Review of Dan Zahavi's *Self and Other: exploring subjectivity, empathy, and shame* (Oxford, 2015)

Philip J. Walsh

Dan Zahavi's newest book, *Self and Other*, gathers together and expands on his rich body of work on selfhood and intersubjectivity. The book is divided into three parts: the experiential self; empathic understanding; and the interpersonal self. A consistent theme throughout is the relevance of classical phenomenology to contemporary debates about consciousness, selfhood, and social cognition. Zahavi's discussions of Husserl, Stein, Scheler, Schutz, Sartre, and Merleau-Ponty (among others) are extremely clear and careful. The exegetical work alone makes this book valuable for anyone interested in the phenomenological tradition and its contemporary relevance. The book's value, however, exceeds its historical acumen by placing explanations of classical phenomenology alongside a comprehensive review of recent work from both empirical psychology and analytic philosophy. In short, this book will be of interest to anyone who cares about the myriad ways that our social existence shapes and is shaped by the mind.

Zahavi begins the book with a justification of the need for philosophy. Concepts like 'self' and 'self-awareness' employed by empirical researchers are too frequently muddled and confused. The role of the philosopher is to clarify such concepts so that empirical (or, better yet, interdisciplinary) research may proceed. A brief survey of the conceptual morass found in characterizations of autism is more than convincing that Zahavi is correct. This theme is repeated throughout the book. We cannot accept "findings" on empathy or shame if those findings use the concepts inconsistently.

The chapters that comprise the first part of the book argue, along Husserlian and Sartrean lines, for a form of pre-reflective self-awareness that is an essential feature of consciousness. This feature is characterized as the "mineness" or "for-me-ness" implicit in all forms of experience (19). Importantly, this is a *phenomenological* feature of experience, and not (as Searle and Prinz argue) a mere formal or conceptual feature of experience. Nor should it be understood along representationalist lines, as awareness *of* something. Zahavi characterizes his phenomenological proposal in Chapter Two ("Consciousness, Self-Consciousness, and Selfhood") with a thought experiment which, I must admit, is not as immediately intuition-pumping as one might hope. Mick and Mack are phenomenologically type-identical, staring at a blank white wall. Zahavi claims that the difference between Mick's and Mack's experiential lives is *different* than the difference between two type-identical but numerically distinct drops of water. Although their experiences of the white wall are type-identical, "each of these experiences will have it own distinct pre-reflective self-manifestation. For Mick, his experience will be quite unlike Mack's experience (and vice versa)" (24). This claim is supposed to amount to something more than the claim that a full characterization of the overall phenomenal character of their respective experiential states necessarily includes a for-me-ness aspect. But one could coherently claim that both Mick's and Mack's overall phenomenology is type-identical, token-distinct, and includes a distinctive "for-me-ness" character. Why not simply say that "for-me-ness" is a type, tokened in both experiences, just as the look of the white wall (in this particular lighting from this particular position, etc.) is a phenomenal type tokened in both experiences? What distinguishes the for-me-ness character of Mick's experience from the for-me-ness character of Mack's experience other

than the (non-phenomenal) distinctness of their bearers? Though I find the notion of pre-reflective self-awareness intuitive, I found it less clear how it does of the work of making Mick's experience "quite unlike" Mack's. This issue proves important for Zahavi's subsequent move from self-awareness to selfhood.

Zahavi argues that his account of pre-reflective self-awareness entails the existence of a pre-reflective self. In Chapter Three ("Transparency and Anonymity") he successfully defends his account of pre-reflective awareness against two objections. The anonymity objection accuses Zahavi of falling prey to the refrigerator fallacy, i.e. the notion that the light of self-awareness is always on just because it is on every time we take a look. The transparency objection contends that the phenomenology of experience is exhausted by the feature of the *objects* of experience. Zahavi skillfully preserves what these objections get right about phenomenology while showing that they do not threaten his notion of pre-reflective self-awareness. It seems that all of the confusions that Zahavi is dealing with ultimately originate in a conception of awareness as necessarily object-directed. Understanding his account requires that one accept a notion of self-awareness as a constitutive feature of phenomenal consciousness that is reflexively self-intimating without requiring a transitive form.

Chapter Four ("Subjectivity and Selfhood") argues that this notion of pre-reflective self-awareness entails the existence of a pre-reflective self. In other words, if one accepts a certain notion of subjectivity one must also accept a certain (thin, minimal) notion of selfhood. Unlike "no-self" accounts (Albahari 2006), Zahavi argues that experiential self-awareness necessitates the existence of an experiential self. This self is ontologically dependent on but irreducible to the stream of consciousness. It is not a formal or logical feature of the stream, but rather "can be identified with the ubiquitous first-personal character of the experiential phenomena" (18). As discussed above, however, I had trouble grasping how the mineness or for-me-ness experience provides experiential life with inherent individuation, independently of any actually existing relation to others. Indeed, in Chapter Six ("Pure and Poor") Zahavi notes Husserl's observation that "even though our experiential life is inherently individuated, we must realize that it is a formal and empty kind of individuation, one that equally characterizes every other possible subject" (84). This brings me back to my initial confusion with the thought experiment in Chapter Two: how are Mick and Mack distinct in any way other than their being numerically distinct? Shouldn't the "pure and poor" mineness of their respective experiences be type-identical? Wouldn't an overall characterization of their occurrent phenomenologies be identical? Does such a thin notion of selfhood really do any additional individuative work? The real question here, then, is "When is it appropriate to call something a self?" (48). If the experiential self is too minimal and deflationary, then perhaps we should look to "notions of ecological, experiential, dialogical, narrative, relational, embodied, and socially constructed selves" (47).

Narrativist and social-constructionist accounts raise two challenges for Zahavi: (i) that the for-me-ness of subjectivity is simply a necessary prerequisite for selfhood but does not itself constitute a minimal form of selfhood, and (ii) whether the first person character of experience remains invariant through development or whether it is fundamentally transformed through enculturation (61).

Animals, infants, and adults all share Zahavi's pre-reflective self-awareness. This is sufficient for a minimal or core self. One might insist that a self is something more robust and that animals and infants have "subjecthood" but not "selfhood". Nothing hangs on this terminological choice. More interestingly, the second idea that enculturation fundamentally changes our form of self-awareness might seem like a plausible way to explain what we have in common with infants and animals while making the distinction between mere subjecthood and full blown selfhood sharper. In response, Zahavi says that while *what* we experience radically changes in virtue of our enculturation and concept acquisition, *how* we experience things (for-me-ness) must remain constant if we are to even make sense of how infants develop into adults. Otherwise we are left with "unbridgeable dualism between the non-conceptual sentience of the infant and the conceptualized mind of the adult" (62, fn. 8). Zahavi makes an analogy between this transformation and dye in water. Though every bit of the water is saturated with the dye, and thus the effects of the dye are pervasive, the water retains its liquid properties. But one can simply invoke a different metaphor to make the narrativist point: enculturation is not like dye in water, it is like a chemical reaction. The resultant substance has completely new properties. The original substances are fundamentally transformed. Likewise, the narrativist might claim, the for-me-ness of the pre-conceptual infant is fundamentally transformed by enculturation.

Zahavi is walking a fine line here. He reminds us that is is "important to not lose sight of how *formal* a notion of selfhood" he is talking about (62, my emphasis). But, as he repeatedly insists, this notion of self is *not* a mere formal or logical feature of experience, *not* a mere necessary posit (à la Kant, Searle, and Prinz). At times, I found myself conceiving of his minimal selfhood as a sort of empty container: the "field" or "space" of experience in which specific experiential episodes play out. In Chapter Five ("Self and Diachronic Unity"), however, Zahavi insists that his identification of the experiential self with the abiding for-me-ness of experience should not be thought of in this way. The experiential self is not reducible to any specific experiences or some relation holding between experiential episodes, but this "is not, of course, to claim that it has a distinct and independent existence, as if there could first be a pure or empty field of experience" where experiences arise and pass (66). "*Rather, the experiencing simply is the invariant dimension of pre-reflective self-manifestation possessed by each and every experience*" (66, my emphasis).

Thus, it seems that the pre-reflective self depends on the presence of pre-reflective self-awareness. There is no pre-reflective self-awareness during deep sleep. So how, then, does Zahavi handle issues of diachronic unity for such a thin notion of selfhood? Even after periods of deep sleep, Zahavi argues, past experiences are first-personally accessible in recollection, and this obtains in virtue of their partaking in the same invariant dimension of for-me-ness. Though I was never quite sure how this invariant dimension of for-me-ness was supposed to be identified with, yet remain distinct from, a pervasive phenomenal feature of experience, this chapter was a rewarding read as it seamlessly traversed Husserl's theory of inner-time consciousness and contemporary metaphysics of personal identity. Furthermore, though Zahavi argues that his notion of selfhood is thicker than Galen Strawson's SESMETs (2009), he acknowledges that the issue of diachronic unity highlights the need for "A Multidimensional Account" of the self (Chapter 7) that draws on resources from narrativist, ecological, and social-constructionist accounts. If the experiential self is admittedly thin, then we must turn to the social world and our experience of it in order to understand selfhood in its more robust forms.

In part II of the book Zahavi turns his attention to the family of contemporary issues under the umbrella of "social cognition". Importantly, his prior emphasis on the ontological and conceptual independence of the self from the social world should not lead one to conclude that he has an overly Cartesian notion of selfhood, whereby our access to other minds is inherently indirect or inferential. Rather, part II of the text is devoted to developing an account of empathic awareness that is perceptually-based and theoretically unmediated (98). Chapter Ten ("Phenomenology and Empathy") presents a detailed overview and synthesis of the phenomenological tradition's treatment of empathy while Chapter 11 ("Empathy and Social Cognition") engages contemporary research. These chapters are the heart of part II and perhaps the most interesting part of the book. Zahavi traces the phenomenological tradition's rejection of the imitation/projection account of empathy through Scheler, Husserl, Stein, and Schutz, with nods to Gurwitsch and Merleau-Ponty along the way.

Zahavi's summary and evaluation of Scheler's argument against any sort of inference, analogizing, or inner imitation is excellent. I found his examples illustrating the difference between emotional contagion and Scheler's notion of emotional sharing (*Mitfühlen*) especially clear and helpful. Zahavi's defense of Scheler foreshadows his critique of accounts of social cognition that overinflate the explanatory scope of mirror-neuron research. This is an inspiring reminder of the relevance of the phenomenological tradition to contemporary philosophy of mind and cognitive science.

Next, Zahavi explores Husserl's career-long preoccupation with empathy. Restricting his investigation to Husserl's more explicit discussions of empathy (rather than his broader ranging work on intersubjectivity), Zahavi carefully explains the similarities and differences between empathy and perception. While empathy does not present us with the mental states of others with the same degree of fullness or originality that perception presents us with ordinary objects, we must be careful not to conflate this with the directness/indirectness distinction. Empathic experience is directed at *expressive* intentional objects. The bodily movement, facial expressions, and intonation of the other are imbued with psychological meaning from the start. A key difference, however, between empathy and perception is that in perception the profiles of the object that are not given in intuitive fullness nonetheless exist as being possibly brought to fullness. When I see an object its facing side is given in fullness whereas its back and unseen sides are "apperceived". The apperceptive content of empathy, on the other hand, can never be brought to the same kind of first-personal givenness found in the other's own awareness of her mental states. Zahavi's discussion is especially helpful here as he dissolves the apparent tensions in Husserl's account. The moral of the story is that different forms of acquaintance with a mental state (first-personal vs. empathic) do not map onto a directness/indirectness distinction. Yes, the other's mental state will never be given to me the way it is given to him. But this does not mean that it is given to him "directly" whereas given to me "indirectly".

The chapter also includes a lengthy analysis of the role of embodiment in empathic awareness. In several places Husserl identifies one's own first-personal bodily self-awareness as a necessary precondition for other awareness. It is in virtue of my own bodily self-awareness that the body of the other shows up as an expressive phenomenon. Husserl also seems to question this account, however, and worries that it may fall back into a kind of imitation/projection model of empathy.

After all, if one's first-personal bodily self-awareness serves as a sort of inner model for understanding the body of the other as an expressive body like one's own, then it seems that empathic awareness of others is a matter of projecting my own first-personal bodily awareness onto the bodily movements of the other.

Zahavi's proposed way out is "to see the self-experience in question as a *necessary foil* on the basis of which others can be experienced as others" (136, my emphasis). In other words, "my self-experience doesn't constitute the model or matrix; rather, it is simply that against which the other's difference and transcendence can reveal itself" (137). But one might ask, if we are to preserve the constitutive asymmetry between first-personal bodily awareness and one's awareness of the body of the other, on what basis would my first-personal bodily awareness serve as the "foil" against which other bodies show up as other *expressive* bodies? A promising insight comes from Husserl, Stein, and Merleau-Ponty: there is always-already an intersubjectivity of the body. One is aware of one's body in two ways: as that which *experiences*, i.e. as a volitional structure and dimension of sensing; and as that which is *experienced*, i.e. as an object given in visual and tactile experience. Thus, self-awareness consists of a subtle interplay between *ipseity* and *alterity* (135).

This insight is indeed fascinating, however, as I have argued elsewhere, I am not sure it solves the problem (Walsh 2014). The interplay between ipseity and alterity in auto-affection establishes an order of functional correlations between one's self as *that which does the experiencing* (inner volitional structure and dimension of sensing) and one's self as *that which is experienced* (one's body as it appears visually and tactilely for one). For example, as I run my right hand over my left forearm, the visual and tactile awareness of my hand on my arm correlates in a regular and systematic with the proprioceptive and kinesthetic awareness that constitutes my feeling of moving and controlling my hand. But in empathic experience this same order of functional correlation could not be in place. The visual information presenting the moving body of the other (that which is experienc*ed*) does not correlate in the same way with my inner bodily awareness (that which experienc*es*), since in this case I am standing still, turning my head and moving my eyes to track the movement of the other. If the same order of functional correlations were operative in the auto-affection case as in the empathic awareness case, then we are back in the realm of resonance or simulation. In other words, even though there may be a two-sided form of first-personal awareness of one's own body, this still does not explain why or how the appearance of the other's body would be automatically integrated into this system.

This point aside, Zahavi's ability to succinctly organize and explain the vast corpus of phenomenological research on empathy is superb. As Zahavi brings phenomenological insight to bear on contemporary social cognition debates in Chapter 11 ("Empathy and Social Cognition"), we see that the two fields are mutually enriching. Our awareness of others is multifaceted, and understanding classical phenomenology's notion of empathy can help us sort out what contemporary accounts of resonance, mirroring, and mind-reading are really explaining. Furthermore, attention to these issues in social cognition paves the way for part III's analysis of how sociality enriches selfhood.

The highlight of part III is Chapter 14 ("Shame"), which ties together the threads of subjectivity and sociality woven in previous chapters. Shame is a complex emotion that essentially self-

refers, but shame is also essentially social in some way. Without insisting on a strictly univocal concept of shame, Zahavi critiques accounts of shame that presuppose an overly sophisticated level of self-awareness as well as those that presuppose an overly sophisticated level of socially-mediated normative commitment and enculturation. Zahavi's account draws heavily on Sartre. In shame, I recognize myself as exposed and vulnerable before the other, regardless of whether an actual other is physically present and observing me. I am not merely aware of how others might view me: "I *am* the way the other sees me and I am nothing but that" (213). I am powerless in this realization precisely because I am never able to fully adopt this objectifying view despite my knowledge that I am subject to it.

Shame is a kind of self-consciousness that is intersubjectively mediated (212). It involves a global decrease in self-esteem (226). Unlike embarrassment and humiliation, however, shame involves a self-imposed negative evaluation. I may feel humiliated before others who mock and degrade me without actually accepting and internalizing their negative evaluation. But in allowing myself to be humiliated and accepting it (perhaps I fear physical punishment), I may subsequently feel ashamed because I assess myself as "exemplifying a cowardice that I despise" (227, fn. 14). This kind of experience is an experience of myself, but one that is "thrust" upon me, thus forcing me to confront my own vulnerability and exposure (222).

This chapter nicely complements the first half of the book because it further clarifies Zahavi's minimalist notion of the self. This thin notion of self is logically prior to the shamed-self. The shamed-self is necessarily a more complex self because it makes essential reference to the social world. The shamed-self is not, however, the more robust socially-constructed narrative self. Rather, the shamed-self is the bridge linking the minimal experiential self to the socially constructed narrative self. The self that is capable of shame is the "interpersonal self" (238). It is socially constituted but "precedes the learning of particular social standards" (238). Here is where the prior discussion of empathy ties in. Empathy is a basic, pre-linguistic capacity to recognize others as others. This experience of others turns out to be transformative: through an iterative process of empathy, I simultaneously come to understand myself as there for others just as they are there for me. Zahavi recruits empirical work on joint attention to flesh this out and explain how it is not an overly cognitively demanding model. In coming to recognize others, infants also recognize themselves as the objects of joint attention. In short, this research supports the idea that empathy is a basic capacity that involves not only recognition, but interaction.

The brief final chapter ("You, Me, and We") explores an exciting new area of recent phenomenological research: plural subject theory. When one ascribes a we-intentional experience, e.g. "We saw the hedgehog", who or what "had" this experience? Zahavi argues that we-intentions do not have a single owner, a view he arrives at through an extended discussion of David Carr's (1986) account. We-intentionality depends on individual minds, but cannot be reduced to a collection or aggregate of I-intentions. Furthermore, we-experiences presuppose "some kind of identification/sharing with others," and the notion of sharing in play here involves a plurality of subjects that are integrated in a unique way (244-245). In other words, we-intentionality is not reducible to what is going on in the minds of the constituent members of the we because the mental activity of the members is uniquely conditioned by the super-personal ontological structure of which they are a part.

Zahavi's example of watching a movie together with a friend is helpful. I find that I rarely re-watch a film. But if I learn that a friend has not seen it, and it is a film I like, I find myself wanting to watch it again *with* my friend. I am not simply enjoying the film as I did the first time I saw it. My enjoyment is founded on a different kind of experience. *We* are watching the film. We are sharing in something. Not only do we enjoy the film, but we also empathically enjoy one another's enjoyment of the film. Our individual experiences are "co-regulated and constitutively bound together" (245).

A virtue of this final chapter, as it ties together themes from the entire book, is that it is a compelling case study in what we might call socialized phenomenological philosophy of mind. We ought to reject overly solipsistic or methodologically individualistic accounts of mind and self, for they make social phenomena overly problematic. That said, we can retain core phenomenological insights into subjectivity without effacing the enormously transformative significance of sociality.

References

Albahari, Miri. 2006. *Analytical Buddhism: The Two-Tiered Illusion of Self*. Palgrave
        Macmillan.
Strawson, Galen. 2009. *Selves: An Essay in Revisionary Metaphysics*. Oxford University Press
        Inc.
Walsh, Philip J. 2014. "Empathy, Embodiment, and the Unity of Expression." *Topoi* 33 (1): 215–
        26.