

Essays in Philosophy

A Biannual Journal

Vol. 6, No. 1, January 2005

Book Review

Being No-one: The Self-Model Theory of Subjectivity, by Thomas Metzinger, MIT Press: Cambridge MA, 2003, 713 pages, 14 illustrations, bibliography, ISBN: 0262134179, A Bradford Book, \$55.

“the conscious self is an illusion which is *no one's* illusion” Thomas Metzinger.

Ever since Rene Descartes proposed his *cogito ergo sum*, we were given a rather robust defense of the self as a distinct, non-physical, substance that does in fact think. If there is one thing of which we can be certain, according to Descartes -- I am. According to Thomas Metzinger, however, perhaps even this apparently rather self-evident claim, besides from the substance dualism, should be doubted.

My claim is that – ontologically speaking – no such things as selves exist in the world. What actually exists is a special kind of self-models and their contents, and this content makes us believe that we actually do have, or are identical to, a self. (626)

As argued by Metzinger, we ought to remove the ‘I’ here and perhaps leave ourselves only with the more modest claim as follows: a ‘think’ exists. Now this ‘think’ or mental state with phenomenal and/or intentional content may give us a world model; for example, a thought with the content of a tree. But the content of the thought might contain a self-model as part of the world model; such as, a thought about general properties of the body; for example a mental state with the content of where the parts of this body are located in the world. Finally, and this is crucial for Metzinger, the thought might be about itself. This thought is just another representational state with content, but the content is such that the thought and its content and the relation between the thought and its content is itself the content of a representational state. The thought acknowledges the existence of the thought and its relations to the world and itself. For Metzinger this thought ‘makes us believe’ the content that there is some *one* that is having the thought - Descartes’ rather natural conclusion; but according to Metzinger, this content that a self exists and is in fact the entity that has that thought is an illusion. All that exists is the world model, the self model and the state that has its own existence and the mental states relationship to its content as its content. Descartes made an inference too far when he concluded the robust existence of a self.

Post (and pre) Descartes, there is a long tradition of denying the existence of the self: for example, Gilbert Ryle suggests that the self itself is never itself the content of attention:

It is not supposed that when I am wondering, say, what is the answer to a puzzle and am *ipso facto* consciously doing so, that I am synchronously performing two acts of attention, one to the puzzle and the other to my wondering about it. Nor, to generalize this point, is it supposed that my act of wondering and its self-intimation to me are two distinct acts or processes

indissolubly welded together. Rather, to relapse perforce into simile, it is supposed that mental processes are phosphorescent, like tropical sea-water, which makes itself visible by the light which it itself emits. Or to use another simile, mental processes are 'overheard' by the mind whose processes they are, somewhat as a speaker overhears the words he is himself uttering. (158-9) Ryle, G. (1949). *The Concept of Mind*. New York: Barnes and Noble.

However, when anyone (including Daniel Dennett as in *Consciousness Explained*, Boston, MA: Little, Brown & Co., 1991) posits a virtual or an illusory self, I am always tempted to ask how real is the illusion? However, I don't think that the question as to the reality of an illusion is illusory or poorly formed because if the self is going to do any explanatory or causal work, there is surely going to have to be an answer to this question. Furthermore, what precisely is the relationship between the phenomenal and mental content?

The illusory self may in fact be a very useful illusion because as soon as I have an 'I', I should perhaps care more about prolonging this 'I'. When an individual dies, does the self – cease, or does the mere illusion of self – de cease? Metzinger hints at some positive ethical implications of this illusory nature of self. He briefly discusses ethics in the context of suffering (p.570). He argues that as soon as there is first stage phenomenal suffering with a weak notion of phenomenal ownership through the mental content of a world model that includes the body, erring on the safe side, we should at least consider the entity as worthy of moral consideration as a moral object. At the end of his book Metzinger rather dramatically suggests an ethical problem with creating artificial entities with partial stages of consciousness comparing them to the creation of mentally handicapped infants. However, Metzinger thinks that the possession of a full blown illusion of a self, despite a self being something that does not exist, is sufficient for the owners possession of an illusory self's treatment as a moral subject. He is trying to derive an 'ought' from an 'is' here, but even more curiously – from an illusory 'is'! He draws sharp distinctions here that appear contrary to his earlier general method of reasoning, but perhaps not quite as sharp as Immanuel Kant's distinction between things and persons. Although for Metzinger, once one has an illusory self, one is able to grasp relations to others and their illusory selves thus creating a robust notion of personhood. But if Metzinger is right, is personhood not as illusory as the self?

Metzinger's ontology of the phenomenal is more about process than substance. Alfred Whitehead's somewhat impenetrable but speculatively rich philosophy comes to mind as another attempt to solve philosophical problems through the assuming of a process and event ontology rather than object and property ontology. One justification for adopting this approach, for Metzinger, is to make the evolutionary development of phenomenal consciousness more plausible. Rather than consciousness being an 'either/or' affair it is more a matter of degree or progression through the various stages of content and the phenomenal that he outlines culminating in the illusory self. He rather grandly talks of dynamic processes of interactions between various stages of the phenomenal at different stages of development.

It is perhaps worth noting that process can perhaps be seen developed much earlier in Hume's discussion of identity, including personal identity, where identity is bound up in the mental construction of constant conjunctions of experiences and events; furthermore, these conjunctions are given a temporal dimension by the faculty of memory. For Hume, an important conclusion is that the self is a fiction discovered by our memories.

But does a memory have to be a memory to someone; must a representation be presented to someone or some 'thing'? Metzinger's reply is that 'presence' is sufficient for mental content and thus mental content

requires no-one or no thing to whom the representation is presented. However, clever analytic development of notions such as ‘presence’ and ‘presentations’ alone do not solve philosophical problems. But Metzinger goes beyond the analytic and argues that a wealth of empirical data supports his conclusion and in a highly informative section of the book discusses a series of ‘neurophenomenological’ case studies.

However, it is not entirely made clear how the ‘empirical’ data Metzinger offers either falsifies all competing theories of mind, or verifies his specific conclusions about consciousness and the self. For example when trying to use the example of ‘color blindsight’ to demonstrate what he consider to be an essential distinction between intentional and phenomenal content Metzinger argues thus:

Color experience, it turns out, can be selectively lost. In cerebral achromatopsia we find a selective vanishing of colors from the conscious model of the world (Zeki 1990). We have a situation in which localized brain lesions specifically destroy the conscious perception of color *without* necessarily being accompanied by object agnosia. However, here and in many other cases, the situation soon gets complicated. Cowey and Heywood report a phenomenally colorblind patient, M.S., who can still detect chromatic borders, perceive shape from color, and discriminate the direction in which a striped pattern moves when the determination of direction requires the viewer to “know” which stripes have a particular color (Cowey and Heywood 1997). Such cases of color blindness, despite an extended remaining capacity for wavelength processing, not only show, again, how a basic category of phenomenal content can selectively be deleted from the conscious model of the world, they also seem to demonstrate to how great a degree *intentional* and *phenomenal* content active in the system can be separated. . . . We seem to be presented with the possibility of wavelength-related information being cognitively available, namely, in terms of becoming the possible contents of judgements about perceptually uncolored events. However this information is not *attentionally* available. . . . It seems safe to say that color has disappeared from the phenomenal model of reality for this patient, however, the respective information obviously is available for the formation of judgments, for inferences, and for control of speech acts. (220-1)

But does the above example really demonstrate a distinction between intentional and phenomenal content? Why is it necessary that the perceptual processes that involve border recognition require all the cognitive processes necessary for detecting blocks of color? Why is it necessary that the *relevant* wavelength related material remains cognitively, but non-attentionally available in the above cases of color blindness? Edges and borders clearly do play a prominent role in our perception, but perhaps not because of whatever complex wavelength information, reflectance properties and processing that one might think is required when one perceives a block of color. Instead, perhaps edge and border detection requires only a cognitive and attentional detection of wavelength (or whatever more complex properties are relevant) *differences* between the one side of a surface and the other. This explains how one could still detect edges whilst losing the *cognitive* processing abilities required for experiencing and identifying blocks of color.

But perhaps more importantly, how does Metzinger get to even the first stage of the phenomenal? Perhaps the ‘magic’ occurs at this point. Metzinger leans heavily on notions of ‘representation’ and ‘content.’ Talk of ‘representation’ and ‘content’ is very much in vogue at the moment in contemporary philosophy of mind; but these are terms that are often wielded like sledgehammers, without any real clarity as to what the terms themselves refer or mean. Fred Dretske is perhaps an exception to the norm

in that he clearly was and is aware that he is invoking concepts that require some serious development and careful analysis. Metzinger follows, perhaps too hastily, in the tradition of most contemporary rhetoric by invoking ‘representations’ and ‘contents’ of various types perhaps hoping that they will answer the deeper problems posed whilst paying the briefest homage to the philosophical history. His specific theory of representation he claims is teleofunctional; David Papineau, Ruth Millikan, Daniel Dennett and lately Michael Tye offer good (but by no means exhaustive) exemplars of this general approach.

. . . I have chosen to make teleofunctionalism one of the background assumptions in this book. . . . The adaptivity constraint demands that all the functional properties I have sketched when discussing those constraints on the respective level of description always have to possess a plausible reading as gradually acquired properties. That is, additional teleofunctionalist constraints must hold on this level of analysis. We then have to assume that such states do not simply have to be individuated by their causal role alone but that one has to introduce the general assumption that they always play this role for the system: they help the organism in pursuing its individual goals—or those of its ancestors. . . . Phenomenal representata are *good* representata if, and only if they successfully and reliably depict those causal properties of the interaction domain of an organism that were important for reproductive success. (203)

Considerable care has to be taken here. It is one thing to say evolutionary history is important in determining content and another to posit some kind of necessary co-extensive relation between the properties of a state that includes its history and its content. When Metzinger claims such states are *individuated* beyond their causal role to include the history of the formation of that causal role, he might be taken to imply a much stronger relation of necessary co-extension. When later he introduces the normative concept of *good*, he appears to weaken this relation. To clarify, we could have one contentful mental state M1, with intrinsic properties IP1, causal roles C1 with causal history ‘A’; and a *distinct* contentful mental state M2 with identical intrinsic properties IP1, causal roles C1. yet a relevantly different causal history ‘B.’ Yet if we individuate these content states as distinct, do we extend that individuation to distinct phenomenal properties? Metzinger offers an answer of no to this question.

He (more controversially than he acknowledges) maintains a ‘local supervenience’ of the phenomenal on the internal and contemporaneous properties of a neural system.

Of course it remains true that phenomenal content as such only locally supervenes, that is, in a given individual it is always determined by the internal and contemporaneous properties of its neural system. (200)

His distinction between intentional and phenomenal content is surely troubling and perhaps requires a more careful exposition. An early Michael Tye (1994), as in “Qualia content and the inverted spectrum”. *Nous* 28, 159-83, might well have agreed with the above claim, but a later Michael Tye, as in his *Ten Problems of Consciousness*. Cambridge: MIT Press, and Daniel Dennett and many others might be more resistant to accepting this latter proposition. There is an important question here concerning what quite Metzinger means by the term ‘content’ when he talks about ‘phenomenal content’? I suspect he might have in mind a distinction between ‘narrow’ and ‘wide’ content, whereby intentional content is taken to be ‘wide’ content and perhaps phenomenal content is taken to be ‘narrow’, perhaps along the lines of an earlier Michael Tye’s and George Rey’s 1998 in “A narrow representationalist account of qualitative experience.” *Philosophical Perspectives* 12, 435-57. This distinction however is not without difficulty,

see Thompson, *Phenomenism and Representationism: A critique of two approaches to explaining the relation between representational and phenomenal content*. PhD Dissertation: The University of Edinburgh (2002).

Metzinger does revealingly give a response to the Donald Davidson's infamous Swampman scenario (A being that emerges fully formed and spontaneously out of a swamp that is identical to a human in all the intrinsic properties, current causal roles, yet lacks the appropriate history). Metzinger's solution is to once again rely on his distinction between intentional and phenomenal content. The swampman would be conscious in a "weaker" sense given distinct intentional contents, whilst having identical phenomenal contents to the original person (p. 206). This concept of 'weakness' as a mental content that has the wrong sort of history is given further distinctions of degree between a being that spontaneously pops into existence, a being that is made by an appropriately historical correct conscious agent, and the consciousness of the appropriately conscious agent itself.

But what is perhaps most troubling is the apparent contradiction between Metzinger's teleofunctionalism and his apparent defense of the phenomenal locally supervening on the intrinsic physical properties of a system. Metzinger's proposed compatibilism with regards to these two positions is by no means uncontroversial and does I think rely on some unstated but implicit notion of 'narrow' content that is nowhere explicitly defended other than brief references to brains in vats surely experiencing a rich phenomenal existence. There is considerable literature, however, that responds to this intuition and arguably needs to be addressed in more depth by Metzinger.

Metzinger does refreshingly pose a hard and excellent question when he asks – when is a representation a phenomenal representation? I am not entirely satisfied that he successfully answers this question in this book. Perhaps he needs to develop further the equally hard question – what is a 'representation' or a 'content'? I am inclined to think that the answer to the latter question is deeply intertwined with the answer to the former.

All this being said, Metzinger's astonishingly ambitious book is well worth the read as a bold endeavor to attempt to apply a teleofunctional, representationalist, with perhaps an implicit nod to narrow content analysis of the phenomenal to the self with an enjoyable eliminativism with regards to the self itself.

Hamish Thompson
East Tennessee State University