To appear in: O. P. John, R. W. Robins, & L. A. Pervin (Eds.). *Handbook of personality: Theory and research* (3rd edition). New York: Guilford.

Naturalizing the Self

Richard W. Robins
University of California, Davis

Jessica L. Tracy   Kali H. Trzesniewski
University of British Columbia   University of Western Ontario

Draft Date: 8/28/08

Author Note: Correspondence should be addressed to Richard W. Robins, Department of Psychology, University of California, Davis, CA 95616-8686. E-mail may be sent to: rwrobins@ucdavis.edu.

Late in his life, Michelangelo began carving what many art historians consider his most mature and provocative sculpture, the Florentine Pieta, an enormous 8 ft. statue he intended to place at the top of his own tomb. After working intensely for a decade on this monumental project, the artist entered his studio one day and, in a fit of rage, assaulted the sculpture with a sledgehammer. He broke off the hands and legs and nearly shattered the work before his assistants dragged him away. Why would Michelangelo attempt to destroy one of his greatest creations, a sculpture that has been described as among the finest works of the Renaissance?

How would a personality psychologist answer this question? A trait researcher might say that Michelangelo was highly impulsive and dispositionally prone to negative emotionality. A biologically oriented researcher might speculate that he had a deficiency in the monoamine oxidase A gene, low levels of serotonin, and an atypical pattern of activation in the frontal and temporal lobes. A motivational researcher might assume that Michelangelo's personal projects shifted and the Florentine Pieta came into conflict with other important goals. Yet none of these explanations provides a completely satisfactory account of Michelangelo's seemingly irrational act. In our view, it is only through a consideration of self-processes—identity, self-esteem, and self-regulation—that one can begin to understand Michelangelo's behavior. An analysis of Michelangelo's "self" allows us to formulate hypotheses concerning a heightened sense of perfectionism and shame that likely accompanied his reputed narcissistic tendencies, a failure to live up to his own expectations and those of his father (who equated sculpting with manual labor), a breakdown in self-regulation, and an identity crisis due to his impending death.

Many aspects of human behavior are inexplicable without the notion that people have a self. In fact, an understanding of the self is necessary for a complete understanding of personality processes—the processes that generate and regulate thoughts, feelings, and

behaviors. An understanding of the self helps explain not only such exceptional behaviors as Michelangelo's destructive act, but also many aspects of everyday social life: Why do some individuals feel shy in social contexts whereas others do not? Why are some individuals boastful in some situations but insecure in others? Why are some individuals preoccupied by achievement concerns whereas others crave intimacy?

<u>Psychology's Most Puzzling Puzzle</u>

In *Principles of Psychology*, William James (1890) referred to the self as psychology's "most puzzling puzzle" (p. 330). For the past century, psychologists have debated whether it is a puzzle worth puzzling about. In an article titled *Is the concept of self necessary?*, Allport raised the possibility that the self is "an impediment in the path of psychological progress" (1955, p. 25). Skinner (1990, p. 1209) argued that "There is no place in a scientific analysis of behavior for a mind or self." Pinker (1997) described self-awareness as an intractable problem that we as a species are not sufficiently evolved to grasp. And, Ramachandran (2007) characterized the "problem of self" as "Science's greatest riddle."

Faced with this daunting level of pessimism, we propose the perhaps overly optimistic thesis that a scientific understanding of the self is not only possible but is, in fact, fundamental to a science of personality. Research over the past few decades has documented many ways in which the self influences how people act, think, and feel in particular situations, the goals they pursue in life, and the ways they cope with and adapt to new environments. Many currently prominent areas of personality research assume a central role for the self, including the study of self-conscious emotions such as pride and shame (e.g., Tracy, Robins, & Tangney, 2007), traits such as narcissism (e.g., Morf & Rhodewalt, 2001), internal working models of attachment (e.g., Collins & Allard, 2001), autobiographical memories (e.g., McLean, Pasupathi, & Pals, 2007;

Sutin & Robins, 2005), self-regulation (Gailliot, Mead, & Baumeister, this volume), and goals and motivation (e.g., Carver, Scheier, & Fulford, this volume).

A Naturalist View of the Self

In the early days of scientific psychology, the self was an integral part of many general theories of the person. Indeed, many "classic" readings on the self come from the writings of the most influential theorists of the first half of the 20th century: James (1890), Baldwin (1897), Cooley (1902), Mead (1934), McDougall (1908), Murphy (1947), Hilgard (1949), and Allport (1955). Three basic themes recur in these broad conceptions of the person. First, the self was seen as fundamental to understanding social behavior and personality processes, and many early theorists attempted to link self-processes to other basic psychological processes. Specifically, the self was seen as an executive body coordinating the thoughts, feelings, and behavior of a highly complex, dynamic organism. Second, many of these perspectives emphasized the interplay between biological and social forces—the self is constructed out of the raw materials endowed by nature and shaped by nurture. Third, the self was conceptualized from an evolutionary and functionalist perspective. The early theorists were working in the immediate aftermath of Darwin, and many drew heavily on evolutionary thinking. In particular, James (1890) was committed to a *naturalistic* explanation of the origin and function of the self-awareness, assuming that conscious mental life "emerged by way of natural selection because it gave our species certain survival, and therefore reproductive, advantages" (p. 52).

For most of the latter half of the 20th century, research on the self moved away from these three themes, as researchers came to conceptualize the self as a social and cultural construction. In the past decade, however, there has been renewed interest in a naturalist view of the self, spearheaded by neuroscientists such as Crick (1994), Gazzaniga (1998), Ramachandran (2004),

Koch (2004), Edelman (2005), and others, who study the neural mechanisms underlying consciousness in an attempt to understand how a sense of self emerges from the activity of the brain. Similarly, in their search for the neural bases of affective experience, emotion researchers such as Damasio (2003) and LeDoux (2003)have discovered basic facts about how the brain is wired, which have profound implications for self researchers. Memory researchers have linked various forms of memory to the conscious experience of self (e.g., Tulving, 2005), and this work has been used to examine neural activation during processing of self-relevant information in normal populations (Kelley et al., 2002; Magno & Allan, 2007) and amnesiacs (Klein, Loftus, & Kihlstrom, 1996). Baron-Cohen (2008) and other neurologists (e.g., Feinberg & Paul, 2005) provide vivid illustrations of how neurological disorders such as autism can produce profound deficits in self and identity.

Thus, researchers from a wide range of perspectives outside of the traditional boundaries of personality and social psychology are grappling with basic issues about the self. Moreover, within personality and social psychology, there has been an explosion of recent research linking brain mechanisms to self-related processes. Indeed, many of the topics mentioned above as integrating the study of self and personality, have now been studied from a neuroscience perspective, including the neural bases of pride (Takahashi et al., 2007), attachment (Gillath, Bunge, Shaver, Wendelken, & Mikulincer, 2005), autobiographical memory (Levine, 2004), and self-regulation (Inzlicht & Gutsell, 2007). Building on this emerging biological account, researchers have also discussed the evolutionary origins of the self (Sedikides & Skowronski, 2003) and explored the genetics of self-esteem and other self processes (e.g., Neiss et al., 2005; Neiss, Sedikides, & Stevenson, 2006).

What unites many of these perspectives is a naturalist view of the self—a belief that the self can be studied like any other natural phenomenon. One goal of this chapter is to help self research recover its roots by reconnecting it with broader scientific concerns. We return to a set of foundational issues that preoccupied William James when he formulated his naturalist perspective of the mind.

Overview of Chapter

The literature on the self is enormous. In a recent survey of personality psychologists (Robins, Tracy, & Sherman, 2007), 43% of respondents indicated that they study the self-concept and 35% study self-esteem. A *Psycinfo* search for the keyword "self" identified 265,161 articles.[1] Even restricting the search to the past 30 years and to a single journal—the *Journal of Personality and Social Psychology*—yielded 2,411 articles with the keyword "self" (33% of all articles published in the journal from 1970 to 2007). As Figure 1 shows, research on the self surged in the 1970s—probably reflecting the rise of cognitive psychology and its emphasis on mental events—and then maintained a high level up to the present.

In this chapter, we do *not* attempt to provide a comprehensive review of the voluminous social-personality literature on the self. Several recent books and chapters provide such reviews: see Chen and Pickett (in press), Leary (2004), and Sedikides and Spence (2007) for general overviews of the self literature; Harter (2006) and Robins and Trzesniewski (2005) for research on the development of the self; Tracy, Robins, and Tangney (2007) for research on self and emotions; Gailliot et al. (this volume) for research on self-regulation; Swann (this volume) for research on the role of the self in social interaction; and Byrne (1996) for an overview of self-concept measures. Instead, our goal is to step back and reflect on some broader questions about the self: What is the self? When does the self first emerge and how does it change over time? Is

the self distinctive to humans? How does the brain build a sense of self? Is the self a product of evolution? What are the adaptive functions of the self? These questions return us to James' initial set of concerns when he formulated his naturalist perspective on self and consciousness.

<div align="center">Basic Questions about the Self</div>

What is the Self? Definitional and Conceptual Issues

The question—what is the self?—has been an abiding concern of philosophers, writers, scientists, and laypeople. Although self theorists have offered numerous definitions, there is no consensual framework for conceptualizing the various aspects of the self. Some researchers embrace the theoretical richness of the field, whereas others bemoan its conceptual muddiness. One source of confusion is that the self does not refer to a single entity but rather to myriad structures and processes (Leary, 2004). Consider the example of happiness. One aspect of the self is the feeling or experience of happiness itself. This is what philosophers refer to as qualia or sentience. A second aspect is the conscious recognition that my feeling of happiness belongs to me—it is I who feels happy. As Flanagan (1991) observed, "thoughts, feelings, and the like do not sit around disembodied. All thoughts and feelings are 'owned,' that is, all thoughts and feelings occur to someone." (p. 31). A third aspect refers to attentional focus—I can be aware (i.e., conscious) or not aware of my current state of happiness. For example, I could be feeling happy but not pay attention to this because I am focused on skiing down the mountain. Finally, I can have a stable representation of myself as happy—that is, I can think of myself as a generally happy person, or have the belief that "I always feel happy when I am skiing." These four examples are only a sampling of the many ways the self has been defined in the literature.

Self-awareness and self-representations. In our view, however, all of the definitions boil down to two basic classes of phenomena: (1) an ongoing sense of self-awareness and (2) stable

<div align="center">7</div>

mental representations. These two aspects of the self correspond to James' (1890) classic distinction between the self-as-perceiver (the "I") and the self-as-object of perception (the "Me").

Figure 2 shows some of the phenomena that relate to these two aspects of the self. What unites the different conceptions listed under *ongoing sense of self-awareness* is a view of the self as an active agent that processes information and regulates behavior. Our ongoing sense of awareness is the one psychological phenomenon for which we seem to have direct and irrefutable evidence—we all know what consciousness feels like from the inside. As Farthing (1993) notes, "casual introspection seems to reveal a self: the unique entity that is the perceiver of our perceptions, the thinker of our thoughts, the feeler of our emotions and desires, and the agent of our actions…" (p. 139). Self-awareness refers to a particular form of consciousness in which the object of attention is the self. Thus, I can be conscious that I am talking, but my consciousness becomes self-consciousness when I reflect on the fact that I am not speaking as lucidly as usual, or when I begin to worry that others are evaluating how intelligent I sound.

A second class of self-phenomena involve *stable mental representations* (see Figure 2). The self as mental representation is a product of reflexive activity ("Me"), rather than the reflexive activity itself ("I"). Mental representations can be of the person as a physical, social, psychological, or moral being in the past, present, or future. This is what contemporary researchers mean by a self-concept, and what Sedikides and Skowronski (2003) refer to as the symbolic self, and Leary (2007) as the conceptual self.

Self-representations can vary in their degree of abstractness. At the most specific level are personal or autobiographical memories, such as "I remember feeling shy when I was with this person," which are related to episodic memory. At the next level are semantic representations or generalized knowledge about the self, such as "I am a shy person." Finally, at

8

the most abstract or conceptual level, are theories about the self, including personal narratives (McLean, et al., 2007) and more specific self-theories such as the *entity-theorist* belief that "My shyness is fixed and I will always be shy" and the *incremental-theorist* belief that "I can overcome my shyness if I really try" (Dweck, 1999).

Both self-awareness ("I") and self-representations ("Me") have been conceptualized as personality variables. That is, people vary in the degree to which they are chronically self-aware, and their self-representations vary in content, structure, stability, clarity, and complexity. Thus, there are meaningful individual differences in both aspects of the self, although, as we will argue, the basic capacity for self-awareness and self-representation is common to all humans.

Another important definitional issue concerns self-esteem. In our ongoing sense of self-awareness we are continually evaluating ourselves (the "I" evaluates the "Me"). At the same time, we also form stable mental representations that have an evaluative component. The former can be thought of as the self-evaluation process (e.g., feeling competent while working on a task) whereas the latter can be thought of as self-esteem (e.g., stable representation of the self as competent or likable). This raises the question of what is the "Me" that is being evaluated. From our perspective it is largely (but not exclusively) one's personality, that is, one's tendency to think, feel, and act in certain ways. However, our self-representations may or may not accurately reflect our personality characteristics (Robins & John, 1997).

Structure of the self. Some theorists question the assumption that each person has a unique stable self that represents his/her distinctive personal experiences and traits (Brewer & Chen, 2007; Turner & Onorato, 1999; Tyler, Kramer, & John, 1999). Instead, these theorists see the self as a socially constructed entity, arguing that society enmeshes the individual in multiple specific social contexts, each imposing a role that confers a distinctive sense of self. From this

"social self" perspective, people have multiple selves reflecting their various group memberships and associated identities. Consistent with this view, when answering the question "Who am I?", people come up with a vast array of responses that encompass everything from beliefs about private thoughts and feelings, to one's place in the larger nexus of relationships, social roles, and cultural institutions (Gordon, 1968).

Within psychology, the idea that we have multiple selves dates back to James's (1890) *The Principles of Psychology*, which included a summary table classifying the levels and structures of the self (p. 329). Figure 3 shows our revised and expanded version of this table (see also Brewer & Gardner, 1996; Brown, 1998; Greenwald & Breckler, 1985). The first row shows the *personal* or individual self, which reflects people's beliefs about their private self, including their traits, values, and abilities. The second row shows the *relational* self, which reflects how people see themselves in intimate relationships. The third row shows the *social* self, which reflects how people see themselves in more general interpersonal contexts, including their social roles and reputation. The fourth row shows the *collective* self, which reflects people's identities concerning their various reference groups, such as their religion, ethnicity, and nationality. Some aspects of identity, such as gender, permeate all levels of the self from the personal (e.g., perceptions of feminine characteristics) to the collective (e.g., one's identity as a "feminist").

Interestingly, research suggests that information about the personal self (e.g., "altruistic") may be represented in memory separately (and thus encoded and retrieved separately) from information about the social and collective self (e.g., "peace activist"; Trafimow, Triandis, & Goto, 1991). People derive self-regard differently depending on which level of the self they are representing. For example, when an individual is focused on the personal self, self-esteem is rooted in meeting personal aspiration. However, when an individual is focused on the collective

self, self-esteem is rooted in the accomplishments and prestige of the social groups to which the person belongs (Rubin & Hewstone, 1998).

A great deal of recent research has demonstrated cross-cultural differences in the degree of emphasis on these various levels of the self. Specifically, Western cultures (e.g., United States) tend to be more focused on the personal self, whereas Eastern cultures (e.g., China, Japan) tend to be more focused on the collective self (e.g., Triandis, 1997). These differences appear to be socialized early in life and persist throughout the lifespan (e.g., Wang, 2006). However, it is important not to overemphasize these cultural differences, given the substantial individual differences that exist within groups. For example, even within Western cultures, women have more collectivistic, interdependent selves, whereas men tend to have more personal, independent selves (Josephs, Markus, & Tafarodi, 1992). Moreover, reviews of the literature on individualism-collectivism have failed to show consistent cultural differences and, in some contexts, individuals from Western cultures do not appear to be any more individualistic or less collectivistic than individuals from Asian cultures (del Prado et al., 2007; Oyserman, Coon, & Kemmelmeier, 2002, but see Schimmack, Oishi, & Diener, 2005). For example, Japanese tend to focus on the personal self (e.g., describe themselves using personality traits) to the same extent as Americans when they are asked to provide self-descriptions in specific contexts such as at home (Kanagawa, Cross, & Markus, 2001). Findings such as these have led to calls for refining the collectivistic construct, including the need to distinguish between relational and group aspects of the collective self (Brewer & Chen, 2007).

Nonetheless, the individualism-collectivism distinction does predict differences in self-views. Individuals from East Asian cultures tend to accept and value negative information about the self more than individuals from Western cultures (Oyserman et al., 2002; Spencer-Rodgers,

Peng, Wang, & Hou, 2004). Similarly, individuals from East Asian cultures tend to have lower self-esteem than those from most other cultures (Schmitt & Allik, 2005). Interestingly, this difference is only observed for explicit (i.e., self-report) measures of self-esteem, not implicit measures such as the implicit association test or preferences for the initials in one's own name (Heine & Hamamura, 2007; Yamaguchi et al., 2007). Thus, it is possible that East Asians report lower levels of explicit self-esteem because they are less prone to self-enhancement and instead adopt a more modest self-presentation strategy. A related possibility is that individuals from both cultures engage in self-enhancement biases, but Easterners tend to self-enhance on collectivistic attributes and Westerners tend to self-enhance on individualistic attributes, which are more common on explicit self-esteem scales (Sedikides, Gaertner, & Vevea, 2005).

It appears that the structure of the self is not the same as the structure of personality, as embodied in models such as the Five-Factor Model (FFM; John, this volume; McCrae & Costa, this volume). Clearly our self-representations include far more than just beliefs about personality traits, as research using the "Who am I?" test has revealed. The FFM was not intended as a complete model of personality, but rather as a model of individual personality traits. Thus, relational, social, and collective aspects of the self were not included in the research that led to the discovery of the FFM, and most lexical studies of personality structure have explicitly excluded social roles and relationships. Moreover, many aspects of the self do not reflect content domains such as the Big Five but rather how the content of the self is organized, including the degree of differentiation, complexity, and compartmentalization (Donahue et al., 1993; Rafaeli-Mor, Gotlib, & Revelle, 1999; Showers & Kling, 1996; Suh, 2002). Consequently, although the FFM effectively captures the trait-relevant aspects of the personal self, it fails to

capture many other important aspects of the self, including its organizational structure and features of the relational, social, and collective self.

When Does the Self Emerge and How Does it Change Across the Lifespan?

Our beliefs about ourselves seem to be relatively enduring. When we wake up in the morning we have the sense that we are the same person we were the previous day. It is unlikely that a person will think he or she is shy and introverted one day and socially bold and extraverted the next. Where does this enduring sense of self come from? At what age does it develop?

Infants, and virtually all animals, have at least one basic aspect of self-awareness—they can distinguish self from non-self, and consequently do not eat themselves when they are hungry. However, it is difficult to determine whether infants and young children have more complex forms of self-awareness and self-representations because they cannot clearly communicate what they are thinking and feeling. Instead, researchers must infer the presence of a self from overt behavioral markers. Hart and Karmel (1996) described three classes of evidence for the existence of a sense of self—linguistic markers, cognitive-behavioral markers, and emotional markers.

Linguistic markers include self-referencing (e.g., use of personal pronouns), narrative language use (e.g., reference to events of personal significance that happened in the past), and declarative labeling speech. These behaviors begin to emerge around the second birthday (e.g., Kagan, 1998). For example, starting around age 2, children will label the self as "me" and identify objects as "mine."

Cognitive-behavioral markers include mirror self-recognition, imitation, and role-taking. Mirror self-recognition (assessed in young children with the "Rouge Test") has been interpreted as evidence for objective self-awareness, subjective self-awareness, and self-representation.

When a child sees its image with an unexpected mark on it, the child needs to focus attention on and identify itself in the mirror and become aware that the current image differs from some stable representation of its typical appearance. Children generally pass the rouge test by 18 months. Importantly, self-recognition seems to be rooted in the capacity to construct a psychological, not a physical, representation of the self; Priel and de Schonen (1986) found that Bedouin children without previous exposure to mirrors showed mirror self-recognition at the same age as children habitually exposed to mirrors. However, the stability of the self-representation in young children is limited. Researchers have found, using a version of the rouge test in which stickers were covertly placed on a child, that children younger than 4 can find and remove the sticker from their body when presented with a live video of themselves, but not when the video is delayed as little as 2 seconds, suggesting they do not have a representation of the self as continuous over time (Myazaki & Hiraki, 2006; Povinelli, & Simon, 1998). This research suggests that children younger than 5 are unable to hold a dual representation of the self in mind—that is, a representation of the present and future or past self—suggesting that young children do not gain a sense of personal continuity until around age 5.

Emotional markers include self-conscious emotions and empathy. Self-conscious emotions such as shame, pride, guilt, and embarrassment require a sense of self. Pride, for example, occurs when individuals construe a positive outcome as relevant to their personal goals and aspirations (i.e., relevant to their identity), and as caused by their own actions or attributes (Tracy & Robins, 2004a). Thus, to feel a sense of pride (or, conversely, shame), a child must have stable self-representations that include knowledge of his/her goals and identity concerns, and he/she must have the capacity to make internal attributions (e.g., to understand that "something about me or my actions caused that to happen"). These capacities seem to emerge

around the age of two and a half to three years (Lagattuta & Thompson, 2007). Three-year-olds show signs of pride after success on a difficult, but not easy, task, and signs of shame after failing to complete an easy, but not a difficult, task (Lewis, Alessandri, & Sullivan, 1992). Similarly, after failing a task, children as young as 4 show a lack of persistence, view their failure as indicating that they lack the ability to complete the task, and report low expectations for their future performance (Dweck, 1999). These findings suggest that children as young as three or four can make internal attributions about their failures, which requires the presence of a stable self-representational system.

Overall, these findings suggest that children have the cognitive skills needed to form specific and stable self-representations around age 4. Does this mean that young children also have a stable *evaluation* of their overall worth as a person (i.e., global self-esteem)? It is difficult to address this question because young children have obvious verbal limitations and there is no non-verbal test such as the mirror self-recognition task. Instead, the typical approach is to use self-report measures, sometimes aided by pictures or puppets, in which children are asked to report on their beliefs about their general competence and likableness. This research suggests that children as young as four can provide reliable evaluations of themselves, but only when they are evaluating themselves in specific domains (e.g., Marsh, Ellis, & Craven, 2002; Measelle, John, Ablow, Cowan, & Cowan, 2005). It is not until later in childhood, sometime between age 6 and 9, that children can reliably report on their *global* self-esteem using standard self-report measures, suggesting that younger children lack the cognitive capacity to integrate their domain-specific evaluations into a generalized, global concept of the self (Harter, 2006). We believe this is because global self-esteem requires a representation of the self as a coherent, integrated entity, which may take longer to develop than individual self-representations. However, it is possible

that young children have a rudimentary evaluation of the self as generally good or bad. One study found that 5-year old children who talked positively about themselves in a puppet interview had significantly higher scores at age 8 on a traditional self-report measure of self-esteem (Verschueren, Buyck, & Marcoen, 2001).

Although a coherent sense of self-worth emerges in childhood, it continues to evolve throughout the lifespan. Individuals who have relatively high (or low) self-esteem at one point in time tend to have high (or low) self-esteem years later, but the stability of self-esteem is relatively low in childhood and far from perfect across the entire lifespan (Trzesniewski, Donnellan, & Robins, 2003). Interestingly, self-esteem and personality show similar levels of stability and follow the same developmental trajectory through much of life: lower stability during childhood and increasingly high stability across adulthood. However, in contrast to personality traits, self-esteem becomes less stable in old age; this developmental shift may be due to greater self-reflection, and thus questioning and reformulation of longstanding self-views, as one approaches the end of life.

In addition to examining the stability of individual differences, researchers have also examined whether self-esteem shows aggregate (or mean-level) increases or decreases over time. As we go through life, our self-esteem inevitably waxes and wanes over time, as part of the process of development. These fluctuations in self-esteem reflect changes in our social environment and maturational changes such as puberty and cognitive declines in old age. When these changes are normative, age-dependent, and influence individuals in a similar manner, they will produce aggregate (i.e., population-level) shifts in self-esteem across developmental periods.

A number of aggregate changes in self-esteem occur from childhood to old age (see Robins & Trzesniewski, 2005, for a review). On average, young children have relatively high

16

self-esteem, which gradually declines over the course of childhood. Researchers have speculated

that children have high self-esteem because it is artificially inflated, and that the subsequent

decline reflects an increasing reliance on more realistic information about the self. Self-esteem

continues to decline in adolescence, producing a substantial cumulative drop from childhood to

adolescence. The adolescent decline has been attributed to physical and neurological changes

associated with puberty, cognitive changes associated with the emergence of formal operational

thinking, and socio-contextual changes associated with the transition from grade school to junior

high school (Harter, 2006).

Self-esteem increases gradually throughout adulthood, peaking sometime around the late

60s. Over the course of adulthood, individuals increasingly occupy positions of power and

status, which might promote feelings of self-worth. However, self-esteem declines again in old

age, beginning around age 70. This decline may be due to the dramatic confluence of changes

that occur in old age, including changes in roles (empty nest, retirement), relationships (spousal

loss, decreased social support), and physical functioning (declining health, memory loss, reduced

mobility). The old age decline may also reflect a shift toward a more modest, humble, and

balanced view of the self in old age (Erikson, 1985). That is, as individuals grow old, they may

increasingly accept their faults and limitations and become less concerned about inflating their

self-worth, which artificially boosts reports of self-esteem earlier in life. Consistent with this

interpretation, narcissism also tends to decline with age (Foster, Campbell, & Twenge, 2003).

Overall, men and women follow essentially the same developmental trajectory: both

genders tend to have high self-esteem in childhood, decline during adolescence, rise gradually

throughout adulthood, and then decline in old age. Despite these similarities, there are three

important differences. First, there is a small but consistent gender difference across the lifespan,

such that men report higher levels of self-esteem at almost every age. Second, girls show a much steeper decline in self-esteem during adolescence. The adolescent gender gap has been attributed to maturational changes associated with puberty (e.g., changes in body shape and image) and social-contextual changes associated with the differential treatment of boys and girls in the classroom and in society at large. Third, men show a sharper decline in old age, perhaps because they are more likely to experience, or more adversely affected by, health problems, retirement, and diminished social support.

The absence of perfect stability of individual differences in self-esteem, combined with normative shifts in average self-esteem levels from childhood to old age, compels researchers to search for factors—both psychological and social-contextual—that promote change in the self across the life-span. In our view, the best way to understand self-esteem change is to understand the self-evaluative mechanisms that drive the self system; that is, the processes that underlie the way self-evaluations are formed, maintained, and changed. Although experimental research has linked a number of self-evaluative processes to short-term changes in self-evaluation, little is known about the influence of such processes on self-esteem change over longer periods of time and in real-world contexts.

We have explored how two self-evaluative factors impact self-esteem change: (a) implicit self-theories and (b) positive illusions. In one study, we found that individuals who believe their intelligence is a fixed quantity (i.e., Entity theorists) tend to decline in self-esteem over the course of college relative to those who believe their intelligence can improve (i.e., Incremental theorists; Robins & Pals, 2002); thus, the college experience had an adverse impact on the self-esteem of Entity theorists but bolstered the self-esteem of Incremental theorists. This divergence was mediated by differences in helpless vs. mastery-oriented responses, with Entity theorists

declining in self-esteem because they tend to become helpless in challenging achievement contexts while Incremental theorists become mastery oriented in the same contexts. In a second study, we found that students who entered college with unrealistically positive beliefs about their academic ability had higher self-esteem at first but then decreased in self-esteem over the course of college compared to those with more accurate self-views (Robins & Beer, 2001). Thus, individuals with more objective beliefs about their ability were better able to maintain their self-esteem during college. These two examples illustrate the need to understand the motives and beliefs that regulate self-evaluative processes and ultimately contribute to developmental changes in self-esteem. Conversely, as Roberts, Wood, and Caspi (this volume) have argued, changes in self and identity are one critical mechanism through which changes occur in more basic aspects of personality such as traits.

The previous sections have shown that the self emerges fairly early in life, is relatively stable, and has universal aspects. These conclusions suggest that self-esteem and other aspects of the self may be a product of our evolutionary history. If so, we might expect to see early forms of self in non-human animals with a shared phylogenetic history.

Is the Self Distinctive to Humans?

If omniscient beings from another planet were to study the inhabitants of the earth, what would stand out about the human species? Our use of tools? Our leisure time and range of creative endeavors? Our complex social organization and interactions? Certainly all of these. But what would be perhaps most salient about the human species would be our inner world—the richness of our mental life. Tapping into the inner life of other species is unlikely to reveal the same rich stream of thoughts, feelings, intentions, and so on.

The capacity for self-awareness and self-representations seems to be a universal characteristic of humans. In all human cultures, people have an awareness of their own thoughts and feelings and have relatively stable mental representations of themselves. The universality of these basic aspects of the self is a striking and highly significant fact that is often overlooked in light of the substantial individual and cross-cultural variability that exists in the way the self is manifested.

Whenever universal (i.e., species-typical) characteristics are found, scientists generally proceed to comparative, cross-species studies. The self may be part of human nature, but is it unique to humans? Comparative psychologists have shown that a number of other species show advanced forms of self-awareness, including the same three classes of evidence used to evaluate a sense of self in young children: linguistic, cognitive-behavioral, and emotional markers.

Linguistic markers. Language-trained great apes (chimps, gorillas, orangutans) show all three linguistic markers (self-referencing, narrative language use, and declarative labeling speech). Koko the gorilla, for example, displayed the capacity for reflexive self-referencing and narrative language use when she signed "me love happy Koko there" after being shown a picture of herself at a birthday party (Patterson & Linden, 1981, p. 86). As Koko's statement suggests, gorillas might also have personal memories that serve as the basis for some forms of linguistic self-referencing. Monkeys and non-primates do not show any of these linguistic markers.

Cognitive-behavioral markers. In a series of seminal studies, Gallup (1970) showed that chimpanzees have the capacity to recognize themselves in a mirror. Gallup allowed chimps to view themselves in a mirror for a few days and then marked their faces with dye while they were anesthetized. Upon subsequently seeing their image in the mirror, the chimps often touched the

marks on their face rather than touching the mirror itself. Based on these findings, Gallup (1977) concluded that "Man may not be evolution's only experiment in self-awareness" (p. 14).

In subsequent research, dozens of animal species have been subjected to the mark test, but only chimpanzees (Gallup, 1970), orangutans (Suarez & Gallup, 1981), gorillas (Parker, 1994; Patterson & Cohn, 1994), dolphins (Reiss & Marino, 2001), and Asian elephants (Plotnik, de Waal, Reiss, 2006) have demonstrated the capacity for self-recognition. Interestingly, the first reactions to mirrors by all of these animals are social in nature—smiling, kissing, and vocalizing to their mirror image (Plotnik et al., 2006). Gallup (1977) found that chimpanzees reared in complete isolation from other chimps do not show mirror self-recognition, consistent with Cooley (1902) and Mead's (1934) theory that the self develops through social interaction and the experience of seeing oneself from the perspective of others.

The other two cognitive-behavioral markers—imitation and role-taking—are present in humans by two years of age but extremely rare in other primates. Nonetheless they have been observed in a few cases, suggesting that chimps, orangutans, gorillas, and possibly monkeys have at least some capacity for imitation and role-taking.

Emotional markers. Like human children, non-human primates exhibit behaviors that suggest the capacity to experience self-conscious emotions, and related social emotions such as compassion and empathy. Much of this evidence, however, is anecdotal. In one incident, Supinah, an orangutan, was observed attempting the difficult task of hanging a hammock from two trees. After successfully hanging the hammock, Supinah "threw herself back in the hammock" and "hugged herself with both arms" in apparent pride (Hayes, 1951, p. 188). In the primate literature, researchers have described dominant, or alpha male, chimpanzees walking with a "cocky" gait, upright posture, and pilo-erected fur, suggesting the possibility of a

precursor to the pride display among these animals (de Waal, 1989). However, more controlled studies are needed.

The evidence for empathy in primates is also mostly anecdotal. Non-human primates clearly show helping behaviors. For example, it is not uncommon for a wounded chimp to be attended to—fed, groomed, and protected—by other members of the social group. Dominant adult males even chase away playing infants or noisy group members to keep them from disturbing the injured chimp (Boesch, 1992). A recent series of experiments provides more direct support for empathy and altruism in chimps (Warneken & Tomasello, 2006). After observing an experimenter "drop" an object and act as if she could not reach it, chimps were found to help out by retrieving the object and giving it to the experimenter, suggesting that they understood the experimenter's dilemma and empathized. However, in tasks where the experimenter's goal was more difficult to understand (e.g., trying to get around physical obstacles), chimps failed to provide help whereas children as young as 18 months did. Thus, like other aspects of human psychology—personality, memory, attachment—we share many capacities with our phylogenetic cousins, but we differ in the complexity and frequency with which we display these capacities.

The findings of cross-species research on the self have several implications. First, they illustrate the value of a comparative perspective in personality psychology (Gosling, Capitanio, & Weinstein, this volume). Comparative research helps identify which aspects of the self are uniquely human and which are common across species. Second, the finding that other species share certain aspects of the human self implies that self-awareness and self-recognition may be evolved capacities. Any capacity that is common to several species of primates is probably not crucially dependent on any species-specific factors, such as culture or language.

Correspondingly, the fact that certain aspects of the self, such as the experience of guilt, seem to be unique to humans raises the question of whether these aspects are newly evolved differences, or the products of human culture. Third, comparisons with other species reveal what our conscious mental life might have been like at an early period in our evolutionary history; that is, our evolutionary ancestors may have had the same level of self-awareness as chimps. A phylogenetic perspective suggests that the self may have evolved from the most basic form of self-awareness (distinguishing self and non-self) to more complex forms of self-awareness seen in non-human primates (e.g., self-recognition) to the most complex forms of human self-representations (e.g., identity).

Finally, the existence of cross-species similarities and differences suggests one route to understanding the neural bases of the self. Can we identify brain regions or neural activation patterns that humans share with other animals who also show evidence of having a self?

How Does the Brain Build a Sense of Self?

The self is clearly dependent, in some manner, on the brain. If we did not have a brain, we would not have a self. But, as Hofstadter and Dennett (1981) pondered, "Who, or what, is the you that has the brain?" (p. 5). At the heart of the mind-body debate is the puzzle of how a mass of tissue and the firing of brain cells can possibly produce a mind that is aware of itself and that can experience the color orange, the feeling of pride, and a sense of agency. Historically, the primary philosophical stance was to accept the mind-body distinction and assume that the self is not a physical entity but rather arises from a soul or spirit. For example, van Leeuwenhoek (1632-1723) believed that the brain contains a special vital animal spirit that embodied consciousness.

Most scientists, however, have rejected dualism and generally champion some form of materialism. Francis Crick (1994) illustrates this position: "You, your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules" (p. 3). Similarly, LeDoux (2003) states that "your 'self,' the essence of who you are, reflects patterns of interconnectivity between neurons in your brain" (p. 2) and Ramachandran (2007) asserts that the self is not a holistic property of the entire brain; it arises from the activity of specific sets of interlinked brain circuits." Reflecting the central message of these quotes, Dennett (2005) asserts that, ultimately, a mechanistic approach to consciousness will explain it just as deeply and completely as other seemingly more concrete natural phenomena, such as metabolism and reproduction.

Unfortunately, attempts to understand exactly how the brain builds a sense of self have not been particularly successful and speculative accounts abound. For example, two Nobel laureates—Crick (1994) and Edelman (1989; 2005)—have each provided accounts of the neural substrates of consciousness, but these accounts have little in common. This problem persists in more recent accounts of the neural bases of the self (e.g., Koch, 2004).

Those adopting a computational view of the mind believe that at least some aspects of the self can be explained by neural information processing (Sejnowski, 2003): "Computation has finally demystified mentalist terms. Beliefs are inscriptions in memory, desires are goal inscriptions, thinking is computation, perceptions are inscriptions triggered by sensors, trying is executing operations triggered by a goal." (Pinker, 1997, p. 78). From this perspective, self-representations, self-awareness, self-regulation, and other self processes can be explained through the same neural mechanisms that account for the way the mind encodes, stores,

retrieves, and manipulates information about the world. The computational view of the mind has also spawned the provocative thesis that the self is an illusion – there is no central executive coordinating our thoughts and feelings (Dennett, 2005). Artificial intelligence pioneer Marvin Minsky (1985) describes the mind as a "society of agents"— the agents of the brain are organized hierarchically into nested subroutines with a set of master decision rules. Although it may seem like there is an agent running the "society", in fact it is just the collective action of neural information processing in multiple parts of the brain. Similarly, some researchers believe that consciousness emerges as different groups of neurons—dealing with vision, memory, or touch—are activated. From this perspective, "there is no seat of consciousness, no internal theater where consciousness is a permanent spectator. Instead, what we experience as consciousness is this constant procession of waxing and waning of neuronal groupings" (Greenfield, 1996, p. 159).

Clearly, our current understanding of the neural mechanisms underlying the self is woefully inadequate. If we accept the materialist position and search for the proximate neural mechanisms involved in self processes, we need to go further than grand speculation about a "society of agents" or the "waxing and waning of neuronal groups." We need to approach the problem of how the brain produces a sense of self using the full array of methods used by cognitive neuroscientists. Below we discuss several neuroscientific methods that have been used to study the self and self-related phenomena.

Neuroanatomical studies. An understanding of the gross anatomy of the brain can help us to better understand cross-species comparisons in self processes. For example, what distinguishes the brains of chimps (who show evidence of mirror self-recognition) from the brains of monkeys (who do not)? Which neuroanatomical areas are common to animals that

show the capacity for self-recognition, or linguistic self-referencing, or self-conscious emotions? One clue is that humans begin to show evidence of mirror self-recognition around 18 to 24 months of age, which is when the prefrontal cortex begins to mature in structure and function. Although there are likely many reasons why the neuroanatomy of two species might be similar or different, cross-species comparisons can at least identify possible anatomical regions that merit further exploration.

Functional neurosurgery and brain lesioning. As with other intrusive techniques, the intentional lesioning of a particular brain region cannot be performed on humans. However, researchers could lesion non-human primates to determine whether damage to a particular area eliminates the capacity for mirror self-recognition and other markers of a sense of self. One kind of brain lesioning that traditionally has been performed on humans is functional neurosurgery. The dramatically altered behavior of lobotomized individuals demonstrates how removing brain tissue can destroy what we normally think of as a person's self. Another example is split-brain (commisurotomized) patients, whose corpus collosum has been severed, and who consequently suffer a host of deficits related to self processes (Gazzaniga, 1970).

Although intentional lesioning of humans is unethical, researchers can now create temporary "lesions" using Transcranial Magnetic Stimulation, which uses a powerful yet noninvasive magnet to alter or suppress activity in specific brain regions. One important advantage of TMS is that it provides a stronger basis for making causal inferences than brain imaging techniques that can only demonstrate correlations with neural activation. To date, we know of only one study that has used this promising method to study self processes; Kwan et al. (2007) showed that TMS stimulation (which serves to suppress activity) of the medial prefrontal

cortex reduces the degree to which participants engaged in self-enhancement, defined as perceiving themselves more positively than they perceive others.

    Neurological disorders and brain damage. Some of the most fascinating avenues for understanding the self have come from the study of neurological patients in whom brain damage has produced cognitive, affective, and behavioral deficits. The complete loss of a sense of self is extremely rare, but many neurological disorders, including autism, Alzheimer's, Parkinson's, and epilepsy, produce profound changes in the self (Feinberg & Paul, 2005). There are also a host of neurological disorders characterized by bizarre distortions in body image, including macropsia and micropsia (an Alice in Wonderland like feeling of having either grown incredibly large or incredibly small in size), asomatognosia (e.g., denial that the left side of one's body is part of one's self and assertions that any actions by one's left side were caused by someone else), and exosomesthesia (a pathological extension of the body image in which touches to the self are experienced as touches to nearby physical objects and vice versa.) Sufferers of Lesch-Nyhan syndrome have a pathological tendency toward self-harm, engaging in dramatic self-injurious behaviors such as biting off their fingers and poking their eyes out; they often feel as if their hands and mouth do not belong to them and are under the control of someone or something else, suggesting that their sense of agency – the feeling that "I" am the one controlling my body – is severely distorted.

    Studies of patients with brain damage due to strokes or accidents suggest that the frontal lobes are involved in self-regulation. Phineas Gage, who had frontal lobe damage after an iron rod penetrated his skull, showed a diminished capacity to self-regulate as well as dramatic personality changes – he became more irreverent, obstinate, and impatient, and less persistent and focused. In general, patients suffering from frontal lobe damage show deficits in the

capacity to monitor and reflect on their own mental states (Beer, Shimamura, & Knight, 2004), suggesting that aspects of the self related to self-awareness (the "I") may be associated with activity in the frontal lobes. For example, Pinker (1997) describes the case of a 15-year old boy with frontal lobe damage who would stay in the shower for hours at a time, unable to decide whether to get out. Patients with damage to a specific region of the frontal lobes known as the orbitofrontal cortex often show highly inappropriate social behaviors, which are believed to be caused by impairments in self-insight and in the capacity to experience self-conscious emotions such as embarrassment and to use emotional information in the service of self-regulation (Beer, 2007; Beer, John, Scabini, & Knight, 2006). Patients with damage to the right parietal or prefrontal cortex are unable to recognize themselves in the mirror even after coaching, suggesting that they have lost the capacity for self-recognition (Keenan, Wheeler, Gallup, & Pascual-Leone, 2000).

Another provocative set of studies has examined self-awareness and self-representations in patients with amnesia. Tulving (1993) relates the story of K.C., an amnesic patient who lost the capacity to form new episodic memories. K.C. has a sense of self-awareness but not autonoetic awareness (autobiograpical remembering, or the feeling that one "owns" one's memories)—K.C. may be conscious in a similar way that a dog is conscious, but not in the same way as individuals without neurological damage. Interestingly, K.C. seems to have the capacity to revise his self-representations even without episodic memories of his specific behaviors and experiences: K.C.'s self-descriptions converge with his mother's descriptions of his current personality more closely than with her descriptions of his pre-amnesic personality (Tulving, 1993). Similarly, Klein, Loftus, and Kihlstrom (1996) found evidence that another amnesic patient, W.J., showed stable and seemingly accurate self-ratings of personality over a period of

time during which she lacked the capacity to remember any personally experienced events. These findings suggest that one's current self-representations are not dependent on memories about personal experiences, and thus have important implications for research on the personality judgment process. In particular, they suggest that trait self-ratings reflect different cognitive and neural mechanisms than self-ratings of specific past behaviors.

Despite the promise of patient studies to elucidate self processes, there are several caveats to generalizing from such studies, including (a) non-random assignment (brain-damaged patients may differ on certain personality dimensions such as risk taking); (b) the brain damage associated with neurological disorders or lesions is often highly diffuse making it difficult to pinpoint the specific brain region involved in any observed deficits; and (c) it is impossible to determine whether the damaged region is responsible for receiving or sending the necessary neural signals, or whether the damaged area simply blocks messages from being relayed between two adjacent brain areas that actually regulate the aspect of self-functioning that is showing a deficit (Beer & Lombardo, 2007).

Studies of neural functioning in healthy individuals. Recent advances in the neuroimaging techniques, including methods based on electrical signals (EEG, ERP, MEG) and those based on functional imaging (PET, fMRI), permit more precise measurement of the structure and function of the brain. Although these methods are essentially correlational, and thus unable to elucidate causal relations, they allow researchers to see which part of the brain becomes particularly active when someone performs a cognitive, affective, or behavioral task. Recently, personality and social psychologists have begun to adopt neuroimaging methods to study self-related processes, including self-recognition, self-referential encoding, self-reflection, self-regulation, and self-conscious emotions. Together with patient studies, new studies on the

brain correlates of the self provide converging evidence that several regions within the frontal and temporal lobes, including the medial prefrontal cortex (MPFC), dorsolateral prefrontal cortex, orbitofrontal cortex, and anterior cingulate, are more heavily recruited when individuals engage in self-related processes (Beer, 2007). For example, there is now considerable evidence that the MPFC plays a critical role in self-referential processing (Kelley et al., 2002); specifically, activity in this region is associated with encoding information in reference to the self ("Does conscientious describe you?") but not encoding information in reference to others ("Does conscientious describe George Bush?") or encoding of general meaning ("Does conscientious have the same meaning as responsible?").

In another study on the neural bases of the self, Inzlicht and Gutsell (2007) used EEG to identify neural correlates of the capacity to control and restrain one's impulses. Their findings show that after engaging in one act of self-control (trying to control one's emotions), participants performed worse at a subsequent task requiring self-control and showed decreased activity in a region of the frontal lobes called the anterior cingulate cortex, suggesting that self-regulation depletes rather than strengthens the brain mechanisms that regulate this important aspect of personality. Sharot, Riccardi, Raio, and Phelps (2007) used fMRI to show that enhanced activity in the anterior cingulate cortex and amygdala are associated with optimistic beliefs about the self. Another provocative study showed that activity in the right middle frontal cortex was greater when participants viewed their own rather than familiar faces, and this difference became larger after their "independent selves" were primed by having them read essays containing 1st-person singular pronouns (e.g., I, mine), compared to when their "interdependent selves" were primed by essays containing 2nd-person pronouns (e.g., we, ours) (Sui & Han, 2007).

One interesting question emerging from this research is what happens in the brain when participants are not performing any mental tasks, but are simply being self-aware or self-reflecting. It turns out that a set of brain regions in the frontal, parietal, and medial temporal lobes consistently become active when people let their minds wander and engage in self-reflection, mentally traveling back and forth through time to learn from the past and plan for the future (Mason et al., 2007). Neuroscientists refer to this activity as the brain's "default mode," suggesting that we spend much of our time exploring past and future selves.

Together, these examples illustrate various ways in which the brain can affect self processes, as well as other aspects of personality functioning. The ultimate goal of neuroscience research on the self is to understand how the brain generates self-awareness and self-representations. This goal is complicated because, like other higher-order mental functions, self processes probably emerge out of a complex interplay among multiple brain regions. Moreover, there is a crucial issue of determining the direction of causal relations; for example, does conscious self-reflection cause behavior or is behavior initiated through unconscious brain processes, which the conscious mind then watches and reflects? Finally, the research to date does not conclusively demonstrate that there are any brain states or structures that are *distinctively* linked to self processes; after reviewing the relevant literature, Gillihan and Farah (2005) concluded that there is little compelling evidence for brain networks devoted to the self that are physically and functionally distinct from those used for more general purpose cognitive processing.

At least in principle, new knowledge about the brain can help us refine our theoretical conception of the self and self-related processes. Specifically, the way we conceptualize the self should be consistent with, and constrained by, what we know about how the brain works (e.g.,

memory researchers used to think that a sense of familiarity was simply a weaker form of recognition memory but research suggests that familiarity and recognition involve distinct brain regions). For example, if one function of the self is to coordinate and regulate inner-body signals and behavioral responses (i.e., self-regulation), then the brain should be doing something different when the self is "in control" than when it is "out of control." If one motive governing the self system is self-enhancement, then there must be some neural mechanism that "tags" the valence of a self-relevant event and causes positive events to be encoded more deeply than negative ones; similarly, if there is selective retrieval of positive autobiographical memories, then there should be identifiable neural process that enables this to occur (e.g., preferential pathways to representations tagged as positively valenced).

Clearly, we have a long way to go before we truly understand the neural mechanisms underlying these and other aspects of the self. However, we believe the new methods of brain science hold a great deal of promise and could ultimately lead to discoveries that provide a foundation for a naturalist view of the self.

<u>Is the Self a Product of Evolution?</u>

The eminent geneticist Theodosius Dobzhansky (1964) remarked that the self is the chief evolutionary novelty possessed by humans. Consistent with this view, behavioral genetic studies have documented the heritability of self-esteem and other self processes (e.g., Neiss et al., 2005; Neiss et al., 2006). Evidence of heritability supports an evolutionary account of the self, but there are four possible interpretations. First, the self may be "genetic junk;" a characteristic that neither contributes to nor detracts from the organism's fitness but is nonetheless passed on to succeeding generations. A second possibility is that the self is a functionless by-product of another adaptation, and does not solve any adaptive problems on its own but is carried along

with the more functional characteristic. For example, the self has been described as an incidental by-product of high-level intelligence and complex sensory processing associated with large brains.

However, it is difficult to write off the self as an evolutionary accident or a functionless by-product. As much as any other component of the mind, the self fulfills the criteria of an adaptive design outlined by Williams (1966)—it is universal, complex, reliably developing, well-engineered, and reproduction-promoting. Aspects of the self are clearly universal (i.e., species-typical); although there is individual variability in self-awareness and self-representations, all humans have both capacities. The self is also clearly complex, reliably developing, and, as we will argue subsequently, promotes survival and reproduction.

Yet, even if we accept that the self meets the criteria of adaptive design, it may not be an adaptation in the technical sense of the word. Instead, the self may be an exaptation—"a feature, now useful to an organism, that did not arise as an adaptation for its present role, but was subsequently co-opted for its current function" (Gould, 1991); for example, a fly's wings were originally selected for thermo-regulation but were later used for flying. However, it seems unlikely that the self is an exaptation because the environmental features (e.g., complex social interaction) that likely created selection pressures for a self existed in our ancestral, as well as our current, environment.

Finally, the self may be a full-fledged adaptation that is part of our genetic programming. If the self-related processes serve an adaptive function, then the mechanisms that underlie these processes should be hardwired into the brain. Thus, to the extent that researchers can identify neural mechanisms that seem to support highly specialized self-processes, this work provides further support for an evolutionary account. The strongest version of this account is that just as

the brain has an evolved module governing language acquisition, it may also have a module governing aspects of the self such as self-awareness and self-deception (Pinker, 1997). A more moderate position is that people have a genetic blueprint for the basic parts of the self, but the self is assembled through interaction with the current environment. It is possible that we share with other great apes the same neurologically rudimentary self, but through language development and complex social interaction we simply do more with the raw materials. This position nicely integrates universalist and cultural relativist positions on the self.

The assumption that the self is an adaptation leads to the question: What function does it serve? What reproductive or survival advantage is conferred by the capacity to reflect on one's internal states and form stable self-representations?

Why Do We Have a Self? What Are its Adaptive Functions?

How did the self facilitate survival and reproduction during our evolutionary history? The two fundamental aspects of the self, self-awareness and self-representations, are believed to be adaptive solutions to the complex social problems that emerged when our ancestors began living in large, flexibly structured social groups (Sedikides & Skowronski, 2003). Individuals who survived and reproduced in our ancestral environment were able to navigate an intricate social structure in which they had to deal with dyadic and group-level cooperation; cheating and detection of cheaters; and intergroup and intragroup (particularly intrasexual) competition.

How does the self help an individual solve these adaptive problems? In our view, the various functions of the self can be subsumed within four broad categories: Self-Regulation, Information Processing Filter, Understanding Others, and Identity Processes. Below we describe each category and discuss how it might be linked to adaptive outcomes.

Self-regulation. One of the unique aspects of human nature is that we are goal directed, and not just toward proximal goals like grabbing food in front of us but toward long-term goals like succeeding at work, finding a romantic partner, and being a good, moral person. These longterm goals are represented in the self system as various forms of self-representations – ideal selves ("to be a good father"), possible selves ("to be an artist"), dreaded selves ("to become like my mother"), and so on. These goal representations serve as reference points for self-regulation, motivating us to engage in behaviors that move us toward the attainment of desired identities and away from undesired or feared identities; they function as both goals to be pursued and standards against which outcomes are measured.

Self-awareness also plays a role in self-regulation, providing a sense of volition that facilitates goal-directed behavior, and allowing for evaluation of goal-relevant outcomes (e.g., awareness of discrepancies between actual and ideal selves). Self-awareness also provides a mechanism for greater flexibility of response in a social environment filled with competing and often conflicting goals. Self-awareness allows us to monitor and regulate not only our overt behavioral responses but also our internal responses (e.g., fear, shame, optimism) to external stimuli and circumstances. Together, the two aspects of the self enable us to prioritize and organize goal-directed behavior, amidst a complex and multiply-nested structure of goals and subgoals. Interestingly, recent research suggests that chimpanzees, our close evolutionary cousins, are also able to engage in fairly complex forms of self-regulation – they use various strategies such as distraction to resist the temptation to eat right away when they know they will get more food later on (Evans & Beran, 2007).

Clearly the capacity for self-regulation does not guarantee problem-free and effective pursuit of goals. Humans often engage in self-defeating and maladaptive behaviors. Apparent

failures of self-regulation may tell us a great deal about how the self-system functions, as in the case of self-handicapping, where individuals set themselves up for failure in a way that, somewhat paradoxically, protects self-esteem.

Information processing filter. In a complex social environment it is inefficient for individuals to attend to and encode all of the information that is constantly bombarding them. The self addresses this dilemma by serving as a filter, or lens, through which the world is experienced. Our self-representations consist of cognitive structures, or schemas, that organize and direct processing of information about the self. Thus, the self serves as a top-down information filter that is guided by four basic motives: accuracy, consistency, popularity (i.e., social status and approval), and positivity. These motives influence which information the self attends to, encodes, retrieves, and acts upon. We have described these motivational orientations in terms of four basic metaphors (Robins & John, 1997).

According to the *Scientist* metaphor, individuals are driven to acquire *accurate* information about themselves and the world (Bem, 1972; Kelly, 1955; Trope, 1986). Just as the scientist develops empirically based theories, people use facts and observations to develop theories about themselves, engaging in a dispassionate search for accurate self-knowledge. Clearly accurate self-representations can serve an adaptive function, helping us to formulate realistic goals and act in accordance with our actual social status, mate value, and other objectively based self-representations. However, as Pinker (1997) notes, "our brains were shaped for fitness, not for truth" (p. 305); consequently, we sometimes adopt somewhat biased information processing strategies .

According to the *Consistency Seeker* metaphor, individuals strive to see themselves in a *consistent* manner, confirming their preexisting self-views regardless of reality (Swann, 1997;

this volume). In fact, there is considerable evidence that people actively seek out and create

contexts in which their self-views will be confirmed, even when these views are inaccurate

and/or negative. Similarly, people selectively remember life events that are consistent with

current self-representations, reconstructing their past to fit the present (Ross, 1989). A.

Although consistency seeking may lead to information processing errors, it can be a useful and

efficient heuristic in a highly chaotic social environment. Consistency also serves an

interpersonal function, ensuring that people will honor the identities they negotiated in previous

social interactions and act similarly over time.

According to the *Politician* metaphor, people strive to present themselves in ways that

create the most favorable impressions on others, thereby enhancing their *social status and*

acceptance. This perspective highlights the reciprocal nature of social interaction: social reality

is constructed and negotiated through interactions with others, in which behaviors represent

public performances that "present images of the self for the social world to see and evaluate"

(Schlenker, 1985, p. 21). Like politicians, people target their public performances to different

audiences (or constituencies), which place multiple and often conflicting demands and

expectations on them. The person-as-politician seeks to "maintain the positive regard of

important constituencies to whom he or she feels accountable" (Tetlock, 1992, p. 332), which

should increase status, reduce conflict, and facilitate coalition building.

Finally, according to the *Egotist* metaphor, people narcissistically distort information to

enhance their self-worth. Virtually every self theory posits some variant of the motive to protect

and enhance self-worth, and a large body of research has documented numerous positivity biases

in self-perception, including unrealistically positive self-conceptions, self-serving attributions for

success and failure, and excessive optimism about the future (Dunning, 2005; Taylor & Brown,

1988). These self-enhancement biases may facilitate goal striving, emotional well-being, mate attraction, and other adaptive behaviors, at least in the short-term (e.g., Lockard & Paulhus, 1988; Robins & Beer, 2001; Taylor & Brown, 1988). For example, in terms of mate selection, evolutionary psychologists view self-esteem as a way of gauging our value to prospective partners (e.g., our mate value). Individuals with higher self-perceived mate value may demand more in a partner and consequently pair up with partners who have higher mate value.

Together the four motives that drive processing of self-relevant information—accuracy, consistency, social status/acceptance, and enhancement—provide a flexible arsenal of tools that help us (and presumably our evolutionary ancestors) adapt to a complex, multi-structured social environment. The ideal mind from an evolutionary perspective would be able to convince itself that it is better, smarter, and faster than it really is when this facilitates survival and reproduction, but switch to reality mode when needed for increased fitness. Consistent with this perspective, Swann and Schroeder (1995) proposed that the various self-evaluative motives can be organized into a hierarchical system in which different motives are prioritized at different stages in the processing of self-relevant information; specifically, positivity or enhancement may drive the first stage ("does it make me feel good?"), consistency the second stage ("is it consistent with how I see myself?"), and more deliberate and effortful cost-benefit analyses the third stage ("is it accurate; does it facilitate my social goals?). This possibility makes sense from an evolutionary perspective and shows how processing of information about the self may indeed reflect the workings of a specialized adaptive design.

Understanding others' minds. In a complex social environment, survival and reproduction depend in part on the ability to explain, predict, and manipulate others. Children with autism show deficits in the ability to understand what other people know, want, or feel, and

correspondingly they have dramatically diminished social interaction skills (Baron-Cohen, 2008). The capacity for self-awareness facilitates introspectively based social strategies such as empathy, sympathy, gratitude, deception and pretense. Some theorists have even argued that subjective awareness evolved for the specific purpose of helping us to understand others (e.g., Leary, 2007). For example, children may learn to understand others by reflecting on their own internal states, feelings, and intentions, and simulating what might be happening in the mind of others (e.g., Harris, 1992). Consistent with these views, recent research suggests that people use the same neural circuits to understand themselves as they use to understand others (Decety & Jackson, 2006).[2]

The capacity to reflect on our internal states and feelings and project them onto others contributes to another capacity: the experience of empathy and self-conscious emotions, which require the ability to evaluate one's self from the perspective of actual or imagined (i.e., internalized) others. Self-conscious emotions are assumed to have evolved because they motivate individuals to protect, defend, and enhance their social reputation and self-image by engaging in behaviors that facilitate social status and acceptance and avoid social rejection (Keltner & Buswell, 1997; Tracy & Robins, 2007b). For example, researchers have argued that embarrassment and shame evolved for purposes of appeasement and avoidance of social approbation, guilt for encouraging communal relationships, and pride for attaining social dominance (Gilbert, 2007; Keltner & Buswell, 1997; Miller, 2007; Tangney & Dearing, 2002; Tracy & Robins, 2004b).

Self-conscious emotions guide individual behavior by compelling us to do things that are socially valued and to avoid doing things that lead to social approbation. We strive to achieve, to be a "good person," or to treat others well because doing so makes us proud of *ourselves*, and

failing to do so makes us feel guilty or ashamed of *ourselves*. Society tells us what kind of person we should be; we internalize these beliefs in the form of actual and ideal self-representations; and self-conscious emotions motivate behavioral action toward the goals embodied in these self-representations. Thus, although we might understand cognitively that working hard is a good thing to do, it sometimes takes the psychological force of emotions like guilt and pride to make us do so. By reinforcing adaptive social behaviors—encouraging us to act in ways that promote social status (getting ahead) and acceptance (getting along)—self-conscious emotions facilitate interpersonal reciprocity, a social arrangement that is highly beneficial in the long term (Trivers, 1971). In summary, self-conscious emotions help us thrive in a social world where attaining status and acceptance is essential to our ability to survive and reproduce. As Kemeny, Gruenwald, and Dickerson (2004) stated, emotions like shame and pride "may be one way that individuals feel their place in the social hierarchy" (p. 154).[3]

Although self-conscious emotions can be linked to adaptive social behaviors they can also be maladaptive. For example, the tendency to become anxiously preoccupied in social situations and excessively worried about being negatively evaluated can contribute to decreased social competence in the form of shyness. Yet shyness can also be functional when it motivates preparation and rehearsal for important interpersonal events, such as planning ahead for the first day of teaching a new class. Moreover, a moderate amount of wariness regarding strangers and unfamiliar or unpredictable situations undoubtedly had adaptive value in our evolutionary history; Wilson, Coleman, Clark, and Biederman (1993) have argued, based on their research on pumpkin sunfish, that it is adaptive for all species to have a mix of shyness and boldness.

The self-conscious emotion of pride also seems to be adaptive in some contexts and maladaptive in others. In the Greek myth, Narcissus ultimately dies from his excessive pride.

From an evolutionary perspective, he acted in a particularly maladaptive manner because he spent all of his time gazing at his own reflection and ignored the love of a beautiful nymph with who he could have produced offspring. Research suggests that narcissistic pride, in the form of inflated beliefs about the self, can have short-term adaptive benefits but longterm negative consequences (Robins & Beer, 2001). Moreover, narcissistic individuals are more inclined to cheat on their partners (Hunyady, Josephs, & Jost, 2007), which could lead to lower relationship stability but also to higher numbers of offspring.[4]

One way to resolve the seeming paradox of pride's combination of beneficial and detrimental effects is to distinguish between two facets of pride. Several lines of research provide converging support for conceptualizing pride in terms of a "hubristic" or narcissistic facet (defined by terms such as "arrogant" and "conceited") and an authentic or achievement-based facet (defined by terms such as "confident" and "accomplished"; Tracy & Robins, 2007c). These two facets do not simply reflect good vs. bad, high vs. low arousal, or trait vs. state aspects of pride. Moreover, they are not distinguished by the *kinds* of events that elicit the pride experience; both occur after success in a range of domains (e.g., academics, romantic relationships). Rather, it is the way in which success is appraised that determines which facet of pride occurs; successes attributed to one's effort and hard work tend to promote authentic pride, whereas successes attributed to more stable (and less controllable) abilities tend to promote hubristic pride.

Authentic and hubristic pride have highly divergent personality correlates. Authentic pride is positively associated with adaptive traits like Extraversion, Agreeableness, Conscientiousness, and genuine self-esteem, whereas hubristic pride is negatively related to these traits but positively associated with self-aggrandizing narcissism, shame-proneness, and

aggression. This pattern suggests that authentic pride is the more pro-social, adaptive facet of the emotion (Tracy, Robins, Cheng, & Trzesniewski, in press).

The examples of shyness and narcissistic pride illustrate the complexities of the evolutionary perspective—certain aspects of the self may be adaptive in some ways but maladaptive in others. These opposing selection pressures lead to individual differences. Shyness and narcissism may be two ways of approaching the conflict between the dominant social goals of getting along and getting ahead—shy individuals have chosen to focus on getting along and seeking approval whereas narcissistic individuals have focused on getting ahead (Roberts & Robins, 2000). Each may be a viable strategy from an evolutionary perspective. In fact, the two facets of pride may solve unique adaptive problems regarding the acquisition of status. Authentic pride might motivate behaviors geared toward long-term status attainment, whereas hubristic pride may provide a "short-cut" solution, granting status that is more immediate but fleeting, and, in some cases, unwarranted. A related possibility is that the second facet (hubristic pride) evolved as a "cheater" attempt to convince others of one's success even in the absence of real accomplishments (Tracy & Robins, 2007a).

Identity processes. Human social life may be viewed as a series of games—the rules are reflected in cultural norms and the parts that individuals play are defined by their social roles. Winning this game requires that humans form dyadic and group coalitions, and generally navigate within a social structure that, more so than any other species, has complex layers of multiple, overlapping, and sometimes non-transitive social hierarchies (e.g., the highest status hunters were not always the highest status warriors). Imagine living in such a complex social environment without a self—that is, without any stable awareness of your position in the social structure and the roles you play in different contexts and with different interaction partners.

As discussed earlier, people's self-representations are comprised of multiple identities—personal, relational, social, and collective. All forms of identity allow us to differentiate ourselves from others, provide a sense of continuity and unity over time, and help us adapt to and navigate complex social structures and hierarchies by prescribing specific values and role-appropriate behaviors. Social identities also facilitate identification with the social group to which a person belongs. In any social group, the young, low status members are tempted to defect to other groups. A sense of identity, and associated in-group biases and out-group derogation, may help keep individuals from leaving their social group and disrupting their kinship network. Finally, stable identities are also efficient. It is more adaptive to have social interactions that are predictable, structured, and even ritualized, and to have identities internalizing the rules of each social context so that individuals do not have to relearn their social roles each day. In some ways, the self provides a bridge between the individual (and his/her personality characteristics) and the collective (and its associated social roles). For example, by eliciting collective feelings of pride when the group with which one identifies has an achievement (e.g., in the Olympic Games, or a high school football team), the self promotes solidarity among group members and helps reinforce the social inclusion of each proud group member.

In summary, we are proposing that the two aspects of the self—self-awareness and self-representations—are evolved mechanisms that serve four adaptive functions: self-regulation, information processing filter, understanding others, and identity formation. It seems plausible that these four functions helped our evolutionary ancestors survive, reproduce, and attain social status and acceptance in a complex social environment characterized by longterm kinship relationships. Yet one may question what an evolutionary perspective, with its emphasis on

ultimate function, can contribute beyond more proximate functional accounts. Thus, one challenge facing researchers working toward a naturalist account of the self is to provide more precise empirical demonstrations of how the specific functions of the self enhance aspects of fitness such as reproductive success.

Toward A Naturalist Approach to Understanding the Self

We hope that this review of theory and research on the self has demonstrated that our understanding of many personality processes would be impoverished without the concept of self. Personality psychology is an unusually broad field because it covers a wide spectrum of phenomena and levels of analysis, from genetic markers of behavioral traits to neural mechanisms underlying emotions, to lexical studies of trait adjectives, to motives in personal life stories, to sociocultural perspectives on the formation of values. What provides coherence to these diverse themes is an emphasis on understanding consistencies in people's thoughts, feelings, and behaviors, and the mechanisms that underlie these consistencies. But, it is the self that ties together these various personality processes and, as Allport (1960) aptly put it, "makes the system cohere in any one person" (p. 308).

In this chapter we have attempted to outline a naturalist approach to the self. We have reviewed the current state of the field with regard to several fundamental questions concerning the structure, development, and function of the self. Our review of the literature was guided by a particular stance toward the self. In particular, we believe that research on the self should be: (a) central to any theory of personality; (b) informed by an evolutionary perspective and organized around functionalist explanations; (c) informed by comparative, cross-species research; and (d) linked to basic psychological processes such as attention, memory, and emotion, and their associated neural mechanisms. Although the self continues to be "puzzling puzzle," we believe

that much progress is being made in the field and that a scientific understanding of the self is fundamental to a science of personality.

The psychology of the self has an important role to play in the integration of evolutionary biology and neuroscience into personality psychology. The self sits in a privileged position, encompassing and integrating all levels of the person from the biological to the social. This privileged position is fundamentally inclusionary: there is ample room, and indeed serious need, for a variety of approaches to understanding the structure and function of the self, and its relation to other psychological processes.

An evolutionary perspective on the self was central to many early theories of personality and social behavior, and it must be considered a central issue for contemporary personality theories. The naturalist agenda outlined by James (1890) remains a worthwhile path for the next century of research on the science of the self. By naturalizing the self, we move the field of personality toward a truly biosocial perspective.

References

Allport, G.W. (1955). Is the concept of self necessary? *Becoming: Basic considerations for a psychology of personality* (Chapter 10). New Haven, CT: Yale University Press.

Allport, G. W. (1960). The open system in personality theory. *Journal of Abnormal and Social Psychology*, *61*, 301-310.

Baldwin, J. M. (1897). *Social and ethical interpretations in mental development*. New York: Macmillan.

Baron-Cohen, S. (2008). *Autism and Asperger's syndrome (the facts)*. Oxford, UK: Oxford University Press.

Beer, J. S. (2007). Neural systems of self-conscious emotions and their underlying appraisals. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds), *The self-conscious emotions: Theory and research* (pp. 53-67). New York: Guilford.

Beer, J. S., John, O. P., Scabini, D., & Knight, R. T. (2006). Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion-cognition interactions. *Journal of Cognitive Neuroscience*, *18*, 871-879.

Beer, J.S., & Lombardo, M. V. (2007). Patient and neuroimaging methodologies in the study of personality and social processes. In R. W. Robins, R.C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology*. New York: Guilford.

Beer, J. S., Shimamura, A. P., & Knight, R. T. (2004). Frontal lobe contributions to executive control of cognitive and social behavior. In M. S. Gazzaniga (Ed.), *The Newest Cognitive Neurosciences* (3rd Edition) (pp.1091-1104). Cambridge: MIT Press.

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.

Boesch, C. (1992). New elements of a theory of mind in wild chimpanzees. *Behavioral and Brain Sciences*, *15*, 149-150.

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

Brewer, M. B., & Chen, Y. (2007). Where (Who) are collectives in collectivism? Toward conceptual clarification of individualism and collectivism. *Psychological Review, 114,* 133-151.

Brown, J. D. (1998). *The self*. New York: McGraw-Hill.

Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation*. Washington, DC: American Psychological Association.

Chen & Pickett (in press). *The self*.

Collins, N. L., & Allard, L. M. (2001). Cognitive representations of attachment: the content and function of working models. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell Handbook of Social Psychology: Vol. 2. Interpersonal Processes (pp. 60-85)*. UK: Blackwell.

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

Crick, F. H. C. (1994). *The astonishing hypothesis: The scientific search for the soul*. New York: Scribner's Sons.

Damasio, A. R. (2003). Feelings of emotion and the self. In J. Ledoux, J. Debiec, & H. Moss, *The self: From soul to brain, Annals of the New York Academy of Sciences*, *1001*, 253-261.

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Psychological Science*, *15*, 54-58.

Dennett, D. C. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT press.

Dobzhansky, T. (1964). *Heredity and the nature of man*. New York: New American Library.

Donahue, E. M., Robins, R. W., Roberts, B. W., & John, O. P. (1993). The divided self: Concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation. *Journal of Personality and Social Psychology*, *64*, 834-846.

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.

Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

Edelman, G. M. (2005). *Wider than the sky: The phenomenal gift of consciousness*. New Haven, CT: Yale University Press.

Evans, T. A., Beran, M. J. (2007). Chimpanzees use self-distraction to cope with impulsivity. *Biology Letters.* doi:10.1098/rsbl.2007.03.

Farthing, G. W. (1993). The self paradox in cognitive psychology and Buddhism: Review of "The embodied mind: Cognitive science and human experience." *Contemporary Psychology, 38*, 139-140. Cambridge, MA: MIT Press.

Feinberg, T. E., & Paul, J. (Ed) (2005). *The lost self: Pathologies of the brain and identity*. Oxford, UK: Oxford University Press.

Flanagan, O. J. (1991). *The science of the mind*. Cambridge, MA: MIT Press.

Foster, J.D., Campbell, W.K., & Twenge, J.M. (2003). Individual differences in narcissism: Inflated self-views across the lifespan and around the world. *Journal of Research in*

Personality, 37, 469–486.

Gallup, G. G. (1970). Chimpanzees: Self-recognition. *Science, 167*, 86-87.

Gallup, G. G. (1977). Self-recognition in primates: A comparative approach to the bidirectional properties of consciousness. *American Psychologist, 32*, 329-338.

Gazzaniga, M. S. (1970). *The bisected brain*. New York: Appleton-Century-Crofts.

Gazzaniga, M. S. (1998). *The mind's past*. Berkeley, CA: University of California Press.

Gilbert, P. (2007). The evolution of shame as a marker for relationship security: A biopsychosocial approach. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 283-309). New York: Guilford.

Gillath, O., Bunge, S. A., Shaver P. R., Wendelken, C., & Mikulincer, M. (2005). Attachment-style differences in the ability to suppress negative thoughts: Exploring the neural correlates. *Neuroimage (Special Issue - Social Cognitive Neuroscience)*, *28*, 835-847.

Gillihan, S.J., & Farah, M.J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin, 131*, 76-97.

Gordon, C. (1968). Self-conceptions: Configurations of content. In C. Gordon & K. J. Gergen (Eds.), *The self in social interaction* (pp. 115-136). New York: Wiley.

Gould, S. J. (1991). Exaptation: A crucial tool for evolutionary psychology. *Journal of Social Issues, 47*, 43-65.

Greenfield, S. A. (Ed.) (1996). *The human mind explained*. New York: Holt.

Greenwald, A.G., & Breckler, S.J. (1985). To whom is the self presented? In B.R. Schlenker (Ed.), *The self and social life* (pp. 126-145). New York: McGraw-Hill.

Harris, P. L. (1992). From simulation to folk psychology: The case for development. Special Issue: Mental simulation: Philosophical and psychological essays. *Mind and language*, *7*, 120-144.

Hart, D., & Karmel, M. P. (1996). Self-awareness and self-knowledge in humans, apes, and monkeys. In A. E. Russon, K. A. Bard, S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (p. 325-347). Cambridge University Press, Cambridge, England.

Harter, S. (2006). The self. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 646-718). New York: Wiley.

Hayes, C. (1951). *The ape in our house*. New York: Harper and Brothers.

Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social Psychology Review*, *11*, 4-27.

Hilgard, E. R. (1949). Human motives and the concept of the self. *American Psychologist, 4*, 374-382.

Hofstadter, D. R. & Dennett, D. C. (1981). Introduction. *The Mind's I*. New York: Bantam.

Hunyady, O., Josephs, L., & Jost, J. T. (2007). Priming the primal scene: Betrayal trauma, narcissism, and attitudes toward sexual infidelity. *Self and Identity*.

Inzlicht, M., & Gutsell, J. N. (2007). Running on empty: Neural signals for self-control failure. *Psychological Science*, *18*, 933-937.

James, W. (1890). *The principles of psychology*. Cambridge, MA: Harvard University.

Josephs, R. A., Markus, H. R., & Tafarodi, R. W. (1992). Gender and self-esteem. *Journal of Personality and Social Psychology, 63*, 391-402.

Kagan, J. (1998). Is there a self in infancy? In M. Ferrari and R. J. Sternberg (Eds.), *Self-awareness: Its nature and development*, (pp. 137-147). New York: Guildford Press.

Kanagawa, C., Cross, S. E., & Markus, H. R. (2001). "Who am I?" The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin*, *27*, 90-103.

Keenan, J. P., Wheeler, M. A., Gallup, G. G., & Pascual-Leone, A. (2000). Self-recognition and the right prefrontal cortex. *Trends in Cognitive Sciences*, *4*, 338-344.

Kelly, G. A. (1955). *The psychology of personal constructs*. New York: Norton.

Kelley, W.M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785-794.

Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin, 122*, 250-270.

Kemeny, M. E., Gruenewald, T. L., & Dickerson, S. (2004). Shame as the emotional response to threat to the social self: Implications for behavior, physiology, and health. *Psychological Inquiry*, *15*, 153-160.

Klein, S. B., Loftus, J., & Kihlstrom, J. F. (1996). Self-knowledge of an amnesic patient: Toward a neuropsychology of personality and social psychology. *Journal of Experimental Psychology: General*, *125*, 250-260.

Klein, S. B., Cosmides, L., Murray, E. R., & Tooby, J. (2004). On the acquisition of knowledge about personality traits: Does learning about the self engage different mechanisms than learning about others? *Social Cognition*, *22,* 367-390.

Koch, C. (2004). *The quest for consciousness: A neurobiological approach.* Woodbury, NY: Roberts and company.

Kwan, V. S. Y., Barrios, V., Ganis, G., Gorman, J. Lange, C., Kumar, M., Shepard, A., & Keenan, J. P. (2007). Assessing the neural correlates of self-enhancement bias: a transcranial magnetic stimulation study. *Experimental Brain Research, 182*, 379-385.

Lagattuta, K. H., & Thompson, R. T. (2007). The development of self-conscious emotions: Cognitive processes and social influences. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 91-113). New York: Guilford.

Leary, M. R. (2004). What is the self? A plea for clarity. *Self and Identity*, *3*, 1-3.

Leary, M. R. (2007). How the self became involved in affective experience: Three sources of self-reflective emotions. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 38-52). New York: Guilford.

LeDoux, J. (2003). *Synaptic self: How our brains become who we are*. New York: Penguin.

Levine, B. (2004). Autobiographical memory and the self in time: Brain lesion effects, functional neuroanatomy, and lifespan development. *Brain and Cognition*, *55*, 54-68.

Lewis, M., Alessandri, S. M., & Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child Development, 63,* 630–638.

Lockard, J. S., & Paulhus, D. L. (1988). *Self-deception: An adaptive mechanism?* Englewood Cliffs, NJ: Prentice-Hall.

Magno, E., & Allan, K. (2007). Self-reference during explicit memory retrieval: An event-related potential analysis. *Psychological Science, 18*, 272-277.

Marsh, H. W., Ellis, L. A., & Craven, R. G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental Psychology, 38,* 376-393.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*, 393-395.

Mclean, K. C., Pasaputhi, M., & Pals, J. L. (2007). Selves creating stories creating selves: A process model of self development. *Personality and Social Psychology Review*, *11*, 262-278.

McDougall, W. (1963; originally published 1908). *An introduction to social psychology* (31st edition). London: Methuen.

Mead, G. H. (1934). *Mind, self, and society from the standpoint of a social behaviorist.* Chicago: University of Chicago Press.

Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (2005). Can children provide coherent, stable, and valid self-reports on the Big Five dimensions? A longitudinal study from ages 5 to 7. *Journal of Personality and Social Psychology*, *89*, 90-106.

Miller, R. S. (2007). Is embarrassment a blessing or a curse? In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 245-262). New York: Guilford.

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry, 12*, 177-196.

Minsky, M. (1985). *The society of mind.* New York: Simon & Schuster.

Murphy, G. (1947). *Personality: A biosocial approach to origins and structure*. New York: Harper.

Myazaki, M., & Hiraki, K. (2006). Delayed intermodal contingency affects young children's recognition of their current self. *Child Development, 77*, 736-750.

Neiss, M. B., Sedikides, C., & Stevenson, J. (2006). Genetic influences on level and stability of self-esteem. *Self and Identity*, *5*, 247-266.

Neiss, M. B., Stevenson, J., Sedikides, C., Kumashiro, M., Finkel, E. J., & Rusbult, C. E. (2005). Executive self, self-esteem, and negative affectivity: Relations at the phenotypic and genotypic level. *Journal of Personality and Social Psychology*, *89*, 593-606.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3-72.

Parker, S. T. (1994). Incipient mirror self-recognition in zoo gorillas and chimpanzees. In S. T. Parker, R. W. Mitchell, M. L. Boccia (Eds.), *Self-awareness in animals and humans: Developmental perspectives* (p. 301-307). New York: Cambridge University Press.

Patterson, F. G. P., & Cohn, R. H. (1994). Self-recognition and self-awareness in lowland gorillas. In S. T. Parker, R. W. Mitchell, M. L. Boccia (Eds.), *Self-awareness in animals and humans: Developmental perspectives* (p. 273-290). New York: Cambridge University Press.

Patterson, F., & Linden, E. (1981). *The education of Koko*. New York: Holt, Rinehart, & Winston.

Paulhus, D. L., & Levitt, K. (1987). Desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology*, *52*, 245-259.

Pinker, S. (1997). *How the mind works*. New York: W. W. Norton.

Plotnik, J. M., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences, 103*, 17053-17057.

Povinelli, D. J., & Simon, B. B. (1998). Young children's understanding of briefly versus extremely delayed images of self: Emergence of an autobiographical stance. *Developmental Psychology, 34*, 188-194.

Priel, B., & de Schonen, S. (1986). Self-recognition: A study of a population without mirrors. *Journal of Experimental Child Psychology*, *41*, 237-250.

Ramachandran, V. S. (2004). A brief tour of human consciousness. New York: Pi Press.

Ramachandran, V. S. (2007). The neurology of self-awareness [1.8.07]. *Edge* (www.edge.org).

Reiss, D., & Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences, 98*, 5937-5942.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and longterm costs. *Journal of Personality and Social Psychology*, *80*, 340-352.

Robins, R. W., & John, O. P. (1997). The quest for self-insight: Theory and research on accuracy and bias in self-perception. In R. Hogan, J. Johnson, and S. Briggs (Eds.), *Handbook of Personality Psychology*, (pp. 649-679). New York: Academic Press.

Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, *1*, 313-336.

Robins, R. W., Tracy, J. L., & Sherman, J. W. (2007). What makes a personality psychologist? A survey of journal editors and editorial board members. In R. W. Robins, R. C.

Fraley, and R. F. Krueger (Eds), *Handbook of research methods in personality psychology* (pp. 673-678) New York: Guilford.

Robins, R. W., & Trzesniewski, K. H. (2005). Self-esteem development across the lifespan. *Current Directions in Psychological Science, 14*, 158-162.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341-357.

Rubin, M., & Hewstone, M. (1998). Social identity theory's self-esteem hypothesis: A review and some suggestions for clarification. *Personality and Social Psychology Review*, *2*, 40-62.

Sapolsky, R. M. (1997). *The trouble with testosterone: and other essays on the biology of the human predicament*. New York: Scribner.

Schimmack, U., Oishi, S., Diener, E. (2005). Individualism: A valid and important dimension of cultural differences between nations. *Personality and Social Psychology Review*, *9*, 17-31.

Schlenker, B. R. (1985). Introduction: Foundations of the self in social life. In B. R. Schlenker (Ed.), *The self and social life*. New York: McGraw-Hill.

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, *89*, 623-642.

Sedikides, C., Gaertner, L., & Vevea, J.L. (2005). Pancultural self-enhancement reloaded. *Journal of Personality and Social Psychology*, *89*, 539-551.

Sedikides, C., & Skowronski, J. J. (2003). Evolution of the symbolic self: Issues and prospects. In M. R. Leary and J. P. Tangney (eds): *Handbook of Self and Identity* (pp. 594-609). New York, NY: Guilford.

Sejnowski, T. (2003). The computational self. In J. Ledoux, J. Debiec, and H. Moss, The self: From soul to brain. *Annals of the New York Academy of Sciences*, *1001*, 262-271.

Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, *450*, 102-105.

Showers, C. J., & Kling, K. C. (1996). The organization of self-knowledge: Implications for mood regulation. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect, and self-regulation* (pp. 151-173). Mahwah, NJ: Erlbaum.

Skinner, B. F. (1990). Can psychology be a science of the mind? *American Psychologist*, *45*, 1206-1210.

Spencer-Rodgers, J., Peng, K., Wang, L., & Hou, Y. (2004). Dialectical self-esteem and east-west differences in psychological well-being. *Personality and Social Psychology Bulletin, 30*, 1416-1432.

Suarez, S. D., & Gallup, G. G. (1981). Self-recognition in chimpanzees and organutans, but not gorillas. *Journal of Human Evolution*, *10*, 175-188.

Suh, E. M. (2002). Culture, identity consistency, and subjective well-being. *Journal of Personality and Social Psychology, 83*, 1378–1391.

Sui, J., & Han, S. (2007). Self-construal priming modulates neural substrates of self-awareness. *Psychological Science*, *18*, 861–866.

Sutin, A. R., & Robins, R. W. (2005). Continuity and correlates of emotions and motives in self-defining memories. *Journal of Personality*, *73*, 793-824.

Swann, W. B., Jr. (1997). The trouble with change: Self-verification and allegiance to the self. *Psychological Science*, *8*, 177-180.

Swann, W. B., & Schroeder, D. G. (1995). The search for beauty and truth: A framework for understanding reactions to evaluations. *Personality and Social Psychology Bulletin*, *21*, 1307-1318.

Takahashi, H., Matsuura, M., Koeda, M., Yahata, N., Suhara, T., Kato, M., & Okubo, Y. (2007). Brain activations during judgments of positive self-conscious emotion: Pride and joy. *Cerebral Cortex*, *xx*, xx-xx.

Tangney, J. P., & Dearing, R. L. (2002). *Shame and guilt*. New York: Guilford.

Taylor, S. E., & Brown, J. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.

Tracy, J. L., & Robins, R. W. (2004a). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry, 15*, 103-125.

Tracy, J. L., & Robins, R. W. (2004b). Show your pride: Evidence for a discrete emotion expression. *Psychological Science, 15*, 194-197.

Tracy, J. L., & Robins, R. W. (2007a). The nature of pride. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds), *The self-conscious emotions: Theory and research* (pp. 263-282). New York: Guilford.

Tracy, J. L., & Robins, R. W. (2007b). Self-conscious emotions: Where self and emotion meet. In C. Sedikides & S. Spence (Eds.), *The self in social psychology. Frontiers of social psychology series* (pp. 187-209). New York: Psychology Press.

Tracy, J. L., & Robins, R. W. (2007c). The psychological structure of pride: A tale of two facets. *Journal of Personality and Social Psychology, 92*, 506-525.

Tracy, J. L**.**, Robins, R. W., & Tangney, J. P. (2007). *The self-conscious emotions: Theory and research*. New York: Guilford.

Trafimow, D., Triandis, H. C., & Goto, S. G. (1991). Some tests of the distinction between the private self and collective self. *Journal of Personality and Social Psychology*, *60*, 649-655.

Triandis, H. C. (1997). Cross-cultural perspectives on personality. In R. Hogan, J. Johnson, and S. Briggs (Eds.), *Handbook of Personality Psychology*, (pp. 439-464). New York: Academic Press.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35-57.

Trope, Y. (1986). Self-enhancement and self-assessment in achievement behavior. In R. M. Sorrentino, E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 350-378). New York: Guilford Press.

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the lifespan. *Journal of Personality and Social Psychology, 84*, 205-220.

Tulving, E. (1993). Self-knowledge of an amnesic patient is represented abstractly. In T. K. Srull & R. S. Wyer (Eds.), *The mental representation of trait and autobiographical knowledge about the self: Advances in social cognition* (Vol. 5; pp. 147-156). Hillsdale, NJ: Erlbaum.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3-56). New York, NY: Oxford University Press.

Turner, J. C., & Onorato, R. S. (1999). Social identity, personality, and the self-concept: A self-categorization perspective. In T. R. Tyler, R. M. Kramer, & O. P. John (Eds.), *The psychology of the social self* (pp. 11-46). Mahwah, NJ: Erlbaum.

Tyler, T. R., Kramer, R. M., & John, O. P. (Eds.) (1999). *The psychology of the social self*. Mahwah, NJ: Erlbaum.

Verschueren, K., Buyck, P., & Marcoen, A. (2001). Self-representations and socioemotional competence in young children: A 3-year longitudinal study. *Developmental Psychology, 37*, 126-134.

Wang, Q. (2006). Culture and the development of self-knowledge. *Current Directions in Psychological Science, 15*, 182-187.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science,* 1301-1303.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton, NJ: Princeton University Press.

Wilson, D. S., Coleman, K., Clark, A. B., & Biederman, L. (1993). Shy-bold continuum in pumpkinseed sunfish (Lepomis gibbosus): An ecological study of a psychological trait. *Journal of Comparative Psychology*, *107*, 250-260.

Yamaguchi, S., Greenwald, A. G., Banaji, M. R., Murakami, F., Chen, D., Shiomura, K., Kobayashi, C., Cai, H., & Krendl, A. (2007). Apparent universality of positive implicit self-esteem. *Psychological Science*, *18*, 498-500.

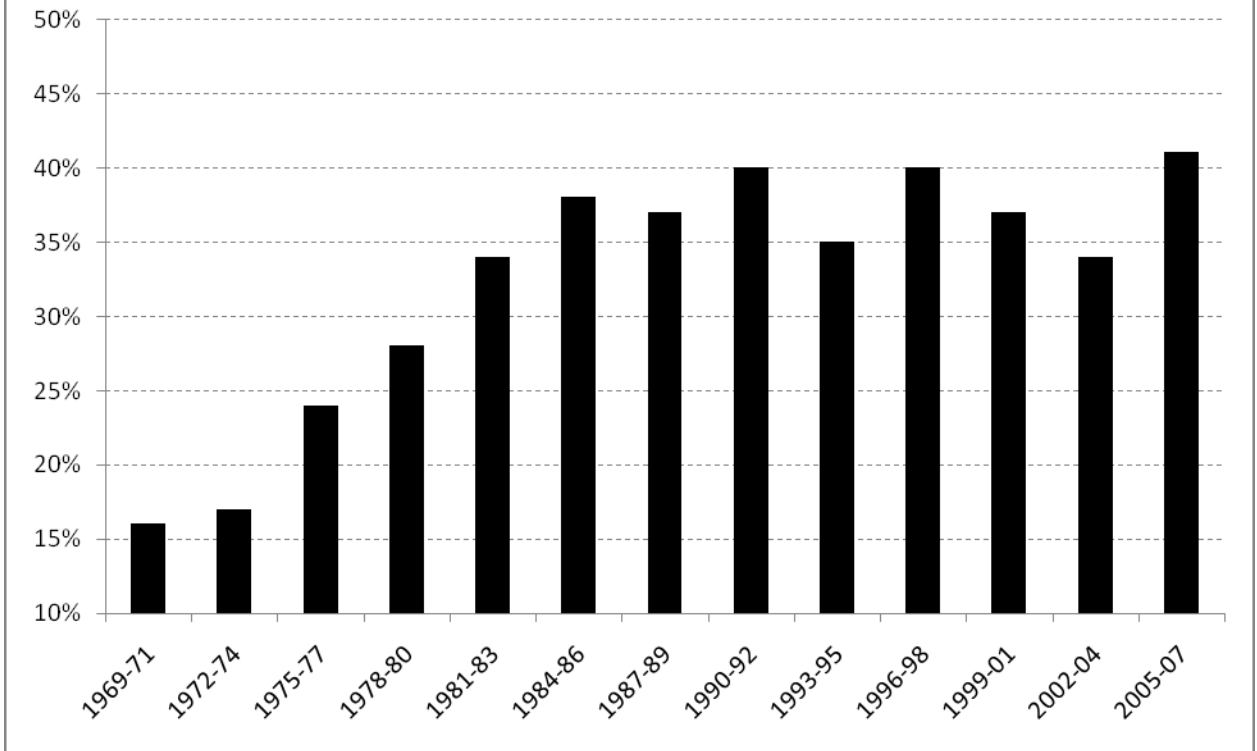**Figure 1**: % of Articles in JPSP with "Self" as a Keyword

Figure 2: Two Classes of Self Phenomena

Ongoing Sense of Self-Awareness	Stable Self-Representations
□ "I"	□ "Me"
□ Self as subject	□ Self as object
□ Self as perceiver	□ Self as perceived
□ Sentience	□ Personal (episodic) memories
□ Qualia	□ Self-knowledge (semantic memory)
□ Phenomenology	□ Linguistic/symbolic self
□ Subjective experience	□ Self-concept
□ Agent/homunculus	□ Self-representation
□ Self-awareness	□ Self-theories
□ Self-consciousness	□ Ideal/ought/possible selves

Figure 3: Layers and Structures of the Self

Levels/Locus of audience	Orientation	Description	Example	Basis of self-regard	Cultural differences	Individual differences
Personal	Private	Traits, values, and abilities	"I am a sensitive person"	Personal aspirations and standards;	Independent/ individualistic	Personal Identity Orientation Rosenberg Self-Esteem Scale Authentic and Hubristic Pride Scales Private Self-Consciousness Individualism Scale
Relational	Intimate	Other people with whom we have direct personal contact	"I am Amy's close friend"	Mutual regard; pride in one's self as a relationship partner, validation from intimate others		Inclusion of Other in Self Scale Mutuality Scale Interdependent Self-Construal Scale Internal Working Models
Social	Interpersonal	Social roles and reputation	"I am a popular professor"	Public recognition; praise from others pride in role		Social Identity Orientation Public Self-Consciousness Scale Social Self-Confidence (vs. Shyness)
Collective	Communal	Social categories to which we belong	"I am Irish"	Ethnic pride; pride in one's social groups	Interdependent/ collectivistic	Collective Identity Orientation Collective Self-Esteem Scale Inclusion of Other in Group Scale Collectivism Scale

Endnotes

[1] The keyword "self" is clearly over-inclusive and will detect articles addressing psychological phenomena beyond the scope of research on the self. However, any other keyword (e.g., self-esteem, self-concept, self-awareness) is necessarily under-inclusive and would fail to detect important aspects of the self literature.

[2] In contrast to the idea that self-understanding is linked to an understanding of others, Klein, Cosmides, Murray, and Tooby (2004) described the case of an individual with autism who has developed normal, consensually accurate knowledge of his own traits but is unable to differentiate accurately between the personalities of his various family members.

[3] The communication of self-conscious emotions to others may also serve an adaptive function. The nonverbal expression of embarrassment draws forgiveness and increases sympathy and liking from onlookers after a social transgression (Keltner & Buswell, 1997; Miller, 2007), and the pride expression may promote social status by increasing an individual's visibility to others following a socially valued achievement (Tracy & Robins, 2004b).

[4] Narcissistic behavior also seems to be present and in some cases adaptive in non-human animals. Sapolsky (1997) describes an orangutan named Hobbes—the "cocky son of a high-ranking female" (p. 83)—who immediately began acting like the alpha male after migrating to a new troop. To Sapolsky's surprise, Hobbes was quickly treated by others as a high status animal, despite his initial low status in the group. Although we clearly do not know whether Hobbes had an overly positive self-representation, his narcissistic behavior does seem to have served the adaptive function of helping him attain social status and consequently mates. Similar benefits may accrue to humans who believe they are more brilliant and powerful than they really are. An interesting point relevant to the positive illusions debate is that Hobbes had unusually high cortisol levels; as Sapolsky pointed out, "it doesn't come cheap to be a bastard twelve hours a day—a couple of months of this sort of thing is likely to exert a physiological toll" (p. 86).