionalism and higher-order acquaintance. *Philosophical Studies, 134*(3), 289–324.

Kentridge, R. (2011). Attention without awareness: A brief review. In C. Mole, D. Smithies, & W. Wu (Eds.), *Attention: Philosophical and psychological essays* (pp. 228–246). New York: Oxford University Press.

Kim, J. (1991). Events: Their metaphysics and semantics. *Philosophy and Phenomenological Research, 51*(3), 641–646.

Kouider, S., & Dehaene, S. (2007). Levels of processing during nonconscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B, 362,* 857–875.

Kriegel, U. (2002). PANIC theory and the prospects for a representational theory of phenomenal consciousness. *Philosophical Psychology, 15*(1), 55–64.

Kriegel, U. (2009). *Subjective consciousness*. Oxford: Oxford University Press.

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15*(8), 365–373.

Lycan, W. (1996). *Conscious experience*. Oxford: Oxford University Press.

Mole, C. (2008). Attention in the absence of consciousness? *Trends in Cognitive Science, 12*(2), 44.

Phillips, I. (2016). Consciousness and criterion: On block's case for unconscious seeing. *Philosophy and Phenomenological Research, 93*(2), 419–451.

Prinz, J. (2012). *The conscious brain*. Oxford: Oxford University Press.

Rosenthal, D. (1997). A theory of consciousness. In Ned Block, Owen J. Flanagan, & Guven Guzeldere (Eds.), *The nature of consciousness*. Cambridge, MA: MIT Press.

Rosenthal, D. (2005). *Consciousness and mind*. Oxford: Clarendon Press.

Rosenthal, D. (2008). Consciousness and its Function. *Neuropsychologia, 46,* 829–840.

Rosenthal, D. (2012). Higher-order awareness, misrepresentation, and function. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1594), 1424–1438.

Shoemaker, S. (1979). Identity, properties, and causality. *Midwest Studies in Philosophy, 4*(1), 321–342.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74,* 1–29.

Stoljar, D. (forthcoming). In praise of poise. In A. Pautz & D. Stoljar (Eds.), *Themes from block*. Cambridge, MA: MIT Press.

Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. Cambridge, MA: MIT Press.

Tye, M. (2000). *Consciousness, color, and content*. Cambridge, MA: MIT Press.

Tye, M. (2009). *Consciousness revisited: Materialism without phenomenal concepts*. Cambridge, MA: MIT Press.

Tye, M. (2010). Attention, seeing, and change blindness. *Philosophical Issues, 20*(1), 410–437.

Tye, M. (2014). Transparency, qualia realism and representationalism. *Philosophical Studies, 170*(1), 39–57.

Weiskrantz, L. (2009). *Blindsight: A case study spanning 35 years and new developments* (2nd ed.). Oxford: Oxford University Press.