I begin with a brief discussion of our idea of the self, then review some of the empirical literature, and end with a brief discussion of the models supported by the empirical work.

II. Four features of our concept of self

There are at least four characteristics of our concept of self that are central to both the ordinary notion and the philosophical concern. To begin, as I mentioned above, there is the fact that selves explain the unity of a human person. Selves explain how people are connected to their past and future. Bodies change, mental states change, but people tend to believe that in some sense of the word they are the same people now that they were in the past and will be in the future. The connection between past, present and future selves is sometimes called diachronic unity. A different sort of unity – synchronic unity – can also be explained by appeal to selves. This is the sense that all the various experiences that occur at a time are occurring to the same subject. For example, the experience I have of the keyboard, the sound of the cardinal singing outside, the sight of the dog on the carpet are all happening now to the same person, myself. The various experiences are unified in a single subject.

A second feature of selves which is important to both philosophers and the general public is the

These are not isolated features; they are related to each other in various ways. For example, the common American belief in the constancy of character is

the external stimuli, their mental states, or even their behavior. (Subjects, of course, realize that they are walking, but they don't realize they are walking slowly.)

A more dramatic demonstration of the connection between conscious and unconscious behavior can be found in the work of Pierre Fourneret and Marc Jeannerod (1998). Fourneret and Jeannerod asked subjects to use a stylus on a graphics tablet to trace a line represented on a computer screen.

2002, 18). In a meta-analysis, Greenwald and colleagues found that the IAT is a better predictor than self-reports for certain topics like Black-White interactive behavior and intergroup behavior (2002, 28). On other words, a subject's conscious beliefs are less useful a predictor of behavior than her unconscious implicit beliefs; moreover, these often will be in conflict.

This is a very small sample of a large and growing literature that undermines some of our most cherished beliefs about ourselves. All four characteristics discussed in the previous section – unity, agency, personality, and self-awareness – are implicated in these studies. Much of our behavior is driven by unconscious processes that are both inaccessible to introspection and often in conflict with our conscious desires. We are less unified, less in control, less stable in personality, and often plain wrong in our assessments of our selves.

IV. What now?

Two models of the self immediately suggest themselves in response to the empirical literature. According to the *multiple selves* model, the human mind possesses multiple processes that end in action and are fairly isolated from one another. This is illustrated by the Titchener illusion. When people with normal vision look at the Titchener illusion it appears that the middle circle surrounded by little circles is bigger than the middle circle surrounded by bigger circles. This being an illusion, the middle circles are the same size. And some part of the human mind knows this. When we reach to grab the inner circles our fingers form a grasping position which is identical in radius in both cases. Some part of our minds is not fooled by the illusion and it is this part which is in control of our grasping behavior. Essentially, the unconscious overrides our conscious experience and directs our behavior.[7] The neuroscientist V. S. Ramachandran uses this example, among others, to argue that our minds consist of lots of "zombie selves" that direct and control our behavior yet remain inaccessible, at least from the first-person perspective.[8]

Endorsing the multiple selves model is to give up the traditional philosophical pursuit of an account of personal identity. According to this model, the self is simply an illusion. Our concept of ourselves might be as a single entity in control of our lives but this concept is empty; there is no single entity that makes decisions, has experiences, and controls our behavior. This is fundamentally a skeptical

---

[7]This example is discussed in Ramachandran (1998).
[8]A more detailed account of the multiple selves can be found in Humphrey and Dennett (1989).

5

with an evolutionary past.[11] Taking our evolutionary past as a starting point for our accounts of the self is helpful for many reasons. First, it acts as a counter to the constant pull toward dualism. Thinking of ourselves in a dualistic way comes easily to us, and we need help in resisting it.[12] Second, it reinforces the essentially social nature of humans, which informs how we think about what kinds of selves we are. Much of the Western tradition has viewed the self in an atomistic, individualistic way. Along with many other philosophers, I think that this has been a mistake. Recent empirical work in the social sciences has begun to emphasize how surprising humans are in their understanding and ability to cooperate with others.[13] Philosophers need to incorporate these findings into our revised conception of the self. Third, making the evolutionary history of humans central to our understanding of our conception of ourselves pushes us to think about human behavior and less about mental life that does not result in behavior.

One of the benefits of this approach is that this conception appears to lead to more accurate self-knowledge. Recent work by Emily Pronin, Jonah Berger and Sarah Molouki (2007) suggests that we would be well served to start paying more attention to behavior and less attention to our mental states. In five studies exploring subjects' conformity judgments, the researchers found a consistent bias when it came to first person attributions of conformity compared to third person attributions of conformity. That is, subjects considered themselves as less likely than others to conform across a range of situations. More interesting for my purposes, Pronin et al. investigated the source of this bias. They found evidence across all five studies that the subjects fell prey to what the researchers call "the introspective illusion." In brief, the introspective illusions occur when subjects pay too much attention to their own mental states and not enough attention to their behavior in self-ascriptions. This is exacerbated by subjects' belief that introspective information about their own states is more valuable than introspective information about others' mental states. The upshot: more accurate self-assessments are made when a subject discounts her mental states and focuses on her behavior.

In conclusion, I hope to have shown that there is empirical literature relevant to philosophers' theorizing about the self, that this literature has many implications for certain conceptions of the self, and also undermines many of the first person methods that philosophers have relied upon in drawing their accounts.

---

[11] Here I am in agreement with Stanovich, whose account is informed by the fact that we are biological organisms shaped by evolutionary forces.

[12] See Paul Bloom, *Descartes' Baby* for an argument that dualistic thinking comes naturally to humans across the world.

[13] See Sarah Blaffer Hrdy, Michael Tomasello, Frans de Waal, and David Sloan Wilson for empirical work that focuses on the significance and genesis of human social abilities.

References

Balcetis, Emily Dunning, David and Miller, Richard (2008) "Do Collectivists Know Themselves Better than Individualists?" *Journal of Personality and Social Psychology*, 95:6, 1252-1267.

Bargh, John and Chartrand, Tanya ("1999) "The Unbearable Automaticity of Being." *American Psychologist,* 54:7, pp. 462-479.

Davies, Martin and Stone, Tony. (eds.) (1995) *Folk Psychology: The Theory of Mind Debate*, Oxford: Blackwell Publishers.

Doris, John. (2009). "Skepticism about Persons." *Philosophical Issues* 19: Metaethics, 57-91.

Dutton, Donald G. and Aron, Arthur (1974). "Some Evidence for Heightened Sexual Attraction under Conditions of High Anxiety." *Journal of Personality and Social Psychology*, 30:4, 510-517.

English, Tammy and Chen, Serena (2007). "Culture and Self-Concept Stability." *Journal of Personality and Social Psychology,* 93:3, 478-490.

Fourneret, Pierre and Jeannerod, Marc (1998). "Limited conscious monitoring of motor performance in normal subjects." *Neuropsychologia*, 36:11, 1133-1140.

Frankfurt, Harry G. (1971). "Freedom of the Will and the Concept of a Person." Reprinted in *The Importance of What We Care About* (1998) New York: Cambridge University Press, 11-25.

Gallagher, Shaun and Shear, Jonathan (eds.) (1999) *Models of the Self.* Exeter: Imprint Academic.