Analytic Philosophy Vol. 57 No. 4 December 2016 pp. 353–379

TRANSITIVITY AND TRANSPARENCY

JOSEPH GOTTLIEB
Texas Tech University

1. Two Purported Truisms about Consciousness

Here's an allegedly intuitive idea:

TRANSITIVITY: Conscious mental states are mental states we are aware of in some way.

Higher-Order theories of consciousness entail TRANSITIVITY. We can individuate Higher-Order theories by the type of higher-order awareness they posit—that is, by how they implement TRANSITIVITY's 'in some way' clause. For example, on the Higher-Order Thought (HOT) brand of Higher-Order theory, we are aware of a mental state when it is the object of an assertoric higher-order thought (e.g., Rosenthal 2005). On the Higher-Order Perception (HOP) brand of Higher-Order theory, we are aware of a mental state when it is the object of a perception-like internal monitor (e.g., Lycan 1996). On the Higher-Order Acquaintance (HOA) brand of Higher-Order theory, we are aware of a mental state when it is the object of a non-intentional form of awareness, viz. acquaintance (e.g., Hellie 2007).[1]

TRANSITIVITY occupies a curious place in the philosophical literature on consciousness. On the one hand, it is a dividing point between two major theoretical families: Higher-Order theorists, who accept it, and First-Order theorists, who reject it. First-order theorists are a motley crew. One can be a naïve realist (e.g., Campbell 2002), a qualia realist (e.g., Block 1996) or a representationalist (e.g., Dretske 1995; Tye 1995, 2000, 2009) and be a first-order theorist. First-order theories are thus best defined negatively: a theory T is a first-order theory of consciousness if, according to T, higher-order awareness of a mental state is not necessary for that mental state's being conscious. Conscious mental states may be mental states we are conscious *with*, but they are not necessarily states we are conscious of (Dretske 1993: 281).

On the other hand, many Higher-Order theorists believe that TRANSITIVITY is *obviously* true. William Lycan has been quoted as wondering whether TRANSITIVITY is "nearly trivially true" (Gennaro 2011: 29).

[1] Despite positing a constitutive relationship between the higher-order awareness and the lower-order mental state that is the object of the higher-order awareness, self-representationalism (e.g., Kriegel 2009) is still a Higher-Order theory on my taxonomy. Self-representationalism can, in principle, accommodate a HOT, HOP or HOA construal of higher-order awareness.

Analytic Philosophy Vol. 57 No. 4

David Rosenthal describes TRANSITIVITY as a "pre-theoretic datum" (1997a: 156) and as "intuitively obvious" (2000: 273). Similarly, Pessi Lyrra describes TRANSITIVITY as a "fundamentatum" (2009: 68).

I think the best way to summarize these sentiments is like this: for the Higher-Order theorist, TRANSITIVITY is a *truism*. A truism is something that is not only deemed true, but also deemed to be obviously true. Truisms have methodological import. Truisms can be assumed at the beginning of inquiry. We theorize *from* truisms, not *to* them. The upshot here is significant. Higher-Order theories are implementations of TRANSITIVITY. So to say that TRANSITIVITY is a truism is really just to say that some form of Higher-Order theory is obviously true.

Here's another allegedly intuitive idea:

> TRANSPARENCY: Our conscious states, or experiences, are in some way *transparent* to awareness.[2]

TRANSPARENCY, or something near enough like it, enjoys widespread support. It is adopted by most First-Order theorists, and many Higher-Order theorists.[3] I am hesitant, however, to say that TRANSPARENCY is a truism, at least as it stands. I have two reasons. First, while many enthusiastically endorse the claim that experience is transparent, it's an open question whether the enthused are actually enthused about the same phenomenon (Schroer 2007). Second, even assuming the enthused have a bead on a single genuine phenomenon, the scope of this phenomenon is another source of widespread debate. Very few reject TRANSPARENCY outright; in most cases, it is the *universality* of TRANSPARENCY that is rejected.

It would be too hasty to conclude that there is no truism in its vicinity, however. We can get there in two steps. The first is a disjunctive step. The basic idea behind TRANSPARENCY is that there is

[2] It is perhaps more common to talk of the transparency of *experience*, rather than the transparency of conscious states. But using 'conscious experience' and 'conscious state' interchangeably is largely harmless, and is common throughout the literature on Higher-Order theory. (For substitutions outside the Higher-Order theory literature, see Wright 2005). If we think of mental states as properties, and experiences as events, then—following something like, say, Rosenthal's usage—an experience is an event that consists in the instantiation of a perceptual mental state with a distinctive qualitative character (Rosenthal 2015). This way of putting things would preserve the Higher-Order theorist's distinction between conscious experiences and unconscious experiences (e.g., Carruthers 2000; Rosenthal 2005), and align conscious states with the former; an experience is conscious only if one is aware of being in the mental state the instantiation of which that experience consists in. Moreover, it would also be in line with what proponents of TRANSPARENCY tell us when they say that experience is transparent, viz. that there is a sense in which our awareness is exhausted by what these (first-order) mental states are *of*, not the states themselves. [We need not take a stance on what events are generally speaking, but something like the Jaegwon Kim's (1991) property-exemplifcation theory seems appropriate here (cf. Bourget 2010)].

[3] A representative but incomplete list of sympathizers includes: Moore 1922; Harman 1990; Dretske 1995; Martin 2002; Tye 2002; Speaks 2014; Schroer 2007; Gennaro 2003, 2008; Kriegel 2009.

an awareness-relation that we seemingly cannot bear to our experiences. Yet the nature of this awareness-relation is a matter of controversy. By treating this awareness-relation as a disjunction of awareness-relations, we can mitigate that controversy. The second is a restrictive step. By restricting the class of things that are transparent to the least controversial cases, we mitigate the controversy surrounding the scope of transparency. The resultant two-step thesis—let's call it *Weak Transparency*, or W-TRANSPARENCY for short—should be a crowd-pleaser. It retains the widely accepted notion that there is *some* sense in which at least *some* experiences are transparent, and so will speak to everyone save for those who think that there is *no* sense in which *any* experience is transparent—a small enclave. I conclude, then, that W-TRANSPARENCY is a truism. Or if it isn't, it's near enough for our sake.[4]

I have two aims in this paper, one primary and one secondary. My primary aim is to investigate a puzzle. The principle goal of a theory of consciousness is to provide *existence conditions* for conscious mental states. That is, a theory of consciousness is to explain what it is in virtue of which a mental state is conscious as opposed to unconscious. And TRANSITIVITY, so it would seem, gives us a highly compelling place to start: conscious states are those mental states we are *aware of in some way*. But we also have another purported truism according to which we are *in some way unaware* of at least some conscious states. The puzzle, then, is to see whether there is a precisification of TRANSITIVITY's 'in some way' clause that is (i) compatible with the various ways of precisifying W-TRANSPARENCY's 'in some way' clause, yet (ii) continues to motivate Higher-Order theory. I'll argue that there is no such precisification.

This result is important. For one thing, even if First-Order theorists reject it, TRANSITIVITY is all the same undeniably extremely intuitive; it is certainly odd to speak of a conscious state of which we are wholly unaware. So it would be nice, given their intuitiveness, if we could consistently ascribe to both TRANSITIVITY and W-TRANSPARENCY. If I am right, we cannot, or at least not if our precisification of TRANSITIVITY is robust enough to sustain a Higher-Order theory. Second, and relatedly, several Higher-Order theorists (e.g., Gennaro 2003, 2008; Kriegel 2009) have explicitly argued that TRANSITIVITY and TRANSPARENCY (and by extension, W-TRANSPARENCY) *are* consistent. So if I am right, these Higher-Order theorists are wrong.

My secondary aim concerns whether TRANSITIVITY and W-TRANSPARENCY really ought to be treated as truisms. I'll suggest that if TRANSITIVITY and W-TRANSPARENCY do conflict, then they should not be treated as truisms. This conclusion, however, does not require that TRANSITIVITY and W-TRANSPARENCY be *ultima facie* incompatible.

[4] W-TRANSPARENCY is different from what Amy Kind (2003) calls "weak transparency," which says that it is difficult (as opposed to impossible) to stand in the relevant awareness-relation to our experiences.

All it requires is something weaker—that it is *extremely difficult* to see how TRANSITIVITY and W-TRANSPARENCY could be compatible. If it is extremely difficult to see how one purportedly obvious thesis could be compatible with another purportedly obvious thesis, then it is imprudent to act as if either thesis can be treated as a fundamentatum in our theorizing about consciousness. Given that both theses *are* often treated in this way, this result is important as well, for it implies that there is something wrong with the way we theorize about consciousness.[5]

2. Getting Transparent on Transparency

The literature on transparency is large. It is also a mess. Thus, some regimentation will be useful. Let a *transparency thesis* be a description of the transparency phenomenon. David Chalmers (2013: 350) points out that a transparency thesis can be schematized like so:

> SCHEMA [for a transparency thesis]: "It is (i) difficult/impossible to stand in (ii) a certain relation to (iii) certain mental entities."

SCHEMA has three dimensions of variation. The dimension of variation involved in (i) is between what we'll call prima facie and ultima facie transparency (Kind 2003). The relation in (ii) may be attention, direct attention, awareness, direct awareness, introspection, and so on. Dimension (iii) could involve intrinsic features of experience, phenomenal character, conscious states, experience, the awareness relation and so on (Chalmers ibid). For now, the point to note is this: there has been a transparency thesis for each permutation of SCHEMA. As I said: a mess.

[5] One might wonder whether Higher-Order theorists are united on the status of TRANSITIVITY. For instance, Lycan (2001) calls it "stipulative", though stresses that the stipulation corresponds to a central use of the term 'consciousness'. And some—even Rosenthal—occasionally describe it as an *intuitively attractive* claim. Both of these remarks seem different from calling TRANSITIVITY a truism. Fair enough. But my point is simply that *many* Higher-Order theorists take TRANSITIVITY to be a truism, and that is the sentiment I'm challenging. (It's also worth noting, that both Lycan and Rosenthal, as quoted earlier, do indicate that they think TRANSITIVITY is a truism, despite what they may say elsewhere). In addition, there is *indirect* evidence that Higher-Order theorists take TRANSITIVITY to be a truism—for example, in Rosenthal's discussion of Ned Block's views on visual extinction (2002: 656), and Josh Weisberg's (2011: 439-40) discussion of Block's argument against HOT theory. These discussions suggest that, according to the Higher-Order theorist, Block is not merely wrong to characterize states of which a subject is wholly unaware as conscious, but in some sense *obviously* wrong—to the extent that he must mean something else by 'consciousness' and 'what it is like' altogether. Yet that's just another way of saying that TRANSITIVITY is a truism.

To clean up, I propose two moves. We'll do the *restrictive* move first. This move actually has two components. The first is a restriction of cases. No matter how we fill in SCHEMA, an unrestricted transparency thesis will be controversial [see, e.g., phosphine experiences (Block 1996); blurred vision (Smith 2008); bodily sensations (Kind 2003)]. Fortunately, we do not need an unrestricted transparency thesis, since TRANSITIVITY is itself unrestricted in scope. However, we fill in (iii)—whatever it is that we cannot (or have difficulty) stand(ing) in a certain relation to—it is sufficient for our purposes to restrict the scope of (iii) to the least controversial cases.[6]

The second restriction consists of treating transparency as a purely *negative* phenomenon. Discussions of transparency often invoke the metaphor of a pane of glass (see, e.g., Kind 2003: 226). Characterizing transparency as a purely negative phenomenon amounts to characterizing the nature of the glass (i.e., what role the glass is supposed to play in the metaphor), and characterizing the relation we evidently don't bear to the glass. Characterizing transparency as a *positive* phenomenon amounts to characterizing what is revealed to us beyond the glass. For instance, Harman's famous case of *Eloise* is not merely a negative illustration, for example, a case of *not* being aware of our experiences. It is also a positive illustration; the colors Eloise experiences are experienced as features of the "tree and its surroundings" (1990: 39). Similarly, Tye claims that, when we have a perceptual experience, the only properties we are aware of are properties of external objects—colors and shapes of surfaces, for instance. (2014: 40; cf. Speaks 2014). These claims are based on a reasonable inference from the way things *seem*. If it seems like what we are aware of are just features of the "tree and its surroundings," then—absent extremely compelling reasons to think otherwise—that *is* what we are aware of. Still, by dropping the positive claim, our understanding of transparency becomes less restrictive. When I have a visual experience of a red apple, the redness that constitutes what it is like for me may be instantiated by an amalgam of sense data. Or, it may just be a represented property of a physical apple. It is unlikely that a Higher-Order theorist will be happy with the former, but it is nice to know that there is nothing presumed here that precludes her doing so.

The second move concerns dimension (ii) of SCHEMA, but let's detour briefly to (i) and (iii). On (i): we want *ultima facie* transparency, in line with the vast majority of transparency theses. Prima facie transparency obviously does not conflict with TRANSITIVITY and so is not of issue here.

[6] Some Higher-Order theorists support transparency in unrestricted form (Gennaro 2008: 50; Kriegel 2009: 182). So we are assuming something weaker than what Higher-Order theorists already take on.

On (iii): we want transparency of conscious states (or experiences).[7] This construal has a strong precedent.[8] It also has the benefit of putting the tension with TRANSITIVITY into sharp relief. When the Higher-Order theorist talks about being aware of our conscious states, she does not just mean that we are aware of what our conscious states represent. Indeed, she *cannot* mean this. Per Higher-Order theory, the way we are conscious of our conscious states is fixed by the content of the state of higher-order awareness. States of higher-order awareness have *higher-order content*. Higher-order content is content that involves representations or representational properties. It is essential to Higher-Order theory that a (first-order) mental state is conscious only when it is represented by a state of higher-order awareness. But a (first-order) mental state is itself a representation: in most cases, it represents, and makes us aware, of things in the world (albeit unconsciously). Suppose we *did* parse TRANSITIVITY as the claim that, in virtue of being in a state of higher-order awareness, we are (merely) aware of what our first-order mental states represent—things like Eloise's tree. Then, we would be forced to say that our states of higher-order awareness don't have higher-order content. In other words, the proposal would collapse into absurdity. Higher-order states might *also* represent what our (first-order) conscious states represent (e.g., redness, trees), but it cannot *only* represent these things. So while we can describe TRANSITIVITY as the claim that conscious mental states are mental states we are aware of as having certain contents, this cannot be shorthand for the claim that we are *just* aware of what contents our conscious states have. A First-Order theorist like Tye or Dretske could *accept* that reading of TRANSITIVITY, as there is nothing here that requires a state of higher-order awareness.[9] So, there is no Higher-Order theory-compatible implementation of TRANSITIVITY on

[7] Another option found in the literature—perhaps due originally to Harman—is transparency of intrinsic qualities to [fill in the relation found in (ii) of SCHEMA]. I'm going to bracket this option, so let me explain why. First, Harman notwithstanding, the point about intrinsicality is now often dispensed with as extraneous in discussion of transparency (e.g., Stoljar 2004; Nida-Rümelin 2007; Tye 2014). Second, the notion of experiencing a feature *as* intrinsic strikes me as obscure. (Even when we consider what we experience weight as—which some say *seems* intrinsic despite not being intrinsic—I'm not sure what this means). But putting these two issues aside, I don't think the focus on intrinsic features matters. Whatever 'features' we are aware of in an experience, be they extrinsic or intrinsic, the point is that they seem to be 'features' (if anything) of objects located in our ambient environment. If being aware of a feature of *x* is a sufficient condition for being aware of *x*, then the fact that all the features we are aware of seem to be features of things other than our conscious states is good evidence that we are (in the relevant sense of 'aware') not aware of our conscious states.

[8] For formulations that take this route, see Wright (2005), Gennaro (2008) and Tye (2002, 2009, 2014).

[9] Relatedly, if we took that reading, there would then be no straightforward argument from TRANSITIVITY to Higher-Order theory, contra what its proponents suggest (see, e.g., Lycan 2001; Kriegel 2009).

which we are merely aware of the contents of our (first-order) conscious states. We are also aware (in some sense) of the conscious states themselves, and their representational properties (e.g., the property of representing red). I discuss 'Higher-Order theory-compatible' readings next in § 3.

So much for (i) and (iii). Finally, regarding dimension (ii), the relation cannot be awareness *tout court* (Stoljar 2004). *That* transparency thesis is implausible. It also makes this paper pointless. Thus the relation must be a specific type of awareness. The issue is what type. This is where the second, *disjunctive* move emerges: we simply fill in (ii) with a disjunction. Thus, collating our earlier points, we have:

W-TRANSPARENCY: For at least one conscious state M, it is impossible to:

(a) $\text{TRANSPARENCY}_{\text{DIRECT}}$: Stand in a *direct* awareness relation to M, or;
(b) $\text{TRANSPARENCY}_{\text{DE RE}}$: Stand in a *de re* awareness relation to M, or;
(c) $\text{TRANSPARENCY}_{\text{INT}}$: Stand in an *introspective* awareness relation to M,

where 'W,' again, stands for 'weak.'[10] What makes it *weak* is that it is defined disjunctively, negatively, and confined to relatively uncontroversial cases. I submit that *this* transparency thesis is a truism. It strikes me as at the very least no less plausible than TRANSITIVITY, and so a conflict with TRANSITIVITY should give the Higher-Order theorist pause.

Critically, W-TRANSPARENCY leaves the Higher-Order theorist daylight. Take $\text{TRANSPARENCY}_{\text{DIRECT}}$. That we can't stand in a direct awareness relation to a conscious state does not preclude standing in an indirect awareness relation to a conscious state. To comport with $\text{TRANSPARENCY}_{\text{DIRECT}}$, then, it must be possible to precisify TRANSITIVITY in terms of indirect awareness. Or $\text{TRANSPARENCY}_{\text{DE RE}}$: that we can't stand in a *de re* awareness relation to a conscious state does not preclude being aware *that* we are in a conscious state. So to comport with $\text{TRANSPARENCY}_{\text{DE RE}}$, it must be possible to precisify TRANSITIVITY in terms of de dicto awareness.

What about $\text{TRANSPARENCY}_{\text{INT}}$? TRANSITIVITY says nothing about introspection on its own terms; it just says that we are *aware* of our conscious states. And in principle one could affirm TRANSITIVITY while denying that there is such a thing as introspection period. Of course, such a move should only come as a last resort. So the real question is whether it is possible to precisify TRANSITIVITY in a way that allows for us to be aware of our conscious states, even though they are not introspectable. If that is possible, then TRANSITIVITY is compatible with $\text{TRANSPARENCY}_{\text{INT}}$. To most, this route might sound like a non-starter; if one is aware of *x*, *x* must be introspectable, at least in principle. But since, as we will see, one Higher-Order theorist (Uriah Kriegel) denies precisely this, it is worth exploring.

[10] For examples in the literature of commitments to (a) see Tye (2002) and Stoljar (2004), for (b) see Tye (2009, 2014), and for (c) see Kind (2003) and Kriegel (2009), among others.

3. Acceptable Consistency Glosses

Let a *consistency gloss* on TRANSITIVITY be a precisification of TRANSITIVITY that renders it consistent with W-TRANSPARENCY. A precisification of TRANSITIVITY is a consistency gloss if it is consistent with $\text{TRANSPARENCY}_{\text{DIRECT}}$, $\text{TRANSPARENCY}_{\text{DE RE}}$, or $\text{TRANSPARENCY}_{\text{INT}}$.

Consistency comes in two forms. The first is unmitigated consistency: a precisification of TRANSITIVITY is consistent with TRANSPARENCY in an unmitigated sense if, on it, we are *not* aware of our conscious states (in the relevant sense). The other is phenomenological: a precisification of TRANSITIVITY is phenomenologically consistent with W-TRANSPARENCY if, on it, we *are* aware of our conscious states, yet it does not *seem* like we are aware. Harman tells us "one might be aware of intrinsic features of experience without being aware of them *as* intrinsic features of experience" (1990: 38). A *projectivist* could accommodate transparency (Schroer 2007).

Mere phenomenological consistency is unattractive. After all, projectivism is an error theory. Unless we have strong reason to think otherwise, we should stick with an approach on which we can accommodate W-TRANSPARENCY without having to convict our experiences of massive error (cf. Tye 2002: 139). Parity seems to force us to place the same constraint on Higher-Order theory; all other things being equal, we should avoid precisifications of TRANSITIVITY that yield mere phenomenological consistency with W-TRANSPARENCY. However, in the present context, I propose we relax this constraint, but add the following proviso: it must be the case that it doesn't seem like we are aware of our conscious states in *both* introspective and non-introspective awareness. That is, for this move to carry any force, TRANSITIVITY must be phenomenologically consistent with $\text{TRANSPARENCY}_{\text{INT}}$ too. The motivation for this rider is clear. Both non-introspective consciousness and introspective consciousness have phenomenologies. Thus incompatibility with the phenomenological deliverances of introspective consciousness is just as problematic for TRANSITIVITY as incompatibility with the phenomenological deliverances of non-introspective consciousness. They both are *in*compatibilities. In any event, I'll discuss a precisification of TRANSITIVITY that permits this sort of scenario in § 4.1.

There are three types of glosses. Each gloss trades on a distinction between the *act* of consciousness and the *object* of consciousness on Higher-Order theories. *Act-Side Precisifications* concern the state of higher-order awareness M*—the act of consciousness.[11] Awareness comes in many forms, so an Act-Side Precisification will fill in the

[11] By calling M* an 'act' of consciousness, I simply mean that it is a state we are conscious *with*, even if—as Higher-Order theorists would insist—it does not have the property *being conscious*, as it is not an *object* of consciousness. Without the act, we would not be conscious of anything (cf. Dretske 1997).

details. Object-Side Precisifications are 'object-side' because they concern the first-order state M itself—the object of consciousness. Are we aware of first-order mental state *qua* mental? Are we aware of merely M, or being in M? And so on. Finally, *Complex Precisifications* combine an Act-Side with an Object-Side Precisification, telling us about the act and the object of consciousness.

All precisifications can be seen as filling in the 'in some way' clause of Transitivity. So 'in some way' should be read broadly; it can be read as referring to the type of awareness relation (where 'way' modifies the act of consciousness), or it can be read as referring to the first-order mental state (where 'way' modifies the object of consciousness).[12]

Consistency glosses are cheap. But a Higher-Order theorist cannot adopt just any consistency gloss. Consider two illustrative cases:

Deflationary Transitivity: Conscious mental states are mental states of which we are aware of automatically (in a way akin to how one automatically smiles one's smile or dance's one dance).[13]

Dispositional Transitivity: Conscious mental states are mental states of which we are disposed to be aware.

Deflationary Transitivity and Dispositional Transitivity are both Act-Side Precisifications. On the former, awareness is automatic. In the same way laughing at a person necessarily involves laughing a laugh at that person, having an experience of something (i.e., undergoing a conscious state) necessarily involves experiencing an experience of something. On the latter, it is dispositional. 'Aware', one might think, admits of a non-occurrent sense; we can be aware, as Daniel Dennett suggests, that "zebras in the wild don't wear overcoats" merely if we are disposed to assent to that proposition in virtue of what we occurrently know (1978: 104).

On Deflationary Transitivity, all there is to the idea that in having an experience one is necessarily aware of it, is the triviality that in having an experience, one experiences one's experience just as one smiles one's smile or laugh's one laugh. The problem is that this analogy is toothless. Laughing a laugh is just *laughing*; it's not as if, when

[12] Our precisifications are partial precisifications. The 'in some way' clause is unpacked only enough to comport with Transparency. The Higher-Order theorist will ultimately want to add to these.

[13] Sosa (2003), Block (2007) and Tye (2009) mention Deflationary Transitivity in passing, but there has been no sustained attempt to unpack it in any meaningful detail, save from the obvious truth that it is the sort of awareness that, if real, is reported by means of an internal accusative (Siewert 2012: 136). Indeed, neither Block nor Tye—the parties relevant here—officially endorse Deflationary Transitivity.

I laugh a laugh at a joke, I am laughing *at* my laugh (Tye 2009: 5). So, if experiencing an experience is like laughing a laugh, then experiencing an experience is just *experiencing* (cf. Rosenthal 2012). This makes Deflationary Transitivity a consistency gloss—it is not as if W-Transparency precludes our having experiences—but it comes at a steep price. The point of invoking higher-order awareness on Higher-Order theories is to distinguish between conscious and unconscious mental states, or conscious and unconscious experiences. Deflationary Transitivity, in effect, just says there *are* experiences, without telling us what *makes* something a conscious experience. No Higher-Order theorist can accept Deflationary Transitivity, since no Higher-Order theorist can accept that the sense in which we are aware of our conscious states *merely* comes down to this automatic, deflationary sense of awareness.

What about Dispositional Transitivity? A mental state might be available to a state of higher-order awareness despite our never actually (i.e., occurrently) being aware of it, just as a salt crystal might still be disposed to dissolve in water despite never getting wet.[14] So W-Transparency isn't violated. However, mere availability to higher-order awareness is a different beast. It is true that dispositional theories under a higher-order guise have been proposed (e.g., Carruthers 2000), but I don't want to quibble about whether a dispositional variant of Higher-Order theory should count as a 'real' Higher-Order theory. The point is that while there may be no tension between Dispositional Transitivity and W-Transparency, Dispositional Transitivity won't cut it as a consistency gloss since it fails to motivate Higher-Order theory, at least as understood here.[15]

The lesson from Deflationary and Dispositional Transitivity is a constraint on a consistency gloss. We alluded to the constraint in § 1. The constraint is this:

> The Compatibility Constraint: For any precisification P of Transitivity, if P is to be adopted as the 'official' precisification of Transitivity (so as to comport with W-Transparency), P must be such that for at least one Higher-Order theory H, P is consistent with H.

[14] There is a wrinkle here. For this to be a consistency gloss, we need to say that something has a dispositional property F even though it is necessarily the case that F is never actualized. I grant that this might sound odd. But as Dispositional Transitivity is just meant to be an illustrative case of a precisification the Higher-Order theorist cannot accept, we can bracket this concern.

[15] That said, it is worth noting that even Block—a staunch opponent of Higher-Order theories—when discussing Sperling paradigm cases, doesn't deny that lower-order mental states must be *accessible* to higher-order awareness, just that necessarily most are unaccessed (2011a: 567).

This allows us to refine our thesis. Let's say that a precisification P of TRANSITIVITY is an *acceptable* consistency gloss if (i) P is a consistency gloss (i.e., it comports with W-TRANSPARENCY), and (ii) P meets THE COMPATIBILITY CONSTRAINT. So, my claim is that TRANSITIVITY admits of no clear acceptable consistency gloss.

Before we press on, let me note two points of strategy. First, I do not intend to give something like a 'master' argument for the claim that TRANSITIVITY admits of no acceptable consistency gloss. Instead, I will look at precisifications on a case-by-case base. This strategy is imperfect, as practical limits preclude exhaustiveness. Yet exhaustiveness is unnecessary. So long as we hit on the most promising precisifications, our case will be strong. And again, establishing the weaker claim that TRANSITIVITY does not *readily* admit of an acceptable consistency gloss (i.e., weaker than the claim that there is *no* such gloss) is still sufficient for our secondary aim, viz. to rebut the notion that TRANSITIVITY and W-TRANSPARENCY are truisms.

Second, precisifications are divided into two groups. Group One (§ 4.1–§ 4.3) covers precisifications presented in the literature. I argue that these precisifications are not acceptable consistency glosses because they are not *consistency* glosses. Group Two (§ 4.4–§ 4.5) covers precisifications that, while not officially aired in the literature, take a more direct approach to W-TRANSPARENCY, say, by appeal to de dicto or indirect awareness. Though these precisifications are consistency glosses—there would be no reason to discuss them otherwise—they are not *acceptable* consistency glosses, as they do not meet THE COMPATIBILITY CONSTRAINT.

4. The Many Varieties of Transitivity

4.1 Unconscious Transitivity

Though not a Higher-Order theorist herself, Isabel Gois suggests that it is a confusion to think that Higher-Order theories have a problem with W-TRANSPARENCY. She says:

> [The] criticism [that there is a conflict with W-TRANSPARENCY] misses the point of the [Higher-Order] strategy…[I]f a first order state is made conscious by a second order state, that second order state will not be conscious unless it is in turn made the object of a third order state…This explains [why] our experiences should seem transparent—we are typically never aware of the higher order mental state that makes a mental state conscious (2010: 148).[16]

[16] See also Gennaro (2003) for a similar move.

On its face, this move is puzzling. W-Transparency is a claim about our (lack of) awareness of our conscious mental states. According to Higher-Order theory, conscious mental states are appropriately targeted *first-order* mental states. So that we are "typically never aware" of our higher-order states is neither here nor there (cf. Textor 2015: 2, fn. 5).

However, my guess is that what Gois actually has in mind is something bit subtler. Rosenthal often points out that experimental work on masked priming and blindsight in individuals with damage in their primary visual cortex reveals that perception can be subliminal (see, e.g., 2008: 837). This suggests a distinction between awareness and conscious awareness: subliminal perception makes us aware of things, but such states of awareness are conscious only when represented by a HOT. Only then is our awareness of the world conscious. But our awareness of our first-*order mental states* is not itself conscious until we *introspect*, until we have—at least on the HOT-theoretic picture—a second HOT directed at the first HOT.

So maybe the 'Rosenthal-Gois move' is to point out that, while we *are* always *aware* of our conscious states, since that awareness is normally unconscious, it does not *seem* like we are aware of our conscious states, and hence why we find transparency compelling. This yields the following Act-Side Precisification:

> Unconscious Transitivity: Conscious mental states are mental states of which we are unconsciously aware.

Unconscious Transitivity will not deliver *unmitigated* inconsistency. It is incompatible with $\text{Transparency}_{\text{de re}}$ since Unconscious Transitivity admits that we are aware *of* our conscious states, and incompatible with $\text{Transparency}_{\text{direct}}$ since there is no reason to think that Rosenthal or Gois have indirect awareness in mind. (In any event, we'll explore appeals to indirect awareness later in § 4.5). $\text{Transparency}_{\text{INT}}$ is also a problem since we are granting that our unconscious awareness of our mental states could become conscious.

But that's fine. For just like the projectivist, what proponents of the Rosenthal-Gois move are after is not unmitigated consistency, but phenomenological consistency. What matters is not whether I am aware of my conscious states, but whether it seems that way. However, the Higher-Order theorist would do well to recall our proviso: for phenomenological consistency, it needs to be the case that, when we introspect, it *still* doesn't seem like we are aware of our conscious states. And it is here, alas, where the Rosenthal-Gois move runs aground. It may be that, in non-introspective consciousness, it does not seem like we are aware of our conscious states because that awareness is not conscious. Yet this doesn't help if, when we do introspect, it *does* seem like we are aware of our conscious states.

Indeed, Gois tacitly admits as much; she says our higher order states are only "typically" unconscious. This concession is damaging, since non-typical cases matter. When we introspect, our higher-order awareness *will* be conscious by the Higher-Order theorist's own lights, and so our conscious states will be phenomenologically manifest, contra TRANSPARENCY$_{INT}$. Rosenthal himself comes close to admitting this:

> Harman's insistence that the only quality one is aware of. . .is that of the tomato. . .is too quick. When one sees a tomato consciously but unreflectively, one conceptualizes the quality one is aware of as a property of the tomato.

But when we introspect:

> One thinks of the quality differently when's one's attention has shifted from the tomato to one's experience of it. One then reconceptualizes the quality one is aware of as property of the experience; one then becomes conscious of that quality as the qualitative aspect of an experience, in virtue of which that experience represents a red tomato (2005: 120–121).

To say that one becomes conscious of the quality 'as. . .an aspect of experience' is to make a phenomenal claim (cf. Harman 1990: 42). It's a claim about the way the quality seems upon introspection: it seems to be a quality of experience. After all, Rosenthal often notes how higher-order concept-acquisition can affect a phenomenological change—that's the point of his frequent discussion of wine-tasting (e.g., 2005: 147). So, by parity of reasoning, we should expect the same thing to occur in *introspective* re-conceptualization. Or, at the very least, we need a reason for why Rosenthal would be entitled to say otherwise. Absent that, it appears that not only can UNCONSCIOUS TRANSITIVITY not give the Higher-Order theorist unmitigated consistency, it cannot give her phenomenological consistency either.

Perhaps another option for the Higher-Order theorist is to simply say that that the higher-order awareness is necessarily unconscious. This is UNCONSCIOUS TRANSITIVITY with a twist:

> UNCONSCIOUS TRANSITIVITY*: Conscious mental states are mental states of which we are, necessarily, unconsciously aware.

UNCONSCIOUS TRANSITIVITY* allows for phenomenological consistency, but it's unsatisfactory all the same. While it is true that Higher-Order theorists disagree over whether our awareness of our mental states is normally conscious (Kriegel) or unconscious (Rosenthal), no Higher-Order theorist says that such awareness is necessarily unconscious. And at least when it comes to Rosenthal's HOT theory—pertinent

here—it's hard to see how they can. For consider: our being consciously aware of a mental state M rather than unconsciously aware is, on this story, a matter of our having a state of higher-order awareness M* (directed at M) that is itself the object of yet another, distinct, state of higher-order awareness M**. So if UNCONSCIOUS TRANSITIVITY* is true, there must be some reason we cannot instantiate M**, despite the fact that we can instantiate M*, and despite the fact that M* and M** are *the very same kind of state*. Perhaps there are limitations on the number of levels of awareness normal humans can possess, but absent empirical evidence, there is no reason to think that the move from M* to M** is too burdensome.

So I propose we bracket UNCONSCIOUS TRANSITIVITY*. I also propose we bracket its even less plausible cousin, where all and only those first-order mental states that are transparent are the objects of a necessarily unconscious state of higher-order awareness. A better option, insofar as TRANSPARENCY$_{INT}$ goes, is to bracket unconscious higher-order awareness, and to argue that what we are aware of must outstrip what we can introspect. That possibility is taken up next.

4.2 Peripheral Transitivity

Kriegel (2009) offers the following precisification of TRANSITIVITY:

> PERIPHERAL TRANSITIVITY: Conscious mental states are mental states of which we are peripherally aware.

PERIPHERAL TRANSITIVITY is an Act-Side Precisification. Peripheral awareness contrasts with *focal* awareness, where the distinction between focal and peripheral awareness is primarily a distinction regarding the distribution of attention. Typically, this is a distinction we apply to outer awareness, that is, awareness of environmental objects and their properties. Kriegel thinks it applies just as well to inner awareness, that is, awareness of mental states. I have doubts about this claim, but I'll grant it here.

On the face of it, PERIPHERAL TRANSITIVITY doesn't jibe with TRANSPARENCY$_{DE\ RE}$ and TRANSPARENCY$_{DIRECT}$. But this won't trouble Kriegel, since he ascribes only to TRANSPARENCY$_{INT}$. Kriegel frames transparency this way: for any experience *e* and any feature F, if F is an introspectable feature of *e*, then F is part of the first-order content of *e* (ibid: 181). A first-order content is a content involving neither representations nor representational properties. Higher-order content, recall, involves representations or representational properties. Since conscious states are representations, Kriegel just means that conscious states are not introspectable. So his construal is equivalent to an unrestricted version of our TRANSPARENCY$_{INT}$.

Kriegel argues that PERIPHERAL TRANSITIVITY is consistent with TRANSPARENCY$_{INT}$ by appealing to an *attention-shift* model of introspection (ibid: 183). Instead of involving a new representational state, introspection reorganizes the center/periphery structure of one's overall experience. Suppose I have a visual experience of a fancy Fabergé egg. According to Kriegel's self-representationalism, in non-introspective cases of consciousness, I have focal awareness of my Fabergé egg but peripheral awareness of my experience of the Fabergé egg. When I introspect, however, I am focally aware of my experience of the Fabergé egg and peripherally aware of the Fabergé egg (ibid).

Putting these pieces together, we get the following case for compatibility. When we introspect our experiences—that is, stand in a relation of focal awareness to them—our prior peripheral awareness is "annihilated" (ibid: 184). We cannot introspect our experiences *qua* their being in the periphery if introspection is a matter of inner focal attention (ibid: 185). We can be aware of our experiences, then, without those experiences being introspectable. As such, there is no tension between PERIPHERAL TRANSITIVITY and TRANSPARENCY$_{INT}$.

I want to make two points in response. First, Kriegel musters little evidence for the attention-shift model of introspection, and PERIPHERAL TRANSITIVITY is not an acceptable consistency gloss without it. We're told that the model has the "advantage of illuminating the fact that introspection does not feel, phenomenologically speaking, like performing a 'dramatic' mental act, an act that creates an altogether new representation" (ibid: 183). For my part, it's hardly clear that entering into a new representational state *must* feel like something, and so it's unclear why the fact—supposing it is a fact—that introspecting does *not* feel this way should make us think that we are *not* entering into a new state. I might, say, come to believe that grass is green, and thus be in a new representational state, but not feel anyway in virtue of doing so.

But I don't want to stake my case on this. For, second, even if we grant Kriegel the attention-shift model, his case for compatibility fails. To make things concrete, lets return to the fancy Fabergé egg case. Call the visual experience *e* and the subject S. On Kriegel's self-representational theory, *e* has the following two essential features: *e* represents the egg, and *e* represents itself. Since it is subjects that are aware of things, and subjects are aware of things *in virtue* of having experiences, when S has *e*, S is thus aware of two things: the apple, and *e* itself. However, given Kriegel's adherence to PERIPHERAL TRANSITIVITY, S is only *peripherally* aware of *e*. So far that's just the view. But here's the problem. For PERIPHERAL TRANSITIVITY to be compatible with TRANSPARENCY$_{INT}$, it has to be that when S introspects *e*, the *only* thing S is aware of is the egg (a feature of first-order content). Kriegel seems to think that this is the case (2009: 185). Alas, he's not entitled to that conclusion. What Kriegel *is* entitled to say is that, upon introspecting:

(1) S's *peripheral* awareness of *e* is annihilated.

Kriegel is entitled to assert (1) because (1) follows from what introspection is on the attention-shift model. You cannot, as Kriegel notes, be introspectively aware of *e* qua item of peripheral awareness if introspection is a matter of being focally aware of *e*. Yet this doesn't mean that, upon introspecting:

(2) S's *introspective* awareness of *e* is annihilated.

Nor that:

(3) S's introspective awareness is *exhausted* by the egg.

But both (2) and (3) are needed to square Peripheral Transitivity with $\text{Transparency}_{\text{INT}}$; if ¬ (2), then there is introspectable higher-order content (*e*, by Kriegel's lights, is a representation), and if ¬ (3), then the introspectable features of *e* are not limited to *e*'s first-order content. Compare an outward-directed case, where in a single visual experience, S has two items in her visual field: the Fabergé egg, and now a mug. And suppose S's distribution of attention is such that S is focally aware of the egg, but only peripherally aware of the mug. It is of course perfectly normal for S's distribution of attention to change, but there is nothing about this change *itself* that should make us thing that, once the change occurs, S is no longer aware of the mug, but only the egg. What's different is the *way* S is aware of the egg. We are given no reason to think this should be different for S's inner introspective awareness of *e*.[17]

4.3 Non-State-Relational Transitivity

As it stands, Transitivity posits a *relation* between a higher-order state and a first-order state. But not all Higher-Order theorists construe Transitivity this way. Instead, some advance:

[17] It's also worth noting that Kriegel's position might be internally unstable. For on his view, peripheral inner awareness is essential to conscious experience. So, if introspection destroys peripheral inner awareness as Kriegel contends, then introspection should destroy the experience we are introspecting. Consequently, it should follow that when we introspect we are not aware of the Fabergé egg either, for without the experience there is no awareness of the egg. But that means introspection, by Kriegel's lights, should reveal *nothing*, which, of course, is an untenable conclusion at best.

NON-STATE-RELATIONAL TRANSITIVITY: A conscious mental state is a mental state whose subject is aware of itself as being in that state.[18]

NON-STATE-RELATIONAL TRANSITIVITY is an Object-Side Precisification. And it appears promising. For it says that we are aware of *ourselves* as being in conscious states, not simply that we are aware of our conscious states.[19] These are different claims.

An immediate worry concerns THE COMPATIBILITY CONSTRAINT. If it follows from NON-STATE-RELATIONAL TRANSITIVITY that being in a conscious state does not consist in being aware of our mental states, NON-STATE-RELATIONAL TRANSITIVITY will fail THE COMPATIBILITY CONSTRAINT. And if NON-STATE-RELATIONAL TRANSITIVITY fails THE COMPATIBILITY CONSTRAINT, it is not an acceptable consistency gloss, even if it is a consistency gloss. One could then argue as follows. Awareness is a relation. What we are aware of is what we stand in a (suitable) relation to. On NON-STATE-RELATIONAL TRANSITIVITY we *don't* stand in such a relation to our mental states—that's the point (see also fn. 19). So, being in a conscious state does not consist in being aware of our mental states. Therefore, NON-STATE-RELATIONAL TRANSITIVITY fails THE COMPATIBILITY CONSTRAINT.[20]

Does this argument work? I see two objections. First, I could see one rejecting the claim that awareness is a relation; adverbialists do just this. Though no Higher-Order theorist that I know of is an adverbialist, there is no in-principle incompatibility here.[21]

The second objection claims that I stack the deck. Yes, NON-STATE-RELATIONAL TRANSITIVITY may not entail TRANSITIVITY, but to complain on that score misses the whole point; after all, given that NON-STATE-RELATIONAL TRANSITIVITY is the thesis some Higher-Order theorists

[18] See Jacob Berger (2014) and Richard Brown (2012, 2015) for discussion. I have added 'state' to Brown's original 'non-relational transitivity' for the reason he notes (2012), viz. that calling this simply 'non-relational' is misleading: there is still a relation between the representation and the subject, despite there being no further relation between the higher-order state and some first-order state.

[19] The motivation for NON-STATE-RELATIONAL TRANSITIVITY stems (in part) from its potential ability to address the objection from *radical misrepresentation*, that is, the possibility of having a HOT that represents a state that one is not in (Neander 1998; Block 2011b). Briefly, here is how the move works. Many think awareness is factive. But we can distinguish between *weak factivity* and *strong factivity*. On weak factivity, if I am aware of *x*, *x* exists. On strong factivity, if I am aware of *x* as being F, not only does *x* exist, but *x* is also F. Proponents of Non-State-Relational Transitivity accept weak factivity, but reject strong factivity. In representing *oneself as* being in a certain state *F*, I am guaranteed to exist, but I'm not similarly guaranteed to be in *F* (Rosenthal, private communication). If for a mental state to be conscious is just for one to be suitably aware of oneself as being in that state, that there is no first-order state in radical misrepresentation is otiose.

[20] See Coleman (2015) for an independent argument that tallies well with this claim.

[21] Although Kriegel (2011) flirts with this option, he ultimately falls short of endorsing it, albeit for reasons different from those relevant here.

adopt, the right reaction seems to be that it is TRANSITIVITY that ought to be read in less restrictive terms, that is, terms that don't make it incompatible with NON-STATE-RELATIONAL TRANSITIVITY.[22]

I'm inclined to think, with Sam Coleman (2015), that this move is misguided: radical misrepresentation aside, if it follows from a Higher-Order theory that being in a conscious state does not consist in being aware of one's mental states, we are changing the subject. But perhaps that's too hasty. When discussing radical misrepresentation, Rosenthal (2011) makes clear that what matters on his view is explaining subjective mental appearance, which may or may not include actual awareness of first-order mental states. The Higher-Order theorist—or at least those sympathetic to a Rosenthalian HOT theory—can rest easy allowing that HOTs, and HOTs alone, supply the subjective appearances distinctive of conscious experience.

None of this changes the underlying issue, however. Suppose the proponent of NON-STATE-RELATIONAL TRANSITIVITY adopts adverbialism. This will help with THE COMPATIBILITY CONSTRAINT, but it will do nothing with respect to W-TRANSPARENCY; the adverbialist has to contend with W-TRANSPARENCY no less than the non-adverbialist. Suppose instead we go the second route, and grant that insisting that NON-STATE-RELATIONAL TRANSITIVITY entail TRANSITIVITY is too strong. Again, the problem with W-TRANSPARENCY is unresolved. For presumably there will be (many!) cases where I *am* in the conscious state I am aware of myself as being in, and so cases where we will still need to ask in *what sense* I am aware of those states, and whether that sense comports with W-TRANSPARENCY. NON-STATE-RELATIONAL TRANSITIVITY doesn't obviously speak to this latter question, though; the awareness we have of ourselves is *de re*, and presumably direct, but whether that's also true of the awareness we have of our conscious states is another issue. So as it stands, NON-STATE-RELATIONAL TRANSITIVITY is not a consistency gloss.

4.4 De Dicto Transitivity

Here is where we are. The precisifications we have encountered—DEFLATIONARY, DISPOSITIONAL, UNCONSCIOUS, PERIPHERAL, and NON-STATE-RELATIONAL TRANSITIVITY—have all been aired in the literature, albeit not always with TRANSPARENCY in mind. None of them worked out. DEFLATIONARY and DISPOSITIONAL TRANSITIVITY were consistency glosses but not acceptable consistency glosses, while PERIPHERAL, UNCONSCIOUS and NON-STATE-RELATIONAL were not even consistency glosses.

At this juncture a more direct approach may prove fruitful. W-TRANSPARENCY is a disjunction with three disjuncts. The HO theorist is thus free to simply pick one disjunct and aim for consistency with that disjunct. Suppose she opted for $\text{TRANSPARENCY}_{\text{DE RE}}$. Then she could say:

[22] Thanks to an anonymous referee for pressing me here.

De Dicto Transitivity: Conscious mental states are mental states of which we have *fact*-awareness.

De Dicto Transitivity is an Object-Side Precisification. Dretske (1993, 1995, 1999) famously distinguished between fact-awareness (or 'awareness of facts') and thing-awareness (or 'awareness of things'). Awareness of things comes in three variants: awareness of objects, properties, and events. For simplicity, I'll speak generically of awareness of things. Unlike awareness of things, awareness of facts implies a deployment of concepts. If I am aware *that* the water is boiling, I must possess, and employ, the concept WATER and BOILING in my awareness of the water (1993).[23]

Now ordinary talk of our being aware of our conscious states is likely neutral on whether such awareness is *de re*; the sentence 'I am aware *of* my desire for tiramisu' can be given a de dicto reading, viz. 'I am aware *that* I have a desire for tiramisu' (cf. Tye 2014). De Dicto Transitivity picks up on this insight. Transparency$_{\text{de re}}$ permits fact-awareness, so De Dicto Transitivity is a clear consistency gloss. The problem is that it doesn't entail Transitivity, and thus doesn't meet The Compatibility Constraint. My being aware *that* I am in a conscious state does not entail I am aware *of* my conscious state. Fact-awareness does not entail thing-awareness. Dretske (1999) notes that one can be aware of the fact that the water is boiling (aware *that* it is boiling) without being aware of the objects or the properties that make it a fact that the water is boiling. I may hear a timer go off, but if I am looking the other way, I am not (*de re*) aware of the water, the flame, or the property of being boiled. Fact-awareness can be *indirect*. Fact-awareness is indirect when it is achieved by awareness of properties and/or objects other than those involved in the fact, and direct otherwise. Awareness that your child has a raised temperature is indirect when you use a thermometer but direct when you feel his forehead (ibid: 108).

The Higher-Order theorist might try to amend De Dicto Transitivity by stipulating that the fact-awareness is *direct*. This would yield the following complex precisification:

Direct De Dicto Transitivity: Conscious mental states are mental states of which we are directly fact-aware.

Since Direct De Dicto Transitivity entails Transitivity the aforementioned problem is rectified. Yet it is a pyrrhic victory, for Direct De

[23] While Dretske's original distinction concerns (outer-directed) perceptual awareness, Giustina & Kriegel (2016) make clear that this distinction applies equally to other forms of inner awareness, that is, the sort of awareness, perceptual or otherwise, we would have our mental states. Giustina & Kriegel (ibid) also provide battery of forceful arguments for the distinction between fact and thing-awareness from a Higher-Order perspective.

Dicto Transitivity is not a consistency gloss. It is not a consistency gloss because it says we *do* have *de re* awareness of our conscious states. Indeed, this is exactly what proponents of transparency often warn against. Here is Tye:

> [W]e can be aware *that* we are having such experiences. No one denies that. But such factive awareness is not supported by awareness of, or attention to, the token experiences themselves. This is the familiar and widely accepted doctrine of transparency (2009: 5).

So De Dicto Transitivity and Direct De Dicto Transitivity seem to fail. The former fails because it does not entail Transitivity, and the latter because it is not a consistency gloss.[24]

What this shows, I take it, is that the Higher-Order theorist will have to reject $\text{Transparency}_{\text{DE RE}}$ outright; if there is hope for an acceptable consistency gloss, it will be had by going in for $\text{Transparency}_{\text{INT}}$ and adopting a 'transparency-friendly' account of introspection, or by going in for indirect awareness, and opting for $\text{Transparency}_{\text{DIRECT}}$. We have already seen why the former move likely fails. We'll broach the latter next.

4.5 Indirect Transitivity

All that remains for the Higher-Order theorist, then, is $\text{Transparency}_{\text{DIRECT}}$. This yields:

> Indirect Transitivity: Conscious mental states are mental states of which we are indirectly aware.

Indirect Transitivity is an Act-Side Precisification. It tells us that the relation of awareness is indirect. Since $\text{Transparency}_{\text{DIRECT}}$ rules out only direct awareness, invoking indirect awareness is fine. The question is whether Indirect Transitivity meets The Compatibility Constraint.

We discussed cases of direct and indirect fact-awareness earlier, but we haven't yet given an explicit definition of direct and indirect awareness. Daniel Stoljar's definition gives us a start: "A subject S is *indirectly* aware of x just in case S is aware of x by being aware of y, where y is distinct from x," and "S is *directly* aware of x just in case (i) S is aware of x and (ii) S is not (merely) indirectly aware of x" (2004: 380).[25] The indirect object of awareness on Indirect Transitivity will just be the

[24] This is true even if the Higher-Order theorist *rejects* the distinction between thing and fact-awareness.

[25] Stoljar is following the work of William Alston (1971) on epistemic immediacy and Frank Jackson (1977) on perceptual immediacy. Stoljar invokes 'merely' to rule out the possibility of being both directly and indirectly aware of something (e.g., being aware of the coffee cup by feeling its heat, and seeing it).

conscious state. That's clear enough. But there must also be something that *anchors* our being indirectly aware of our conscious state; there must be a direct object of awareness. The obvious option—indeed the option proponents of TRANSPARENCY$_{\text{DIRECT}}$ like Tye have pushed—is that the anchor is whatever that conscious state represents.[26] So the emerging picture would now be this. Suppose Carmichael has a first-order visual perception of a blue sky. By TRANSITIVITY, if that perception is conscious, Carmichael is in a state of higher-order awareness M* that has as one of its objects the first-order visual perception. INDIRECT TRANSITIVITY would then add to TRANSITIVITY the following qualification: Carmichael is conscious of his visual perception by being conscious of what that perception represents, viz. a blue sky.

INDIRECT TRANSITIVITY entails TRANSITIVITY. Conscious states are mental states we are aware of, albeit indirectly. But we might ask whether we are aware of a fact ('indirect fact-awareness') or a thing ('indirect thing-awareness'). This yields two more complex precisifications:

INDIRECT-FACT TRANSITIVITY: Conscious mental states are mental states of which we have indirect-fact awareness.

INDIRECT-THING TRANSITIVITY: Conscious mental states are mental states of which we have indirect-thing awareness.

We already saw one reason why INDIRECT-FACT TRANSITIVITY is out. Like DE DICTO TRANSITIVITY, INDIRECT-FACT TRANSITIVITY does not entail TRANSITIVITY, since awareness *that* I am in a conscious state does not entail awareness *of* a conscious state.

But bracket that. There's still another problem. INDIRECT-FACT TRANSITIVITY is based on Dretske's displaced model of introspection. By Dretske's lights, this makes the higher-order awareness in INDIRECT-FACT TRANSITIVITY *inferential* (1995: 42). The idea is this. Suppose I am aware that an experience is F. Then there has to be a premise connecting our perceptual *de re* beliefs about the environment (e.g., based on my seeing that my experience's object has some property F*) to our introspective belief that we are having an experience with F (e.g., that the object would not have F* unless that experience had F). Yet many Higher-Order theorists would *reject* that the relevant the relevant higher-order awareness is inferential. One can infer from publicly accessible evidence that one is angry without the anger being thereby rendered conscious (Rosenthal 1997b: 737). Thus the higher-order awareness we have of our anger must be non-inferential for that anger to be conscious. Rosenthal's HOT theory stops a bit short of this conclusion, since on his view the higher-order awareness need only be *subjectively*

[26] See, for example, Dretske (1995, 1999); Tye (2002, 2003). Tye (2014) now denies that we even have indirect awareness of our conscious states.

unmediated, which is not to say that it isn't, in fact, inferential.[27] Even so, this doesn't help. The HOT theorist who adopts INDIRECT-FACT TRANSITIVITY is claiming that being in a conscious state consists in our being (indirectly) aware of our (first-order) mental states by being (directly) aware of what those states represent (e.g., Fabergé egg), where the inference from the latter to the former seems unmediated. Alas, the HOT theorist, at least as traditionally understood, *cannot* say this. For part of HOT theory is the claim that having a conscious experience, and so being in a conscious state, can be *decoupled* from the presence of a first-order state. This is the heart of the HOT-theoretic response to the problem of radical misrepresentation (see fn. 19). According to Rosenthal, 'is in a conscious states' is non-extensional (2011: 432). If that's so, the HOT theorist cannot *build in* a reliance on direct awareness of what our first-order states represent, because, *ex hypothesis*, there won't be *any* first-order state in cases of radical misrepresentation. However, we'll still have state-consciousness. As such, INDIRECT-FACT TRANSITIVITY won't work.[28]

This leaves us with INDIRECT-THING TRANSITIVITY. INDIRECT-THING TRANSITIVITY has at least two advantages over INDIRECT-FACT TRANSITIVITY. First, it straightforwardly entails TRANSITIVITY. Since there is no appeal to fact-awareness, there is no worry about whether awareness-that entails awareness-of. Second, although we are still dealing with indirect awareness, thing-awareness strikes me as providing more flexibility as to what the attendant indirect awareness could be. We are indirectly aware of y only if we are aware of y *by* being directly aware of x. The proponent of INDIRECT-FACT TRANSITIVITY treats 'by' inferentially, but a more natural option for the proponent of INDIRECT-THING TRANSITIVITY is to gloss the 'by' with the *in virtue of*-relation. To say that I am aware of y by being aware of x, then, will be to say that I am aware of y *in virtue of* being aware of x.

To see what this amounts to, we can start with a more mundane example, adapted from Frank Jackson (1977). Suppose that when Carmichael parks his car on a busy street, his car makes actual physical contact with curb. It is natural to say that Carmichael's car touches the curb in virtue of its tire touching the curb. William Fish (2004) offers a nice way of cashing this out. We have two states of affairs: a *tire-touching-curb* state of affairs and a *car-touching-curb* state of affairs. The latter clearly depends on the former: Carmichael's car will come into contact with the curb only when a part of the car does (ibid: 6). But the converse does not hold: whether the former state of affairs constitutes the latter depends, as Fish notes, "upon the other relationships—relationships that do not alter the underlying *kind* of state of

[27] Thanks to an anonymous referee for reminding me of this point.

[28] That inference is required is not uncontroversial. Tye denies the need for inference (2002: 144). Yet as Murat Aydede points out (2003: 62, fn. 5), Tye just flatly asserts that inference is not needed.

affairs it is—it enters into" (ibid). The *tire-touching-curb* state of affairs is a type of state of affairs that could have obtained even if it fails to constitute a *car-touching-curb* state of affairs. This suggests, to quote Fish, the following take on the perceptual 'in virtue of:' "an S φ's in virtue of p only if p could have obtained otherwise, and p's obtaining would not thereby have constituted S's φ'ing" (ibid: 7).[29]

Applying Fish's account of the (perceptual) 'in virtue of' relation to our construal of indirect awareness, the pertinent question becomes the following. Let M be Carmichael's (first-order) perceptual state, the state in virtue of which Carmichael is aware of the Fabergé egg. Can the Higher-Order theorist say that Carmichael's being in a conscious state (M's being conscious) consists in his being (indirectly) aware of M in virtue of being (directly) aware of what M represents, in the sense that Carmichael could have been aware of what M represents *without* his thereby being aware of M itself?

No. Carmichael's (first-order) visual awareness of the Fabergé egg is either conscious or unconscious. If it's conscious, we have a reversal of the Higher-Order theorist's order of explanation. On Higher-Order theory, the *Carmichael-consciously aware-of-Fabergé egg* state of affairs is not the kind of state of affairs that can obtain independently of Carmichael being aware of his visual awareness of the Fabergé egg. If the (first-order) awareness is unconscious, the same problem about radical misrepresentation mentioned earlier resurfaces: having a conscious experience can be decoupled from the presence of a first-order state, and so absent *any* thing-awareness of the Fabergé egg. It is true that the unconscious awareness of the Fabergé egg can obtain independently of our being in a conscious state—that's subliminal perception—but the converse is also true for Rosenthal: 'is in a conscious state', recall, is non-extensional. And as Josh Weisberg (2008: 169) has nicely argued, though a feature on HOT theories, this decoupling is actually a *bug* for self-representational theories. So there is likely no recourse there either.

Consequently, Indirect-Thing Transitivity fails to meet The Compatibility Constraint, and therefore is not an acceptable consistency gloss. Moreover, absent a more amenable way of cashing out the 'by' in

[29] Fish is explicitly working with a *theoretical* sense of 'in virtue of'. As a general analysis of the 'in virtue of' relation, there are clear counterexamples, which Fish himself recognizes. For instance, S can be a sibling in virtue of S's having the same parents as a different offspring, yet it's not the case that S's having the same parents with a different offspring could have obtained otherwise and would not have thereby constituted S's being a sibling. (Thanks to an anonymous referee for this example). Fish wants to give an analysis that captures the epistemic asymmetry essential to indirect perception, where what we directly perceive is on surer epistemic footing than what we indirectly perceive. Although Fish rightly points out that the scope of the direct-versus-indirect distinction goes beyond classic sense-datum theories, we can use them to make the point: S perceives a material object y in virtue of perceiving a sense datum x only if S's perceiving x is the sort of experience she could have had even if she were to fail to perceive y (2004: 7–8).

indirect awareness, this point generalizes: any precisification of TRANSITIVITY that appeals to indirect awareness is not an acceptable consistency gloss.

5. Two Purported Truisms: Redux

The best theory of consciousness, all other things being equal, will be a theory that can grant TRANSITIVITY and W-TRANSPARENCY. If I am right, and Transitivity does not admit of an acceptable consistency gloss, then our theorizing about consciousness is in a bit of a bind, for there will be no such theory. One of these theses will have to go, and it's far from obvious which one.

But even if we are only entitled to the weaker claim that it is exceedingly difficult to find an acceptable consistency gloss, something interesting still follows. The moral is methodological. Let me close by spelling this out.

Truisms help us theorize. They help by giving us somewhere to *start* —somewhere that we do not have argue *to*, but can argue *from*. They function like this because part of what it is to be a truism is to be seen as obvious in some sense, and we often don't expend energy arguing for things that are obvious. Now, we noted at the outset that TRANSITIVITY and W-TRANSPARENCY are, for many, truisms about consciousness, and that because of this, we can theorize from them, and often with far-sweeping, substantive results.[30] But if it seems very likely that TRANSITIVITY and W-TRANSPARENCY are likely incompatible, this approach is ill-advised, for TRANSITIVITY and W-TRANSPARENCY ought *not* be treated as truisms. Suppose we have two propositions p and p', both of which, when taken individually, are seen by S to be obviously true. (We can add that S's credence in p and p' are the same). Though their conjunction does not form a straightforward contradiction, when scrutinized, however, S has an exceedingly difficult time seeing how $[p \wedge p']$ could possibly true. What should S do? It's seems clear, or at least prudent, that S ought to *downgrade* her credence in p and p'. That doesn't mean giving up on either altogether. That would be bad, as it is out of line with practice not just in philosophy, but theorizing everywhere. But it does likely mean that S should downgrade her credence enough such that p and p' no longer function as

[30] Examples are plentiful. Just focusing on TRANSITIVITY, variants of the so-called 'Simple Argument' are rife (e.g., Lycan 2001; Rosenthal 2005; Kriegel 2009). This takes TRANSITIVITY as a self-evident thesis—who could deny that we can be inconscious state yet be unaware of it?—in essence delivering Higher-Order theory for free. Rosenthal expresses a similar sentiment when he tells us that TRANSITIVITY makes it "fairly easy" to get started on the question of what makes a mental state conscious (2005: 4). Again, I am not claiming that no arguments that TRANSITIVITY have been aired. (In an appendix to his [2009], Kriegel does a nice job canvassing several arguments for TRANSITIVITY). My point is only that this attitude (of acting like TRANSITIVITY is a truism) is widespread, not that it has no exceptions.

truisms in her theorizing since there is a *very* good chance that at least one is false. By extension, when it comes to TRANSITIVITY and W-TRANSPARENCY, this means re-thinking how we approach our theorizing about what is the most familiar aspect of our lives: consciousness. Given the pride of place TRANSITIVITY and W-TRANSPARENCY currently enjoy in our theory building, we have our work cut out for us if I am right.[31]

References

Alston, William. "Varieties of Privileged Access," *American Philosophical Quarterly* 8 (1971): 223–41.

Aydede, Murat. "Is Introspection Inferential?" *Privileged Access: Philosophical Accounts of Self-Knowledge*. Ed. Brie Gertler. Ashgate, 2003.

Berger, Jacob. "Consciousness is Not a Property of States: A Reply to Wilberg," *Philosophical Psychology* 27 (2014): 829–842.

Block, Ned. "Mental Paint and Mental Latex," *Philosophical Issues* 7 (1996): 19–49.

Block, Ned. "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," *Behavioral and Brain Sciences* 30 (2007): 481–548.

Block, Ned. "Perceptual Consciousness Overflows Cognitive Access," *Trends in Cognitive Sciences* 15 (2011a): 567–575.

Block, Ned. "The Higher-Order Approach to Consciousness is Defunct," *Analysis* 71 (2011b): 419–431.

Brown, Richard. *Review of 'The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts' by Rocco J.* Notre Dame Philosophical Reviews: Gennaro, 2012.

Brown, Richard. "The HOROR Theory of Consciousness," *Philosophical Studies*. 172 (2015): 1783–1794.

Bourget, David. *The Representational Theory of Consciousness*, Dissertation. Australian National U, 2010.

Campbell, John. *Reference and Consciousness*. Oxford: Oxford UP, 2002.

Carruthers, Peter. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge UP, 2000.

Chalmers, David. "The Contents of Consciousness: Reply to Hellie, Peacocke and Siegel," *Analysis* 73 (2013): 345–368.

Coleman, Sam. "Quotational Higher-Order Though Theory," *Philosophical Studies* 172 (2015): 2705–2733.

Dennett, Daniel. *Brainstorms*. Cambridge, MA: MIT Press, 1978.

Dretske, Fred. "Conscious Experience," *Mind* 102 (1993): 263–283.

Dretske, Fred. *Naturalizing the Mind*. Cambridge, MA: MIT Press, 1995.

Dretske, Fred. "What Good is Consciousness?," *Canadian Journal of Philosophy* 27 (1997): 1–15.

Dretske, Fred. "The Mind's Awareness of Itself," *Philosophical Studies* 95 (1999): 103–24.

Fish, William. "The Direct/Indirect Distinction in Contemporary Philosophy of Perception," *Essays in Philosophy* 5 (2004): 1–13.

Gennaro, Rocco. "Representationalism, Peripheral Awareness, and the Transparency of Experience," *Philosophical Studies* 139 (2008): 39–56.

Gennaro, Rocco. "Papineau on the Actualist HOT Theory of Consciousness," *Australasian Journal of Philosophy* 81 (2003): 581–586.

Gennaro, Rocco. *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. Cambridge, MA: MIT Press, 2011.

[31] Many thanks to an anonymous referee, Mahrad Almotahari, Colin Klein, Jessica Gottlieb, Saja Parvizian, David Rosenthal, Marya Schechtman, Jeff Speaks, Daniel Stoljar and especially David Hilbert for helpful feedback.

Giustina, Anna, and Uriah Kriegel. "Fact-Introspection, Thing-Introspection, and Inner Awareness," *Review of Philosophy and Psychology* (2016) doi:10.1007/s13164-016-0304-5.
Gois, Isabel. "A Dilemma for Higher-Order Theories of Consciousness," *Philosophia* 38 (2010): 143–156.
Harman, Gilbert. "The Intrinsic Quality of Experience," *Philosophical Perspectives* 4 (1990): 31–52.
Hellie, Benj. "Higher-Order Intentionalism and Higher-Order Acquaintance," *Philosophical Studies* 134 (2007): 289–324.
Jackson, Frank. *Perception*. Cambridge: Cambridge UP, 1977.
Kim, Jaegwon. "Events: Their Metaphysics and Semantics," *Philosophy and Phenomenological Research* 51 (1991): 641–646.
Kind, Amy. "What's So Transparent About Transparency?," *Philosophical Studies* 115 (2003): 225–244.
Kriegel, Uriah. *Subjective Consciousness: A Self-Representational Approach*. Oxford: Oxford UP, 2009.
Kriegel, Uriah. *The Sources of Intentionality*. Oxford: Oxford University Press, 2011.
Lycan, William. *Conscious Experience*. Cambridge: Oxford UP, 1996.
Lycan, William. "A Simple Argument for a Higher-Order Representation Theory of Consciousness," *Analysis* 61 (2001): 3–4.
Lyyra, Pessi. "Two Senses for 'Givenness of Consciousness'," *Phenomenology and Cognitive Science*. 8 (2009): 67–87.
Martin, M. G. F. "The Transparency of Experience," *Mind and Language* 4 (2002): 367–425.
Moore, G. E. The Refutation of Idealism. *Philosophical Studies*. Ed. G. E. Moore, London: Routledge, 1922: 1–30.
Neander, Karen. "The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness," *Philosophical Perspectives* 12 (1998): 411–434.
Nida-Rümelin, Martine. "Transparency of Experience and the Perceptual Model of Phenomenal Awareness," *Philosophical Perspectives* 21 (2007): 429–455.
Rosenthal, David. "Phenomenal Consciousness and What It's Like," *Behavioral and Brain Sciences* 20 (1997a): 156–57.
Rosenthal, David. "A Theory of Consciousness." *The Nature of Consciousness*. Eds. Ned Block, Owen J. Flanagan, Guven Guzeldere. MIT Press, 1997b.
Rosenthal, David. "Consciousness and Metacognition." *Metarepresentation*. Ed. Dan Sperber. Oxford UP, 2000.
Rosenthal, David. "How Many Kinds of Consciousness?," *Consciousness and Cognition* 11 (2002): 653–665.
Rosenthal, David. *Consciousness and Mind*. Oxford: Clarendon Press, 2005.
Rosenthal, David. "Consciousness and its Function," *Neuropsychologia* 46 (2008): 829–840.
Rosenthal, David. "Higher Order Theories of Consciousness." *Oxford Handbook of the Philosophy of Mind*. Eds. B. McLaughlin, A. Beckermann. Oxford: Clarendon Press, 2009. 239–52.
Rosenthal, David. "Exaggerated Reports: A Reply to Block," *Analysis* 71 (2011): 431–437.
Rosenthal, David. Conscious Awareness, Higher-Order Theories, and Overflow. *Towards a Science of Consciousness Conference*, Presentation. Tucson, Arizona, 2012.
Rosenthal, David. "Quality Spaces and Sensory Modalities." *Phenomenal Qualities: Sense, Perception, and Consciousness*. Eds. Paul Coates, Sam Coleman. Oxford UP, 2015.
Schroer, Robert. "Reticence of Visual Phenomenal Character: A Spatial Interpretation of Transparency," *Australasian Journal of Philosophy* 85 (2007): 393–414.
Siewert, Charles. "On the Phenomenology of Introspection." *Introspection and Consciousness, 129–168*. Eds. Declan Smithies, Daniel Stoljar. Oxford UP, 2012.
Smith, A. D. "Translucent Experiences," *Philosophical Studies* 140 (2008): 197–212.
Sosa, Ernest. Privileged Access. *Consciousness: New Philosophical Perspectives*. Ed. Q. Smith and A. Jokic, Oxford: Oxford UP, 2003. 273–294.
Speaks, Jeff. "What Are Debates About Qualia Really About?," *Philosophical Studies* 170 (2014): 1–26.

Stoljar, Daniel. The Argument From Diaphanousness. *Language, Mind and World: Special Issue of the Canadian Journal of Philosophy*. Ed. M. Escurdia, Robert J. Stainton and Christopher D. Viger. U of Alberta P, 2004. 341–90.
Textor, Mark. ""Inner Perception Can Never Become Inner Observation": Brentano on Awareness and Observation," *Philosophers' Imprint* 15 (2015).
Tye, Michael. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press, 1995.
Tye, Michael. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press, 2000.
Tye, Michael. "Representationalism and the Transparency of Experience," *Noûs* 36 (2002): 137–51.
Tye, Michael. *Consciousness and Persons: Unity and Identity*. Cambridge, MA: MIT Press, 2003.
Tye, Michael. *Consciousness Revisited: Materialism Without Phenomenal Concepts*. Cambridge, MA: MIT Press, 2009.
Tye, Michael. "Transparency," *Qualia Realism and Representationalism. Philosophical Studies* 170 (2014): 39–57.
Weisberg, Josh. "Same Old, Same Old: The Same-Order Representational Theory of Consciousness and the Division of Phenomenal Labor," *Synthese* 160 (2008): 161–181.
Weisberg, Josh. "Abusing the Notion of What-It's-Like-Ness: A Response to Block," *Analysis* 71 (2011): 443–448.
Wright, Wayne. "Distracted Drivers and Unattended Experience," *Synthese* 144 (2005): 41–68.