



Zahavi versus Brentano: A Rejoinder

Kenneth Williford
Department of Philosophy
St. Cloud State University
720 4th Avenue South
St. Cloud, MN 56301
U.S.A.
© K. Williford
kwilliford@stcloudstate.edu

PSYCHE 12 (2), May 2006

KEYWORDS: Brentano, Gurwitsch, Husserl, Infinite Regress, Intentionality, Naturalism, Objects, Phenomenology, Self-acquaintance, Self-consciousness, Self-knowledge, Self-representation, Shoemaker, Zahavi

Abstract: Dan Zahavi has argued persuasively that some versions of self-representationalism are implausible on phenomenological and dialectical grounds: they fail to make sense of primitive self-knowledge and lead to an infinite regress. Zahavi proposes an alternative view of ubiquitous prereflective self-consciousness—the phenomenological datum upon which Zahavi and self-representationalists agree—according to which it is a primitive, *sui generis*, non-relational property of consciousness. I argue that some Brentano-style, self-representationalist theories of consciousness are not subject to Zahavi’s criticisms. I articulate a version according to which consciousness involves self-acquaintance. This allows one to account for primitive self-knowledge and still maintain that ubiquitous, prereflective self-consciousness has a relational structure. I also unearth the premise upon which the regress objection depends and show that no self-representationalist need be committed to it. I end by discussing the kinds of considerations that might allow one to decide between the two theories and the prospects for “naturalizing” them.

I. Introduction

Brentano held that every conscious act aims at some object other than itself, its “primary” object, *and* at itself, its “secondary” object.¹ Any theory positing that all conscious episodes represent or are acquainted with themselves I’ll call a B-theory.² Zahavi prefers the view that all episodes of consciousness are self-conscious but not their own *objects*.

This mode of self-consciousness is *sui generis* and doesn't involve self-directed intentionality.³ I'll call any such theory a Z-theory.

On B-theories, ubiquitous prereflective self-consciousness is a matter of an episode⁴ bearing the same kind of relation to itself that it bears to its primary objects. Minimally, only one episode, one relation, and one primary object are required. On Z-theories, ubiquitous prereflective self-consciousness is a matter of an episode exhibiting the *sui generis* property of being non-intentionally self-conscious, while, of course, it is also intentionally directed at some object other than itself. Minimally: one episode, the *sui generis* property, one intentional relation, and one intentional object. B-theorists arguably work with less. If the two theories are phenomenologically equivalent, then, *ceteris paribus*, B-theory is preferable. I address the objections Zahavi raises against B-theories and return to the question of advantage.

2. Gurwitsch's Abandoned Regress⁵

Gurwitsch's regress follows from the assumption of reflexive representation only if one *also* assumes that (1) the episode represents *all* of its representational properties,⁶ and that (2) no episode can be *identified* with such a property.⁷ Let's accept (2).

Suppose an episode (S) represents itself. It will thus have the *property of representing itself* (=R). By (1), it represents *all* its representational properties, so it will represent R as well as S. In virtue of representing R it will have the *property of representing its property of representing itself* (=R*). By (1) it represents R* too. But now it has the *property of representing R** (=R**). And we're off! We get not an infinite ladder of distinct conscious states but one episode representing an infinity of representational properties.

But why couldn't a self-representing episode represent only *some* of its representational properties? There seems to be no good reason for thinking that it would have to represent them *all*. As a physical system, consciousness is under capacity constraints and limited in representational power. Premise (1) seems very doubtful.

Gurwitsch's regress is *not* a refutation. Bracketing some perhaps stranger possibilities, it just implies that the representational powers of consciousness are limited.⁸ This seems independently plausible but does mean that consciousness has some unrepresented representational properties.⁹

3. Consciousness and "Object"

On Z-theory, an episode of consciousness, though self-conscious, is not its own *object*. Given Zahavi's notion of objecthood, this is plausible. Episodes of consciousness are not given to themselves the way "external" objects are. Primitive self-awareness is importantly different.¹⁰

'Object,' in Zahavi's sense is primarily (though not exclusively) intended to pick out ordinary mundane objects. Such objects reveal themselves diachronically through a series of profiles. Consciousness doesn't reveal itself via profiles. You can't flip an episode of consciousness in order to see its bottom the way you can stones. Objects thus exhibit a kind of transcendence episodes of consciousness don't. Objects are relatively stable and "infinitely plunderable." My piano remains stable relative to my streaming

piano-revealing experiences. I can open the piano, look underneath it—all to multiply piano profiles. But I cannot do *that* with an episode of consciousness. But does B-theory imply that consciousness is always its own object *in this sense*?

There are modes of self-awareness in which one can be aware of oneself as object in this sense: seeing oneself in a mirror, theorizing about consciousness, and, possibly, watching a computer image of what one's brain is doing while one is watching what one's brain is doing. But when a B-theorist says that consciousness is its own object, it can't mean that consciousness is self-aware in these ways.

4. Shoemaker and Self-Knowledge

Here is a version of Shoemaker's famous argument.¹¹ I could not, for *all* properties P, know that I have P by inference from a known identification of myself with an object O that is P. In order to know that I=O, I need to know that O has a property—by hypothesis distinct from P—*only* I have. I must *already* know that I have that criterial property. But if for every property P that I know myself to have by inference from such an identification, there is some distinct property P* that I must have known myself to have antecedently, and if all such knowledge of my properties depends on knowledge of such identifications, then in knowing myself to have even one property, I must already know myself to have infinitely many. So there must be a regress-stopping property we non-inferentially know ourselves to have.

Zahavi believes this supports Z-theory. Primitive, non-objectual self-consciousness satisfies the need for primitive, non-objectual self-knowledge. If we suppose that primitive self-consciousness is epistemic, then it can ground other sorts of "objectual" self-knowledge. Could primitive self-consciousness serve as a basis for primitive self-knowledge on a B-theoretic construal? Zahavi is right that it could not if we consider the sort of representation involved to be the same sort that is involved in the representation of objects (as understood above). Suppose it is like this. A strange possibility opens up. Your primitive self-consciousness could be phenomenologically indistinguishable from, say, the visual perception of two balls on a table. One of the balls is just an ordinary ball and the other, strangely enough, is your consciousness itself. You might have no idea of this identity and would only be able to learn of it, *per impossibile*, by inference. We can see here that the characterization of ubiquitous self-consciousness as marginal or secondary or inattentive, while not false, is insufficient.

There is a looser sense of 'object' in which if x represents (is acquainted with, etc.) y, then y is an object for x—no matter what y is and no matter the species of awareness. *Pace* Zahavi, this does not empty the term 'object' of meaning. It is just a very abstract sense.¹² The fact that the term excludes *nothing* is the very point of the term. After all, is there *anything* we cannot, in some way or other, think about? If there is some sort of awareness relation that holds between an episode of consciousness and itself, then, in this wider sense of 'object,' an episode of consciousness will be its own object. Further, I submit, the episode need not be transcendent with respect to itself if the relation is of just the right sort. Zahavi seems to maintain that there is no kind of awareness relation that an episode could bear to itself that would not make the episode a transcendent object with respect to itself—which, we agree, would have the absurd

consequences just noted. Zahavi thus argues that primitive ubiquitous self-consciousness must be *non-relational*.

But the argument from the impossibility of self-transcendence (at this level) to the non-relational nature of primitive self-consciousness fails. There is a kind of awareness relation that would not imply self-transcendence, is epistemic, and is compatible with the phenomenological fact that primitive self-consciousness is inattentive (marginal or secondary).¹³

5. Acquaintance

The relation I have in mind is what Russell called *acquaintance*. It's not a straightforward representation relation. Importantly, if x is acquainted with y, then y exists. This does not hold for representation generally. It holds, trivially, for all cases of self-representation (whatever the variety). But the mere representation of an x such that if it's represented it exists doesn't really capture *presence*. There are theories of representation on which something could represent itself without being present to itself. The satisfaction of the existence constraint in such cases just stems from the fact that the thing has a property (assuming that only existents really have properties). That property just *happens to be* the property of representing. And the thing it represents just *happens to be* itself. The fact that it exists and the fact that it represents itself scarcely suffice to capture the notion of "being before the mind" or presence.

The acquaintance relation *assumes* presence.¹⁴ If x is acquainted with y, then y is *present to x*. *Presence to* is taken as *phenomenologically* primitive.¹⁵ Whether it is *ontologically* primitive is a question returned to below.

If episodes of consciousness are self-acquainted, they are primitively and epistemically self-aware. So self-acquaintance could base derivative forms of self-knowledge. Unlike other forms of self-representation, acquaintance provides the right sort of intimacy. The episode couldn't just *happen* to bear the relation to itself. But there are worries.

Paradigmatic examples of acquaintance feature sensory qualities. Given the facts of misperception and hallucination and given the existence constraint, an acquaintance theorist should reject direct realism and intentionalism about sensory qualities.¹⁶ The acquaintance theorist should hold: (A) There are layers to perceptual experience—an acquaintance layer and a super-imposed representational layer; (B) sensory qualities, as objects of immediate acquaintance, are properties of one's consciousness or brain. (A) allows one to capture the direct realist *flavor* of experience, the facts of misperception and hallucination, perceptual presence, and the existence constraint simultaneously. (B) is the price.¹⁷

But if an episode of consciousness is acquainted with, say, sensory qualities *and* itself, don't we get a new version of Shoemaker's problem? Is my consciousness just one object of acquaintance among others? Given that sensory qualities do exhibit *some sort* of transcendence with respect to consciousness, by what right do we claim that one object of acquaintance is immanent and the others transcendent? How could consciousness primitively distinguish itself from what is, apparently, not consciousness? If we thought that this relation holds *only* between consciousness and itself, we'd have a *relational Z-*

theory, one that posits a special reflexive *relation*, rather than a special monadic property, that only consciousness exhibits. Hardly an advance.

A possible solution: Treat sensory qualities like parts of an episode and the episode like their whole. By hypothesis, the whole bears the acquaintance relation to all of its parts *and* to itself. If there is always some intrinsic feature that distinguishes it from all its other objects of acquaintance, the problem is solved. It has such a feature. The whole is intrinsically different from its proper parts. Its self-acquaintance involves the simultaneous acquaintance with its parts.¹⁸ And the *difference* between the parts and the whole is manifest via acquaintance. This is similar to being acquainted with the difference between a sound and a taste or two different sounds. Put abstractly, if x is acquainted with y and z at t , and $y \neq z$, then x is acquainted with the *difference* between y and z (or the *fact* that they are different). Such acquaintance with difference does not require but bases judgments of difference. And it is easy to see that if an episode is self-acquainted and acquainted with its parts (sensory qualities, by hypothesis), it will be acquainted with its *difference* from them. In this way, sensory qualities preserve their long-puzzling aspect: they are transcendent in that they are *not* the episode of consciousness; yet they belong to the episode that comprises them.

Thus the relation involved in the self-acquaintance of an episode of consciousness need not be thought different from that between the episode and other objects of acquaintance, though it is, strictly speaking, different from object-representation (in Zahavi's sense of "object"). The relation is epistemic, and the difference between the relata is immediately known. Just as one knows by acquaintance a difference between two sensory qualities, one knows by acquaintance the difference between the episode and *all* such qualities. Z-theory and this version of B-theory both *posit* immediate self-knowledge. But the latter has the modest advantage of employing only one relation, not a relation *and* a special monadic property. If intentionality can be constructed on the basis of acquaintance, then this B-theory seems preferable.

VI. Conclusion

Everyone concerned agrees that ubiquitous primitive self-awareness isn't attentive or reflective, is importantly *unlike* naïve perceptual consciousness of ordinary objects, importantly connected to reflective capacities, temporality, the ego, the mastery of indexicals, primitive self-knowledge, and the unity of consciousness. I thus suspect that the only way to decide between the two accounts is at the level of theory not phenomenology.

Perhaps this B-theory is here in a better position. Its relational account of primitive self-awareness suggests that we might gain by developing an abstract theory of the relation.¹⁹ Treating primitive self-consciousness as non-relational precludes this, though one can, as Zahavi does elsewhere, trace out the role the primitive property plays in the overall phenomenological "economy."²⁰ Does the relational construal allow one deeper insight into that role? We'll have to see.

What about naturalization? Both theories entail the rejection of the attempt to build consciousness out of representation understood in causal covariance or teleosemantic terms. An acquaintance relation surely seems unlikely to be understood this

way. But even if we take acquaintance as, in some sense, *ontologically* primitive, one can still hold that it is instantiated in the brain; one can even hold that it is, ultimately, entirely a matter of structure. And perhaps the relation can be identified *a posteriori* with a relation or structure describable in neuroscientifically relevant terms.

We pick out concrete relations via our grasp of the relata. And, by abstraction, we can often figure out if a relation is transitive, symmetric, wellfounded, etc. The lesson here is that we should stop worrying about whether the relation is mysterious—in a sense, all relations are mysterious. Instead, we should develop an abstract theory of the relation drawing upon phenomenological and empirical considerations and then attempt to isolate the relata neuroscientifically. If those relata conform to our more abstract theory of the relation (and other constraints) then in isolating them we will have isolated the *relation* in the only (non-phenomenological) way possible.

The modified B-theory perhaps stands a better chance of fitting into this program simply because it begins with more structure and might connect with some existing mathematical theories.²¹ It could, of course, be a dead end, and both sorts of theory should be pursued. But what is beginning to become clear, I hope, is that a theory of consciousness that does not take ubiquitous primitive self-awareness into account is a non-starter. On this, I think, Zahavi and I wholeheartedly agree.²²

References

- Brentano, Franz. 1995. *Psychology from an Empirical Standpoint*. London and New York: Routledge. (First edition originally published in 1874, second edition in 1924.)
- Caston, Victor. 2002. "Aristotle on Consciousness." *Mind*, 111, 444, 751-815.
- Gurwitsch, Aron. 1964. *The Field of Consciousness*. Pittsburgh: Duquesne University Press.
- _____. 1979. *Human Encounters in the Social World*. Pittsburgh: Duquesne University Press. (Originally published in 1977.)
- _____. 1985. *Marginal Consciousness*. Athens, Ohio: Ohio University Press.
- _____. 2002. *Esquisse de la phénoménologie constitutive*. Paris: J. Vrin.
- Kriegel, Uriah. 2003. "Consciousness as Intransitive Self-Consciousness: Two Views and an Argument." *Canadian Journal of Philosophy*, 33, 1, 103-132.
- Williford, Kenneth. 2005. "The Intentionality of Consciousness and Consciousness of Intentionality." In G. Forrai and G. Kampis (eds.) *Intentionality: Past and Future*, Amsterdam: Rodopi.
- _____. 2006. "The Self-Representational Structure of Consciousness." In U. Kriegel and K. Williford (eds.) *Self-Representational Approaches to Consciousness*, Cambridge: The MIT Press.
- Zahavi, Dan. 1999. *Self-Awareness and Alterity*. Evanston: Northwestern University Press.

Notes

1. Brentano, 1995, pp. 101-154.
2. E.g., Kriegel 2003, Williford 2005, 2006, Caston 2002.
3. See, esp., Zahavi 1999.
4. I prefer 'episode' to 'state' and 'act.' On my theory, the episode, properly individuated, is the transient *phenomenological* subject of consciousness. This explains some peculiarities of expression.
5. "Abandoned" because Gurwitsch apparently came to think that this argument fails. The habilitation essay in which he made the argument was completed toward the end of 1931 but not published until 1977, four years after Gurwitsch's death. Gurwitsch had agreed to its publication "...[a]fter thinking it over for some time...[and] not without reservations" (1979, p. xi). Curiously in his *Esquisse de la phénoménologie constitutive* composed between 1933 and 1940, Gurwitsch referred to Brentano's doctrine favorably without mentioning this regress (Gurwitsch 2002, pp. 240-241). And in *Marginal Consciousness*, written circa 1950 and published in 1985, Gurwitsch effectively adopts Brentano's theory as his own, changing terminology only (1985, pp. 3-13). Gurwitsch even says, "...[W]e are persuaded to abide by it [Brentano's doctrine] in the face of these [James's and Husserl's] reservations, which are not refutations" (1985, p. 22). For recent discussions of this regress see Caston 2002, pp. 797-798, and Kriegel 2003, pp. 125-126 and my 2006.
6. Brentano says things that suggest this premise, see, e.g., 1995, pp. 154. Other remarks suggest the "limited representational powers" view; see 1995, p. 277. *Strictly speaking*, we could restrict this to self-representational properties alone.
7. It's actually more complicated. See my 2006.
8. One can also avoid the regress *and* even accept (1) either by rejecting (2) or by supposing that the episode represents itself by quantifying over all its representational properties. The property of so quantifying is itself representational and would belong to the domain. It's analogous to a proposition about *all* propositions. It would be a representational property in virtue of which consciousness represents all its representational properties.
9. I discuss this in my 2006. This could be problematic for some versions of B-theory.
10. I use 'awareness' and 'consciousness' interchangeably.
11. Shoemaker 1968.
12. Cf. Gurwitsch 1964, p. 158, "The term 'object' is here used in the most inclusive sense so as to denote whatever may be a topic of discourse, thinking, or apprehension of any kind whatsoever."
13. I prefer 'marginal' 'peripheral' and 'inattentive' to terms like 'tacit,' 'implicit,' and 'intransitive.' The former pick out a central aspect of the phenomenology of primitive self-consciousness that everyone should accept and are otherwise neutral. 'Tacit,' and 'implicit,' are apt to suggest that prereflective self-consciousness is dispositional—it's not. 'Intransitive,' if we take the grammatical analogy seriously, is inappropriate to B-theories—prereflective self-consciousness does "take an object."

14. Tying intrinsic intentionality to consciousness might be another (perhaps equivalent) way to capture presence.
15. Talk of “presence” might seem obscure. The term is meant to pick out a phenomenologically primitive notion. All primitive notions are obscure in the precise sense that they cannot be understood in still more primitive terms. But, phenomenologically, having something before your conscious mind via perception is familiar. And replacing talk of ‘presence to’ with talk of ‘acquaintance with’ might highlight the claim that presence, including self-presence, is relational. It perhaps better suggests the central claim about the *categorical* status of presence.
16. By ‘direct realism’ I mean views according to which we immediately perceive physical ordinary objects (or their surfaces) and their properties. By ‘intentionalism’ I mean any view according to which sensory qualities are *fundamentally* just special objects of representation (and thus might not exist even when apparently perceived).
17. Some of us are willing to pay it.
18. The *property* of being so acquainted need not itself be considered a part.
19. I make some steps in my 2006.
20. See his excellent 1999.
21. See my 2006.
22. I’d like to thank Uriah Kriegel for helpful comments on an earlier version.