SELF AS FUNCTIONAL FICTION

William B. Swann, Jr. and Michael D. Buhrmester
University of Texas at Austin

Beginning with James, self-theorists have entertained two complementary visions of the self. The me has denoted people's representations of themselves, or the characteristics that they impute to themselves. The I has referred to the sense of self, the homunculus-like entity that experiences and reacts to the world, constructs mental representations and memories, guides action, and so on. Here we contend that both the Me and the I are cognitive constructions or fictions. People rely on these reality based fictions to know what is expected of them, to understand how they fit into the world, and to guide behavior. We review past and contemporary work through the lens of our self as functional fiction perspective. Finally, we discuss the implications of this perspective for several issues, including the nature of accuracy in social perception and the experience of free will.

Some have contended that knowing the self is no different from knowing other people. That is, just as we know others by observing their behavior and forming corresponding mental representations of them, so it is with knowing ourselves. Yet there is something more to the self than merely a set of mental representations. When people peer inward, they have the visceral sensation of a sentient being, a homunculus-like entity that experiences reality, regulates action, represents itself to itself, and so on. In this article, we suggest that although these complementary modes of self-knowledge seem qualitatively distinct, in reality they are structurally and functionally similar. In particular, both forms of self-knowledge are constructions that people generate to organize their thoughts and actions. Although these constructions are based on a true story (i.e., the person's activities and experiences), they are fictional in that they are typically partial and incomplete guesstimates regarding this true story (Wegner, 2002). Nevertheless, because these fictions roughly approximate the reality they are designed to represent, on balance they are highly adaptive. Hence, our label, *functional fictions*.

To be sure, some readers may regard the concept of a functional fiction as dubious, if not oxymoronic. Indeed, the history of Western thought is replete with testimonials to the virtues of truth over fiction. Plato, for example, asked "Is not to have lost the truth an evil, and to possess the truth a good?" (380BC/2008, p. 97). In this tradition, theorists developed models of accuracy of human thought based on the assumption that inferential activities should follow formal mathematical models of logical correctness (e.g., Edwards, 1961). Nevertheless, over the years some have critiqued such formal models. Some critics have acknowledged the utility of normative principles but noted that people routinely disregard such principles (Kahneman & Tversky, 1973; Nisbett & Ross, 1980). Others have questioned the applicability of such models to human decision making, suggesting instead that people should rely on inferential rules that are functional despite being logically incorrect (e.g., Gigerenzer, 2007; McArthur & Baron, 1983). The pragmatist philosophers have gone even further by actually defining the truth value of beliefs in terms of instrumentality. James (1907), for example, contended that:

> The true is the name of whatever proves itself to be good in the way of belief ... Surely you must admit this, that if there were no good for life in true ideas, or if the knowledge of them were positively disadvantageous and false ideas the only useful ones, then the current notion that truth is divine and precious and its pursuit a duty, could never have grown up or become a dogma. In a world like that, our duty would be to shun truth rather. But in this world, just as certain foods are not only agreeable to our tastes, but good for our teeth, our stomach, and our tissues; so certain ideas are not only agreeable to think about ... they are also helpful in life's practical struggles. (James, 1907, p. 96)

In this pragmatist tradition, Swann (1984) introduced the construct of pragmatic accuracy, a form of accuracy that varies as a function of the degree to which beliefs offer predictions that are confirmed by their experiences. Our self as functional fictions argument is in keeping with this pragmatic accuracy perspective (e.g., Gill & Swann, 2004). We use our functional fictions formulation to organize our review of past and contemporary work on the self. We begin with William James's (1890) distinction between the I and the me.[1]

JAMES'S I AND ME

In conceptualizing the self, James cast a wide net:

> In its widest possible sense, however, a man's self is the sum total of all that he can call his, not only his body and his psychic powers, but his clothes and his house, his wife and children, his ancestors and friends, his reputation and works, his lands and horses, and yacht and bank account. (James, 1890, p. 291)

---

1. Note that although James is usually credited for authoring the distinction between the I and Me, like most theoretical notions, this distinction appeared much earlier in ancient eastern philosophy amd religious texts such as the Bhagavad Gita and *Tao Te Ching*.

To bring some modicum of order to this long and ungainly list, he distinguished the two broad aspects of the self (we alluded to these aspects above). One such aspect was the Me, the mental representations that people have of themselves. These representations included all the characteristics that individuals associate with themselves, such as personal qualities (extrovert, tall, intelligent), qualities that align them with groups (e.g., female, Democrat, astronaut), and even external entities that they associate with themselves (e.g., one's personal belongings).

James also distinguished the I, which he described as the agent, thinker and knower. Worried that critics would be tempted to equate the I with a religious soul or Kant's transcendental ego, James insisted that the I had an actual empirical referent in the human body. Although adamant about this much, he was less clear about the precise nature of the I. For example, while asserting that the I was the thinker, he also argued that "the thoughts themselves are the thinkers" (James, 1892/1985, p. 216). To the modern eye, the notion that the I might be both the thinker and the thought is perplexing. Nevertheless, we believe that James was on to something: To most people, the notion that there exists an I who serves as the psychological master of the house is intuitively appealing. Moreover, theorists and researchers have continued to grapple with the construct, as evidenced by several contemporary formulations such as Baumeister's (1998) ego or executive agent and Wegner's (2002) conceptualization of free will. We will have more to say about these modern-day incarnations of the I later.

Although flawed, James's distinction between the Me and the I proved to be a powerful springboard for subsequent explorations of the self. Now, more than 12 decades later, the literature on the self has grown exponentially. On balance, this is a welcome development, as it attests to the richness of the ideas under scrutiny as well as the vigor with which they have been pursued. Yet, as authors who have been charged with the task of providing insight into the question "What is the self?", the task of translating the more than 100 contemporary self-constructs into a comprehensible answer is daunting. In light of the sheer enormity of this task, we have followed the example of several of our esteemed predecessors (e.g., Leary & Tangney, 2003) by conducting a review that is designed to be illustrative rather than exhaustive. We have, however, attempted to at least allude to most of the central issues that occupy center stage in the contemporary literature.

THE ME

James imagined that when children are born, they are shocked by a world of "blooming, buzzing confusion" (James, 1890, p. 462). At least part of this confusion may stem from the near certainty that infants have yet to construct a sense of self—or a sense of most anything else for that matter. They quickly begin to address the deficit by searching for regularities and patterns in the information they encounter. As they discover such regularities and patterns, they use them to form expectations, beliefs, and other mental representations of the world around them (see Guidano & Liotti, 1983; Heine, Proulx, & Vohs, 2006; Popper, 1963;).

At first, the mental representations that children develop involve relatively ordinary objects (toys, food), animals (the family dog), and other people (parents, siblings). Representations of the self emerge somewhat later generally around 18 to 24 months of age (Rosen & Patterson, 2011). The first representations of the self are

quite rudimentary, involving a mere capacity to recognize the physical self, something various other species are capable of doing (Lewis & Brooks-Gunn, 1979). As children mature, they will notice more and more regularities in their own behavior and the reactions they elicit from others. Gradually, they will use these regularities as a basis for constructing representations of themselves. James labeled these mental representations the me; in this essay, we also use the more contemporary term, self-views (Swann & Bosson, 2010). As we note above, these self-views are fictions in that they are partial distillations of abstractions, yet they are functional in that they are accurate more often than not.

Self-views are cognitive/affective structures that contain information about who we are. They appear to be based on at least three distinct sources. One source of self-knowledge is feedback from others (e.g., Cooley, 1902; Mead, 1934). Thus, for example, people who discover that others treat them with warmth and respect will develop positive self-views; those who are treated shabbily will develop negative self-views. A second source of self-knowledge is people's observations of their own behavior. People who notice that they effortlessly master athletic activities will conceive of themselves as superior athletes; those who observe that they are awkward and clumsy will develop more modest conceptions of their athleticism (Bem, 1972). Finally, people will also infer who they are by taking note of their standing relative to others (Festinger, 1954). As James (1890) put it:

> So we have the paradox of a man shamed to death because he is only the second pugilist or the second oarsman in the world. That he is able to beat the whole population of the globe minus one is nothing; he has "pitted" himself to beat that one; and as long as he doesn't do that nothing else counts. (p. 310)

Much ink has been spilled regarding the degree to which self-knowledge is unique and distinctive. One popular notion, for example, has been that representations of the self are distinct from representations of others because people have direct access to their own stream of conscious experiences of themselves (Funder, 1995; Vazire, 2010). Such direct access could produce exceptionally rich, detailed representations of oneself because, in principle at least, people are privy to unique data about themselves that are unavailable to others. In practice, however, this source of data may be limited by the fact that people spend remarkably little time (approximately 8% of total thoughts) engaged in introspection (Csikszentmihalyi & Figurski, 1982). Further, the very notion of direct access may be somewhat illusory, for people's conscious self-reflections are surely merely memories that serve as the basis for constructing self-views rather than unfiltered snapshots of some actual inner selves. Insofar as representations of both the self and others are products of a cognitive-construction process, the notion that people have direct access to their real selves loses force (Bargh & Morsella, 2008; Nisbett & Wilson, 1977; Wegner, 2002) and the notion that self-views are functional fictions gains traction.

Others have contended that although self-knowledge may be derived from the same sources as knowledge of others, the fact that we know more, and care more, about the self may imbue self-knowledge with unique properties. For example, the sheer volume of self-knowledge that people possess may enhance its role in information processing and memory (Kuiper & Rogers, 1979). Although it is certainly true that self-knowledge can be helpful in organizing and remembering information, the relative advantage enjoyed by self-knowledge tends to disappear as ac-

quaintanceship with the comparison other increases (e.g., Keenan & Baillet, 1980). Hence, in its capacity to assist information processing, the difference between self and other knowledge appears to be quantitative rather than qualitative.

In short, if there are qualitative differences in the nature of people's representations of themselves and others, they have yet to be demonstrated (Kihlstrom, Beer, & Klein, 2003; Swann & Bosson, 2010). This has given license to researchers interested in mental representations of the self to borrow conceptual distinctions from work on person perception or social cognition more broadly. We will discuss several such distinctions below.

EPISODIC VERSUS SEMANTIC REPRESENTATIONS IN MEMORY

Self-knowledge is presumably stored in memory (Klein & Gangi, 2010). One form that such memories may take is autobiographical scripts: relatively literal recollections of specific events from one's past, including episodes that unfolded in a specific time and place. In addition to such episodic memories, people may also possess two types of semantic memories of themselves. One form of semantic memory consists of context-free factual knowledge (e.g., remembering one's alma mater or home address). The other form of semantic memory consists of knowledge of one's dispositions and traits (e.g., outgoing vs. shy). Such trait summaries are highly adaptive because they allow people to make quick judgments about themselves, thereby obviating the effortful retrieval of autobiographical knowledge (Klein, Cosmides, & Costabile, 2003).

The foregoing forms of self-knowledge appear to be independent. Support for this proposition comes from both experimental research and case studies. The experimental research has shown that activating one form of self-knowledge does not necessarily facilitate memory for the other forms of self-knowledge (e.g., Klein & Loftus, 1993; Klein, Loftus, Trafton, & Fuhrman, 1992). This evidence has been buttressed by evidence from a series of intriguing case studies of individuals who have lost one type of memory but not another (Klein & Gangi, 2010; Klein & Lax, 2010). For example, some amnesiacs lose a portion of their episodic and semantic factual self-knowledge, but report their traits accurately (as indicated by convergence between their self-reports and the impressions of people who are acquainted with them). Even more surprising is a recent case of a severe amnesiac who could recall past autobiographical events but felt as though he was recalling someone else's experiences rather than his own experiences! This dissociation of the episode from the experience of one's personal involvement (or ownership) in the episode suggests the presence of an ownership tag that is linked to one's episodic memories of self. In theory, these ownership tags may play a crucial role in maintaining a temporally stable sense of self, a feeling that we later argue serves a crucially important function (Klein & Nichols, in press).

SELF-ESTEEM VS. SELF-CONCEPTS

Trait summaries and other self-views can vary in specificity, and these variations can alter the properties of the self-view. For example, several theorists have conceptualized the distinction between self-esteem and self-concepts in terms of dif-

ferences in specificity (Marsh & Craven, 2006; Shavelson, Hubner, & Stanton, 1976; Swann, Chang-Schneider, & McClarty, 2007). The most global such representation involves people's estimates of how worthwhile they are or self-esteem. The lack of consensus regarding what makes people worthwhile introduces the possibility of disagreement between the esteem that target individuals have for themselves vs. that which others hold them in, leading others to conclude that targets have fictional assessments of their own self-worth. In contrast, self-concepts are relatively specific: people evaluate themselves on a wide range of qualities, including intelligence, sociability, athleticism, attractiveness, and so on. The relative specificity of self-concepts will mean that they are more tightly tethered to reality than self-esteem.

Some conceptualizations suggest that self-knowledge is organized hierarchically (Marsh & Hattie, 1996), with highly global concepts (i.e., self-esteem) sitting atop mid-level concepts (i.e., specific identities) which are likewise atop highly specific concepts (i.e., personally nuanced attributes; McConnell, 2011; Shavelson et al., 1976). Other formulations (e.g., Franks & Marolla, 1976; Gecas, 1971; Tafarodi & Swann, 2001) eschew the notion of a single global self-esteem, suggesting instead that there are two forms of global self-esteem, self-liking (the degree to which one regards oneself as good and worthy of love) and self-competence (the degree to which one regards oneself as able to bring about desired outcomes). Still other formulations have framed the distinction between self-esteem and self-concept in terms of affective involvement (Baumeister, Campbell, Krueger, & Vohs, 2003), with self-esteem referring to affective evaluations of self and self-concepts referring to cognitions about the self.

In addition to interest in the manner in which self-knowledge is organized, researchers have also attempted to link self-concept and self-esteem to important outcomes. One important and influential research tradition has shown that perceptions of self-efficacy predict a host of performance outcomes, such as academic grades and test scores (e.g., Bandura, 1986). At the same time, consistent with the specificity matching principle (Swann et al., 2007), others have shown that global self-esteem is predictive of broader, aggregate outcomes such as life satisfaction and depression (e.g., Multon, Brown, & Lent, 1991; Robbins, Lauver, Le, Langley, & Carlstrom, 2004; Valentine, DuBois, & Cooper, 2004; Werner & Smith, 1992). Clearly, although researchers have identified many forces that distort self-views (e.g., accessibility, defensiveness, demand characteristics), they are usually based on actual evidence rather than simply fabricated.

IMPLICIT VS. EXPLICIT SELF-ESTEEM

During the last two decades social-personality psychologists have devoted increased attention to nonconscious processes (Bargh & Morsella, 2008; Greenwald & Banaji, 1995; Wilson, 2002). A burgeoning literature has made clear that nonconscious processes are related to wide-ranging and surprisingly complex behaviors. Students of the self have joined this trend and designed several measures of implicit self-esteem. At this stage in the development of the construct, there is little agreement regarding several key issues, including even the meaning of the term implicit. According to a dual systems perspective (Chaiken & Trope, 1999; Greenwald & Banaji, 1995), implicit constructs themselves (i.e., self-esteem, attitudes)

exist in implicit forms that are separate from their explicit, conscious counterparts. A rival perspective (Fazio & Towles-Schwen, 1999; Hetts, Sakuma, & Pelham, 1999), however, holds that there exists only a single construct (i.e., self-esteem or attitude). In this instance, features of the construct (e.g., causes, contents, or consequences) may be inaccessible to, or distorted by, self-report measures, necessitating the use of an implicit (i.e., indirect) measure.

The debate regarding what is implicit about implicit self-esteem notwithstanding, interest in the construct has burgeoned, leading to the development of several measures. An initial examination of the psychometric properties of these measures (reliability and convergent validity) revealed that the reliability of only two measures approached acceptable levels (Bosson, Swann, & Pennebaker, 2000). More recently, a thorough review of the convergent and predictive validity of the two surviving implicit measures, the Name Letter Test and Implicit Association Test, revealed that neither measure displayed strong evidence of validity (Buhrmester, Blanton, & Swann, 2011). Apparently, although there surely are aspects of self-esteem that people are unwilling or unable to report, it does not appear that the methods developed up to date are effectively capturing these aspects.

SOCIAL VS. PERSONAL SELVES

Social identity theorists (e.g., Tajfel & Turner, 1979) distinguish representations of self that align people with groups (e.g., social selves, such as Democrat, Texan) versus representations that make people unique (e.g., personal selves, such as competitive, resilient). Later theorists added the concept of the relational self to refer to self-representations associated with the close bonds formed in interpersonal relationships (Brewer & Gardner, 1996). Perceiving individuals through the lens of social and relational group categories has been a longstanding and important framework for understanding behaviors as wide reaching as the violence of crowds and mobs (e.g., Diener, 1980) to the bonding activities of people engaged in close relationships (Aron, Aron, & Smollan, 1992).

From a functional fictions perspective, both social and personal identities are likely to be accurate insofar as they are defined in ways that link them to a particular set of circumscribed, observable behaviors. For example, whereas those with the social identity plumber are expected to display a specific repertoire of highly visible behaviors (e.g., repairing leaks), those with the personal identity athletic are expected to seek and excel in athletic activity. Identities that are only loosely associated with particular patterns of behavior (e.g., American, iconoclast) will tend to be more loosely associated with a given subset of behaviors.

IMAGINED SELVES

As the name implies, possible selves refer to fictional self-representations associated with one's hopes, dreams, and dreads (Markus & Nurius, 1986). A related distinction has been made between one's thoughts about how one actually is from thoughts about how one ideally could be and how one ought to be (Higgins, 1987). Still others have distinguished the feared self—an undesirable but possible future self that people wish to avoid (e.g., Carver, Lawrence, & Scheier, 1999).

Self-discrepancy theory has played a key role in understanding self-related emotions and motivations (Higgins, 1987, 1989). For example, people are more likely to experience specific emotions such as guilt and dejection when they perceive relatively large discrepancies between their actual self and various aspirational selves (Higgins, Shah, & Friedman, 1997). Similarly, abundant evidence supports the significance of the degree to which people possess a promotion-focused versus prevention-focused orientation to their activities. Within this conceptual paradigm, some people routinely construct fictional selves related to either approach or avoidance, and these fictions have important motivational consequences.

METACOGNITIVE ASPECTS OF SELF

The extensive set of cognitive representations people can hold about themselves grows exponentially when one considers meta-cognitive qualifiers of these representations. Imagine a woman struggling with her career as a musician. As memories of her past triumphs and tribulations flood her consciousness, she evaluates each one, imbuing each with a positive—or more likely given her struggles—negative, valence. She might then start to reconsider how important her career is to her, how certain she feels about her musical talent, and how contingent her feelings of self-worth are pegged to her career. Feeling increasingly despondent, her clarity about herself may plummet.

Research has demonstrated that such meta-cognitive fictions about the self can have self-fulfilling consequences. For example, meta-cognitive factors such as belief certainty and importance moderate the competing motives for self-verification and self-enhancement, which may in turn steer one into a rewarding or abusive relationship (Kwang & Swann, 2010; Swann & Ely, 1984; Swann & Pelham, 2002). In addition, people who rate themselves on the extreme ends of an attribute dimension and also regard that attribute as important are more resistant to challenges to their self-views than their counterparts (Markus, 1977). Similarly, endorsements of extreme positions on measures of self-view malleability (i.e., whether a trait such as intelligence can change or is fixed) are associated with unique motivational and performance patterns (Dweck, 2000).

In summary, since James (1890) introduced the Me, researchers have made great strides in both broadening and deepening their understanding of the construct. This interest in the content of people's fictions about themselves is not surprising, for content-rich representations of people—both others and the self—are obviously important. For example, one could argue that such representations represent the cognitive glue that enables people to initiate and maintain relationships through the process of identity negotiation (Swann, 1987; Swann & Bosson, 2008). Another reason why the Me has received so much attention is that it is highly intuitive and easy to grasp. In this regard the Me is quite distinct from the I, which many scholars have struggled to understand.

THE I

If Descartes's dictum "I think therefore I am" immediately engulfed the I with a luminous intellectual fog, that fog only grew denser when James shared his thoughts

on the matter. As we mentioned earlier, while distancing the I from opaque constructs such as transcendental ego, the terms James added instead (the agent, thinker, and knower) unfortunately referred to psychological processes that were themselves vague and inscrutable. Even today, the I is rarely characterized in any detailed way by researchers but is instead understood in terms of what it does.

The I is believed to do several things, each ambitious in its own right. Various theorists have proposed that the I oversees regulatory processes, awareness processes, perceptual processes, and so on (Wegner, 2005). In each case, the properties of the I that theoretically orchestrate the process are left unspecified and, in some instances, the nature of the I is never even discussed. For example, in self-perception theory (Bem, 1972), there is no discussion of the properties of the entity that is doing the perceiving; instead, the analysis is entirely focused on the parameters that govern the types of inferences people make about themselves. In other instances, the I or self is simply the entity that makes the process unfold. Self-control, for instance, is defined as "the exertion of control over the self by the self" (Muraven & Baumeister, 2000). In what follows, we take a closer look at these and other processes that researchers have linked to the I.

THE EXECUTIVE AGENT, THE CONTROLLER, THE REGULATOR

With a nod to James's I as agent or Freud's (1923/1961) ego, several researchers have proposed that one core function of the self is that of active agent that controls or regulates various reactions and activities—thoughts, emotions, impulses, and behaviors (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Carver & Scheier, 1981). These regulatory processes are thought to have important consequences. For example, a widely cited early study tracked children who varied in their capacity to resist eating a tasty treat (e.g., Mischel, Shoda, & Rodriguez, 1989). Those children who were able to successfully resist temptation later enjoyed better life quality as adults than those who were less capable self-regulators.

More recent work has confirmed and extended these early findings. That is, among adult participants, those who scored high on a self-report measure of self-control enjoyed more positive outcomes (e.g., higher grades, higher self-esteem, fewer mental health problems) than their low scoring compatriots (Tangney, Baumeister, & Boone, 2004). Experimental manipulations of ego depletion produce results that parallel the effects of individual differences (Muraven & Baumeister, 2000). That is, persons who engaged in activities that were designed to induce either mental or physical fatigue subsequently performed more poorly on tasks that required inhibiting their impulses. Furthermore, providing participants with glucose rejuvenated individuals who were fatigued, thereby restoring their regulatory functioning (Gailliot et al., 2007).

As intriguing as this work on ego depletion may be, thus far there has been no documentation of the mediating role of ego in this work. Furthermore, evidence from several studies raises questions regarding the putative role of the self in these processes. Consider, for example, that ego depletion effects have been replicated using dogs as participants (Miller, Pattison, DeWall, Rayburn-Reeves, & Zentall, 2010). Moreover, according to research guided by the auto-motive model of self-regulation (Bargh, 1990), the sequence of goal setting, activation, and completion can all be triggered by the environment and remain outside of conscious aware-

ness. Together, such findings imply that if the self is underlying the effects of ego depletion, it is a rather rudimentary, nonconscious sense of self. Such a self seems a rather pale image of the omniscient agent that James discussed.

THE EXPERIENCER, THE INTERPRETER, THE PERCEIVER

In the Jamesian tradition of the relatively passive I as knower and thinker, others have proposed that the I experiences life and attempts to make sense of it. For example, the experiencer I (Butterworth, 1992; Wegner & Erskine, 2003) may notice what is happening and react to it in some way. These reactions can be to anything, including one's surroundings (I love the feel of the wind on my face when I go sailing), body (My tooth aches), and emotions (I felt nervous walking to the podium). In contrast, the interpreter I is intent upon finding meaning in the world. This version of the I is highlighted in Gazzaniga's (1989, 1998) research on split brain patients (individuals whose corpus callosa were surgically severed as a treatment for epilepsy). Patients were instructed to engage in a series of actions, the content of which they were consciously unaware because the instructions were available to the right side of the brain only. Gazzaniga discovered that their left brain interpreter confabulated plausible but factually incorrect explanations for the behavior. This semi-clueless interpreter is also featured in self-perception theory (Bem, 1972). Here, people impute meaning to their own behavior in the same way that an objective observer might: They take stock of the behavior and the contexts in which it occurs and make causal attributions for the behavior. If the behavior can be chalked up to contextual pressures, they conclude that the behavior says little about their own idiosyncratic characteristics. In contrast, if contextual pressures cannot account for the behavior, people decide that the behavior was indeed a reflection of their own qualities.

THE MOST IMPORTANT THING THAT THE
ME AND I HAVE IN COMMON

On the face of it, the me and the I seem strikingly different. For starters, the nature of the me is relatively transparent but the nature of the I is not. The Me is a cognitive representation (or construction) pure and simple, parallel to other cognitive representations, especially representations of other people. Furthermore, knowing that the me is composed of cognitive constructions tells us exactly what it is: a set of fictions that we construct to make our worlds coherent, predictable, and controllable (e.g., Mead, 1934).

The nature of the I is more obscure. Mysterious even. A familiar presence to all sentient individuals, the I is the entity that enables us to simultaneously navigate and experience our lives, that often makes choices that win us approval but sometimes gets us into hot water. But defining the I as "an entity that performs various functions" is unsatisfying, as it merely tells us what the I *does;* it does not explain what it *is.* Referring to the I as an entity may be preferable to spirit or transcendental ego, but without specifications as to the nature of the entity this definition merely escorts us down the path of infinite regress.

A solution to this conundrum is offered by Wegner's suggestion that like the me, the I is a cognitive construction (Wegner, 2002, 2008; Wegner & Wheatley, 1999). In his penetrating analysis of free will, Wegner makes a convincing case that the I is a convenient fiction that is constructed to refer to the subjectively experienced agent which experiences and acts upon the world. The I may also serve various social-communicative functions. For example, by noting that "I did X" rather than "My body did X," people claim authorship of their actions, thus making them accountable to the larger social community. It is hard to imagine how any society could function without imputing intentionality to social behavior on occasion.

But Wegner raises an additional, more controversial, possibility: If the I (Wegner actually uses the term Will, which refers specifically to the conscious I) is merely a cognitive construction, then perhaps it does not possess the causal power that people typically ascribe to it. Within Wegner's framework, most human behaviors are controlled by nonconscious mental processes—even relatively complex human behaviors that humans have historically credited to the conscious self (see also Bargh, 1999; Bargh & Ferguson, 2000; Baumeister, Schmeichel, & Vohs, 2007; Dijksterhuis & Nordgren, 2006). In reality, these remarkable achievements are produced not by consciousness but by a diverse collection of relatively unsophisticated nonconscious processes that are collectively responsible for the stunningly sophisticated activities that are uniquely human. Dennett (1978) explains this phenomenon using the analogy of a termite colony. The termite colony as a whole can do many smart looking things—raid other colonies, build elaborate mounds, and so on—without the benefit of a brilliant maestro orchestrating their efforts. Rather, individual termites of quite modest intelligence work together, each one performing very simple tasks that collectively result in products that appear to have required substantial intelligence. Similarly, simple, nonconscious components of people's brains may work together to produce behaviors that have heretofore been considered fruits of consciousness. In fact, when people are solving complex problems, nonconscious processes may outperform their conscious counterparts (Dijksterhuis & Nordgren, 2006). Wegner concludes that if nonconscious processes are actually responsible for regulating activities that people believe that they have consciously willed, then the notion of conscious will—one of our most precious pearls of cultural wisdom—is an illusion.

Does this mean that the I is an illusion? With respect to many usages that refer to the conscious I ("I am writing this sentence," "I chose to marry my wife"), Wegner's assertion seems spot on. Yet some usages of I refer to nonconscious processes that are, by definition, outside of the purview of Wegner's analysis of conscious will. For example, when people say "I fell asleep", or "I slipped in the shower", or "I referred to one of my daughters using her sister's name," they are invoking a nonconscious I as the agent that produced unintentional actions. Therefore, Wegner's assertion that conscious will is an illusion rings true but, by definition, it does not extend to usages of the I that refer to activities that are regulated nonconsciously. Furthermore, we suspect that even in those instances in which people erroneously impute causal potency to the I, the illusion is ordinarily a relatively benign one that does not impair their ability to function effectively. In fact, we suspect that on balance, illusions regarding the I are less problematic than illusions associated with the Me.

Like the I, the Me is a construction that may or may not capture reality. Accuracy of these constructions will vary as a function of several variables, such

as the degree to which the relevant self-views rely on social validation for their survival (Vazire & Mehl, 2008). Most people's social identities, for example, do require social validation and are hence almost always on the mark: Self-described Democrats rarely mistake themselves for Republicans, fathers for daughters, or fire-fighters for librarians. The accuracy of relatively private self-views, such as self-esteem or self-worth, may be more modest, in part due to ambiguities inherent in defining the value of a person. This surely helps explain why people suffering from depression render overly harsh judgments of their self-worth while narcissists consistently inflate their personal value and accomplishments (Bosson & Swann, 2009). Despite this, even self-esteem is not completely illusory, as evidenced by systematic relationships between self-esteem and the ratings of peers (e.g., Swann & Predmore, 1985).

More generally, we suggest that although both the I and the Me are fictions, in many instances these fictions are high in pragmatic accuracy. In fact, the utility of the self hinges fully on its pragmatic accuracy. Wegner (2002), for example, reviews numerous case studies in which people feel that a different person, thing, or multiple selves have taken over their bodies. Not surprisingly, these cases are considered pathological. There are several reasons why society should designate such individuals this way, not the least of which is the psychological and interpersonal anarchy that ensues when a single I fails to claim responsibility for the individual's behaviors. No less important are the temporally stable representations of the self that comprise the Me. Deprived of a coherent I and Me, people quickly lose a sense of place and a feeling that the world is a coherent, predictable setting in which they can successfully manage their affairs. A stable, coherent self also reassures people that they are on the right track, thus emboldening them to take actions. Moreover, once they initiate actions, the self may encourage people to consistently enact the same patterns of behavior. This will make people more predictable to others, thus greasing the wheels of social interaction (Swann, 1983, 2011). In these and other ways, the functional fictions that are the self may prove to be quite indispensable.

REFERENCES

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596–612.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition* (Vol. 2, pp. 93–130). New York: Guilford.

Bargh, J. A. (1999). The cognitive monster: The case against controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 361-382). New York: Guilford.

Bargh, J. A., & Ferguson, M. L. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.

Bargh, J. A., & Morsella, E. (2008). The unconscious mind. *Perspectives on Psychological Science*, *3*, 73–79.

Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680–740). New York: McGraw-Hill.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is

the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252–1265.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, *4*, 1–44.

Baumeister, R .F., Schmeichel, B. J., & Vohs, K. D. (2007). Self-regulation and the executive function: The self as controlling agent. In A. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 516–539). New York: Guilford.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.

Bosson, J., & Swann, W. B., Jr. (2009). Self-esteem: Nature, origins, and consequences. Chapter prepared for R. Hoyle & M. Leary (Eds.), *Handbook of individual differences in social behavior* (pp. 527–546). New York: Guilford.

Bosson, J., Swann, W. B., Jr., & Pennebaker, J. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*, 631–643.

Brewer, M. B., & Gardner, W. L. (1996). Who is this "we"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, *71*, 83–93.

Buhrmester, M. D., Blanton, H. & Swann, W. B. Jr. (2011). Implicit self-esteem: Nature, measurement, and a new way forward. *Journal of Personality and Social Psychology*, *100*, 365–385.

Butterworth, G. (1992). Origins of self-perception in infancy. *Psychological Inquiry*, *3*, 103–111.

Carver, C. S., Lawrence, J. W., & Scheier, M. F. (1999). Self-discrepancies and affect: Introducing the role of feared selves. *Personality and Social Psychology Bulletin*, *25*, 783–792.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer-Verlag.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford.

Cooley, C. H. (1902). *Human nature and the social order.* New York: Scribner's.

Csikszentmihalyi, M., & Figurski, T. J. (1982). Self-awareness and aversive experience in everyday life. *Journal of Personality*, *50*, 15–28.

Diener, E. (1980). Deindividuation: The absence of self-awareness and self-regulation in group members. In P. B. Paulus (Ed.), *The psychology of group influence* (pp. 209–242). Hillsdale, NJ: Lawrence Erlbaum.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, *1*, 95–109.

Dennett, D. (1978). *Brainstorms: Philosophical essays on mind and psychology.* Cambridge: Bradford Books/MIT Press.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Psychology Press.

Edwards, W. (1961). Behavioral decision theory. *Annual Review of Psychology*, *12*, 473–49.

Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 97–116). New York: Guilford.

Festinger, L. (1954). *A theory of social comparison processes.* New York: Human Relations Press.

Franks, D. D., & Marolla, J. (1976). Efficacious action and social approval as interacting dimensions of self-esteem: A tentative formulation through construct validation. *Sociometry*, *39*, 41–58.

Freud, S. (1961/1923). The ego and the id. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19; pp. 3–66). London: Hogarth Press. (Original work published 1923)

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652–670.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., & Tice, D. M. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*, 325–336.

Gazzaniga, M. S. (1989). Organization of the human brain. *Science*, *245*, 947–952.

Gazzaniga, M. S. (1998). The split brain revisited. *Scientific American*, *279*, 50–55.

Gecas, V. (1971). Parental behavior and dimensions of adolescent self-evaluation. *Sociometry*, *34*, 466–82.

Gigerenzer, G. (2007). *Gut feelings*: *The intelligence of the unconscious*. New York: Viking Press.

Gill, M. J., & Swann, W.B., Jr. (2004). On what it means to know someone: A matter of pragmatics. *Journal of Personality and Social Psychology*, *86*, 405–418.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Guidano, V. F., & Liotti, G. (1983). *Cognitive processes and emotional disorders.* New York: Guilford.

Heine, S. J., Proulx, T., & Vohs, K. D. (2006). The meaning maintenance model: On the coherence of human motivations. *Personality and Social Psychology Review*, *10*, 88–110.

Hetts, J. J., Sakuma, M., & Pelham, B. W. (1999). Two roads to positive regard: Implicit and explicit self-evaluation and culture. *Journal of Experimental Social Psychology*, *35*, 512–559.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1989). Self-discrepancy theory: What patterns of self-beliefs cause people to suffer? In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 93–136). New York: Academic Press.

Higgins, E. T., Shah, J. Y., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, *72*, 515–525.

James, W. (1890). *The principles of psychology*. New York: Holt.

James, W. (1907). *Pragmatism*. New York: Longmans-Green.

James, W. (1985/1892). *Psychology: The briefer course*. Notre Dame, IN: University of Notre Dame Press. (Original work published 1892).

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Keenan, J. M., & Baillet, S. D. (1980). Memory for personally and socially significant events. In R. S. Nickerson (Ed.), *Attention and Performance VIII* (pp. 651-669)*.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kihlstrom, J. K., Beer, J., Klein, S. B. (2003). Self and identity as memory. In M. Leary and J. Tangney (Eds.). *Handbook of self and identity.* (pp. 68–90). New York: Guilford.

Klein, S. B., Cosmides, L., & Costabile, K. (2003). Preserved knowledge of self in a case of Alzheimer's dementia. *Social Cognition, 2,* 157-165.

Klein, S. B., & Gangi, C. E. (2010). The multiplicity of self: Neuropsychological evidence and its implications for the self as a construct in psychological research. *The Year in Cognitive Neuroscience 2010: Annals of the New York Academy of Sciences*, *1191*, 1–15.

Klein, S. B., & Lax, M. L. (2010). The unanticipated resilience of trait self-knowledge in the face of neural damage. *Memory*, *18*, 918–948.

Klein, S. B., & Loftus, J. (1993). The mental representation of trait and autobiographical knowledge about the self. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 5, pp. 1–49). Hillsdale, NJ: Erlbaum.

Klein, S. B., Loftus, J., Trafton, J. G., & Fuhrman, R. W. (1992). Use of exemplars and abstractions in trait judgments: A model of trait knowledge about the self and others. *Journal of Personality and Social Psychology*, *63*, 739–753.

Klein, S. B., & Nichols, S. (in press). Memory and the sense of personal identity. *Mind*.

Kuiper, N. A., & Rogers, T. B. (1979). Encoding of personal information: Self-other differences. *Journal of Personality and Social Psychology*, *37*, 499–514.

Kwang, T., & Swann, W. B., Jr. (2010). Do people embrace praise even when they feel unworthy? A review of critical tests of self-enhancement versus self-verification. *Personality and Social Psychology Review*, *14*, 263–280.

Leary, M. R., & Tangney, J. P. (2003). *Handbook of self and identity*. New York: Guilford.

Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self.* New York: Plenum.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, *35*, 63–78.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Marsh, H. W., & Craven, R. G. (2006) Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives, *Perspectives on Psychological Science*, *1*, 133–163.

Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept* (pp. 38–90). New York: Wiley.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*, 215–238.

McConnell, A. R. (2011). The multiple self-aspects framework: Self-concept representation and its implications. *Personality and Social Psychology Review*, *15*, 3–27.

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Miller, H. C., Pattison, K., DeWall, C. N., Rayburn-Reeves, R., & Zentall, T. R. (2010). Self-control without a "self"? Common self-control processes in humans and dogs. *Psychological Science*, *21*, 534–538.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*, 933–938.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*, 30–38.

Muraven, M. R., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, *126*, 247–259.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Plato. (2008). *The republic* (B. Jowett, Trans.) New York: Cosimo. (Original work published 360 B.C.)

Popper, K. (1963). *Conjectures and refutations*. London: Routledge.

Robbins, S. B., Lauver, K., Le, H., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*, 261–288.

Rosen, L. H., & Patterson, M. M. (2011). The self and identity. In M. K. Underwood & L. H. Rosen (Eds.), *Social development* (pp. 73-100). New York: Guilford.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407–441.

Swann, W. B., Jr. (1983). Self-verification: Bringing social reality into harmony with the self. In J. Suls & A. G. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. 2, pp. 33–66). Hillsdale, NJ: Erlbaum.

Swann, W. B. Jr. (1984). The quest for accuracy in person perception: A matter of pragmatics. *Psychological Review*, *91*, 457–477.

Swann, W. B., Jr. (1987). Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology*, *53*, 1038–1051.

Swann, W. B., Jr. (2011). Self-verification theory. In P. Van Lang, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 23–42). Sage: London.

Swann, W. B., Jr. & Bosson, J. (2008). Identity negotiation: A theory of self and social interaction. In O. John, R. Robins, & L. Pervin (Eds.), *Handbook of personality psychology: Theory and research I* (pp. 448–471). New York: Guilford.

Swann, W. B., Jr., & Bosson, J. (2010). Self and identity. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 589–628). New York: McGraw-Hill.

Swann, W. B., Jr. Chang-Schneider, C., & McClarty, K. (2007) Do our self-views matter? Self- concept and self-esteem in everyday life. *American Psychologist*, *62*, 84–94.

Swann, W. B., Jr., & Ely, R. J. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology*, *46*, 1287–1302.

Swann, W. B., Jr., & Pelham, B. W. (2002). Who wants out when the going gets good? Psychological investment and preference for self-verifying college roommates. *Journal of Self and Identity*, *1*, 219–233.

Swann, W. B., Jr., & Predmore, S. C. (1985). Intimates as agents of social support: Sources of consolation or despair? *Journal of Personality and Social Psychology*, *49*, 1609–1617.

Tafarodi, R. W., & Swann, W. B., Jr. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences*, *31*, 653–673.

Tajfel, H., & Turner, J.C. (1979). An integrative theory of intergroup relations. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–48). Monterey, CA: Brooks-Cole.

Tangney, J. P., Baumeister, R. F., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*, 271–324.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*, 111–133.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*, 281–300.

Vazire, S., & Mehl, M. M. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*, 1202–1216.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D. M. (2005). Who is the controller of controlled processes? In R. Hassin, J. S. Uleman, & J.A. Bargh (Eds.), *The new unconscious* (pp. 19–36). New York: Oxford University Press.

Wegner, D. M. (2008). Self is magic. In J. Baer, J. C. Kaufman, & R. F. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 226–247). New York: Oxford University Press.

Wegner, D. M., & Erskine, J. (2003). Voluntary involuntariness: Thought suppression and the regulation of the experience of will. *Consciousness and Cognition*, *12*, 684–694.

Wegner, D. M., & Wheatley, T. P. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*, 480–492.

Werner, E. E., & Smith, R .S. (1992). *Overcoming the odds: High risk children from birth to adulthood*. Ithaca, NY: Cornell University Press.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Harvard University Press.