

## Nobody's Problem

A Response to the Ideas Presented in Thomas Metzinger's *Being No One*

Stephen Ware  
Dr. Briscoe  
Philosophy of Mind

14 December 2006

“But immediately upon this I observed that, whilst I thus wished to think that all was false, it was absolutely necessary that I, who thus thought, should be somewhat; and as I observed that this truth, I think, therefore I am, was so certain and of such evidence that no ground of doubt, however extravagant, could be alleged by the skeptics capable of shaking it [...]”

These famous words by Descartes have been disputed by several philosophers, but perhaps never so convincingly as by Thomas Metzinger and his theory of the Transparent Self-Model. While most critics look for logical errors and claim that Descartes is begging the question, Metzinger goes so far as to say that Descartes, and indeed every other “self” ever to exist, quite simply *is not*. I can imagine Metzinger's pithy reply: “You only *think* you think.” And it's not the thinking that is at issue here, but rather the you. In *Being No One* and *The Subjectivity of Subjective Experience*, Metzinger posits that the self is nothing more than an elaborate fabrication of the mind which lends it's subject selective advantage by allowing for higher forms of thought. This fabrication is generated by the brain in a way that is no more mysterious than the other complex workings of the nervous system, and thus it can be fleshed out through empirical observation and experimentation.

Metzinger's theory is, at its core, representational. According to this camp, man does not have direct, conscious contact with the world around him; instead, the senses and their associated processing streams create mental objects which can be observed and manipulated by consciousness. These mental objects need not be physical objects, but objects in the sense of a file on a computer. The mind is constantly processing input, creating objects, and passing them to consciousness.

In *Being No One*, Metzinger presents three types of content. *Presentational* content is analogous to Block's idea of P-Consciousness. Raw sense data is given to the mind to flag the

presence of a stimulus, but this sort of content either does not manifest itself in the form of objects, or at least not objects that can be manipulated. *Representational* content, on the other hand, are the objects and data that represent present, real-world things to the mind. Closely akin to this is *simulational* content, which represents contingent things—objects that *could* be.

Simulational content is used for decision making and goal-oriented action. These three make up the *what* of consciousness, and along with these Metzinger presents ten *how* constraints, three of which are of primary importance. These constraints must be applied to mental objects before they can be considered conscious. Firstly, content must be represented *globally* in the sense that it has an orientation to other objects in the world, and cannot be isolated. It must also be represented *presentationally*, meaning that it has an orientation in time—specifically that it is happening now. Lastly, and most importantly, it must be represented *transparently*. This transparency implies that the subject receiving the representations does not recognize them as representations, but experiences them as real things. A simple example can illustrate how this all comes together. When I place my hand on a table, the touch receptors in my fingers send presentational content to my mind. I then subconsciously construct a representational object of the table with I can consciously observe. This might also lead to simulational content if my intention is to move the table or modify it in some way. And I do not commune with the table alone as if with a Platonic Form; the table is represented globally as being an object located in a room and presentationally as existing right now in time. But I am oblivious to this process as it happens because it occurs transparently. I experience my representation of the table as if it were the real table.

These contents used under these constraints provide an excellent framework for describing how the mind observes and understands the outside world. At this point, Metzinger introduces two new ingredients which he believes can account for the existence of a first person

perspective. The first is the *Phenomenal Self Model*, or PSM. This is a mental object that represents the self. The second is the *Phenomenal Model of the Intentionality Relation*, or PMIR. This can be thought of as a link between the self and some other object. It is an arrow extending from the self and pointed toward something, or even back at the self. So, while the “old theory” was that the mind constructed a representational world around a real self, Metzinger proposes that the mind simply constructs a world *and* a self, places the self inside the world, and allows the self to be pointed at any number of objects in the world, including itself. The representational world is not built around us. We are built into it... or perhaps along with it. The self is just as much a fabrication as the rest of the representations that fill the mind, but because transparency applies to the self object just as to every other conscious object, we do not realize this to be so. We experience ourselves to be real, just as we experience the objects we observe to be real. Metzinger is simply taking the theory of representationalism one step further, a step that until now was perhaps too traumatizing or too well concealed for others to recognize. This is what he means by “being no one.”

In *The Subjectivity of Subjective Experience*, Metzinger identifies three important self-related characteristics to conscious experience. The first is *mineness*, the notion that certain things belong to the self. Mineness is applied to specific mental objects: my arm, my thoughts, my will. The second is *selfhood*, the philosophically troublesome but naturally intuitive idea that “I am a person” and that “I maintain a certain single identity through time.” The third and perhaps most interesting is *perspectivalness*. This is closely tied to both globality and the PMIR, and it is not a property of individual objects but of conscious space as a whole. Perspectivalness is a certain understanding of orientation in the representational space. It is what gives the self its position in relation to the other representations. It is, in a sense, the egoness of egocentricity. These three closely related qualities can be applied to representational objects that are part of the

self. When I look down at my hand, my mind creates (or calls upon) a representation of a hand, and the quality of mineness is then applied to that object, creating the notion that what I am viewing is not just any hand, but specifically my hand. That hand is part of my self, and I understand its relation to me and to the world of which I am the center. I have a connection to it, and can control it.

Metzinger offers two examples that suggest the existence of the PSM. After being in space for a few days, astronauts lose their sense of top and bottom. However, this intuition can be easily regained by applying pressure to the bottoms of the feet, presumably because it simulates the feeling of gravity. Obviously these astronauts have not forgotten which end of them is the top and which is the bottom, but the feeling of disorientation despite this knowledge is evidence that their senses are not receiving enough directional information to properly orient their representational selves in mental space. The self models that their brains maintain are, in a sense, misrepresented. An even more compelling argument can be found in the studies of phantom limbs. Some amputation patients continue to feel as if their missing limbs are attached, though they have no control over them and obviously cannot see them. In one experiment, amputation patients were placed in front of mirrors so that their existing limbs created the illusion that they still possessed their phantom limbs. Several patients reported that they once again had a feeling of control over their missing limbs and could move them. These two cases suggest that one's mental concept of self is governed more by sense input than by rational knowledge. Obviously, nothing real is moving when the amputee stands in front of the mirrors and moves about; but still movement is experienced. Metzinger argues that what moves is one's phenomenal body—the mental object corresponding to the body that has been built by the patient's mind—and since that object is transparent, the subject is fooled into experiencing actual movement.

This framework is very attractive for several reasons. Firstly, it is empirically rooted. Metzinger claims that the workings of this mental system can be correlated to actual neurological activity in the brain. He firmly believes that science will be able to pinpoint exactly which structures in the brain are responsible for the creation and observation of these mental objects. He provides something that most other theories of first person perspective do not: a scientific starting point for investigation and criterion for validation. He also explores parts of his theory from an evolutionary standpoint, suggesting that more advanced forms of the self model offer greater advantages for survival, self-maintenance, and eventually contemplation and complex goal-oriented action. Lastly, he offers some very convincing evidence for the usefulness of his theory by analyzing mental disorders in terms of his model. These examples demonstrate another attractive feature of Metzinger's framework: its usefulness for explaining how the mind functions (or in these cases, malfunctions).

When the brain is operating properly, it creates objects analogously to the way machines are built on an assembly line: as the unit passes each station, the proper parts are added so that the machine can carry out its designated function. If the wrong parts are added, or if necessary parts are left off, the machine malfunctions. This kind of misassembly is what occurs in schizophrenic patients. They experience thoughts the same way we do, however as the mental objects corresponding to those thoughts are being manufactured, some of them do not get the aforementioned quality of mineness added on. Thus, the thoughts are experienced as someone else's thoughts, because the electrochemical flags that signify those thoughts as *mine* are not present. This can go the other way as well: mania and certain types of religious experiences result from all action objects get labeled with mineness. Thus, the subjects feels as if everything that occurs is his own doing. Just as the label of mineness can be used improperly, the the mental concept of selfhood can become warped. Multiple Personality Disorder can be explained

by duplicating the property of selfhood and creating multiple self objects. Also, the self object can be misrepresented in cases like blindness denial and phantom limbs, when one's self object does not reflect the actual state of the body. Perspectivalness, a property of mental space as a whole, can disintegrate. This might account for feelings of transcendence—the feeling that one is no longer in the world but occupies a sort of ubiquitous, monistic perspective. Similar explanations can be formulated for almost any sort of psychological crisis that involves the self, and this makes Metzinger's framework very attractive.

Interestingly enough, he never offers an example of what might happen if the quality of transparency should fail. Transparency is vitally important to the self model theory, so the fact that he never offers an example of its failure or posits what that would be like might be an indication that transparency isn't actually employed by the mind at all. In Metzinger's defense, transparency is listed as the most important of the constraints for a mental object to become conscious, so he might argue that transparency can never fail—or at least never *noticeably* fail. Perhaps opaque objects never make it to consciousness at all. Or perhaps those in whom it fails become so incredibly insane that they cannot be fruitfully studied. Still, one cannot help but wonder, what would it be like if an organism were suddenly disillusioned and began to recognize its representations as representations? Such a disorder, if it is ever discovered, might aptly be named Metzinger Syndrome. Now, it should be noted that a philosopher who comes to agree with Metzinger (if he is right) would *not* be achieving this transparency disorder. What one judges to be true via intellection is different from what one feels to be true in one's phenomenal self, just as the astronauts knew which way was up but still felt disoriented. One would have to be intuitively gazing into conscious representational mental space, and if Metzinger's theories are correct, this would be a very traumatizing experience. Even worse than being logically convinced that one does not exist, one would have a deeply-rooted gut feeling that he is only a

construct in a world of mental objects. What an awful proposition!

Metzinger's theory, despite its many commendable aspects, has its share of weaknesses. The first major issue that might come to mind is the question of what the self model represents if no self exists. Representational objects are supposed to reflect real objects if the brain's architecture is working properly, so how can one deny the existence of a real self if a self model is constructed? What exactly is this model modeled after? The simple reply is that no self ever existed: it is and has always been an illusion. Nothing in Metzinger's representational framework prevents the mind from creating completely unique objects that have no correlate in reality. Such illusions are generally associated with mental disorders, but that is probably because such illusions are detrimental to a being's survival. The fabrication of the self is actually quite useful for an organism because it adds important new components to the computational building blocks of the mind. Theoretically, the self illusion was the product of mutation sometime in the evolution of animals, and it provided some creatures with selective advantage and was passed down. The self model actually models nothing. According to Metzinger, "the conscious self is an illusion which is *no one's* illusion" (43).

Alex Gamma offers a few interesting objections in his review of *Being No One*. Firstly he notes that some of the examples Metzinger uses to prove the usefulness of his theory are suspect, particularly the ones which rely on explaining the workings of patients with mental disorders. Gamma shrewdly points out that these patients' minds are not normal minds, and so while their abnormalities might be of interest to the study of brain malfunction, the results can be only so conclusive when applied to normal minds. Attempting to flesh out the workings of normal minds based on the functioning of *mal*functioning minds is a dangerous project and should be handled with care. Gamma also takes issue with Metzinger's handling of phenomenology, arguing that it will be much more convincing when based on large-scale studies

instead of personal intuition:

[... We] don't have anything like a good phenomenological evidence base. Philosophers today often do phenomenology with a sample size of one. They serve as their own subjects. Disagreements over the phenomenal character of conscious experience, such as [one of Gamma's disagreements], highlight the blatant need for systematic large-scale studies of phenomenal consciousness in normal individuals. (391)

The particular disagreement that Gamma references is, I feel, trivial in regard to Metzinger's theory as a whole, but Gamma does touch on the importance of studying *minds* rather than simply one's own mind. This is especially true if Metzinger is correct and the mind is particularly good at deceiving the person to which it belongs. Before closing, Gamma makes another well-directed criticism: Metzinger's writing style is extremely cryptic, at times unorganized, and often unclear. This does not affect the soundness of his theory, but unless he finds a way to express his ideas much more clearly, he will have a hard time getting across the important nuances of his theory, and there are many of those. He will also have a hard time finding readers.

But all of that aside, in my opinion, Gamma has missed the largest and most important hole in Metzinger's theory. This framework, however useful it may be, does not explain the epiphenomenal nature of first person experience. He does not solve what David Chalmers would call the *hard problem*. At best, the self model theory can provide a functional explanation of consciousness. It can explain how the brain conceptualizes objects and the self, but it cannot explain the ever-mysterious cause of why we have a distinctively first-person outlook on the world. The question still remains: why is there something that it is like to be a conscious human being? The self model theory might explain how the brain's electrochemical parts work together to digest self and other and present those concepts for processing, but it still does nothing to

explain the what-it-is-like-ness of consciousness. Metzinger might say that what-it-is-like-ness is all a part of the self illusion, but we must ask who that illusion is supposed to be deceiving. His answer that it is no one's illusion is witty, but ultimately unreasonable. Illusions are useless without an audience; if it is no one's illusion, then it is no illusion. It is possible that Metzinger did not intend to account for what-it-is-like-ness, and only meant to put forth a functional, phenomenal view of consciousness. But if that is true, then the self hardly needs to be considered an illusion. In this case the self model would not be a model at all, but the self itself! I see no need to refer to the self model as an illusion or a fabrication when speaking of the brain in functional terms. Even if the self *is* just a representational mental object, what better functional candidate could there be to call the self? The self is simply me, as I think of myself. It doesn't matter how I think of myself, and it doesn't matter if I understand how that thinking works. What else has the self ever been other than a construct of the mind? A mental representation is—if Metzinger's theories are correct—the real self in all its selfliness. There is no need to call it an illusion; the “illusion” might as well just be considered the real thing.

So if the self model theory is intentionally non-epiphenomenal, its claims seems much less interesting. But it cannot be anything more. It simply cannot account for what-it-is-like-ness in terms of an illusion if there is no one to deceive. Perhaps a concrete example will illustrate the epiphenomenal shortcomings of this theory. Most people would agree that a present-day Von Neumann serial processor computer does not have epiphenomenal experience. There is nothing that it is like to be a computer. But what if a programmer were to create a file on that computer and then reprogram the machine to treat the file as *itself the computer*. All processing that the computer had to do that involved itself would reference this file, and the file would be constantly updated to reflect the internal status of the computer's parts. Any input the computer received would be routed through that file. To extend the analogy even further, a

virtual machine could be added to the computer that would not be allowed to notice how that file was accessed or changed or even that it was a file at all. The file would be transparent as far as the virtual machine was concerned. The virtual machine would behave as if the file were itself. It could even be programmed to answer questions about itself. *Functionally speaking*, it would have a self (and, I would argue, not just an illusion of a self, but as far as the computer is *functionally* concerned, a real self). Most people would, presumably, still not attribute epiphenomenal experiences to the computer. There still would not be anything that it is like to be that computer. Even if we say that the computer is under an illusion, it is still just a machine. We have not deceived a conscious being—we have simply reprogrammed a computer. I think the ultimate problem with Metzinger's theory is that he is trying to describe what it is like to have an illusion of what-it-is-like-ness, which is circular at best and ridiculous at worst. Even if there is something that it is like to have that illusion, there is still *something that it is like*. There is still some as-yet unexplainable here-ness about consciousness that I certainly experience and that I can only assume others also experience.

The self model theory offers some interesting functional and phenomenal insights, and seems very useful for explaining psychological abnormalities that involve the self. It is refreshingly empirical and seems evolutionarily plausible. But it, like so many before it, still cannot account for the unavoidable epiphenomenal experiences that make consciousness such a mystery. If the self is a transparent mental object, then it can still be functionally equivalent to the self without being an illusion. In that sense, it might as well be the real self. But true selfhood in the most meaningful sense cannot simply be an illusion, because if there is no one to fool then there is no fooling to be done.

Works Cited

Gamma, Alex. *Review of Thomas Metzinger's Being No One. The Self-Model Theory of Subjectivity. Brain and Mind.* Netherlands: Kluwer, 2003. 385 - 93.

Metzinger, Thomas. *The Subjectivity of Subjective Experience: A Representational Analysis of the First-Person Perspective.* Draft.