77 The Cognitive Neuroscience of Knowing One's Self

STANLEY B. KLEIN

ABSTRACT The unified self of everyday experience may actually be composed of several functionally and neurally isolable components. These include episodic memories of one's own life, representations of one's own personality traits, facts about one's personal history (semantic personal knowledge), the experience of personal agency and continuity through time, and the ability to reflect on one's own thoughts and experiences (Klein, 2001). One component of the self—knowledge of one's own personality traits—is surprisingly resilient in the face of brain damage and developmental disorders. Personality knowledge can be preserved and even updated without any retrievable episodic memory. More strikingly, a pattern of category-specific dissociations within semantic memory suggests that the human cognitive architecture may include a subsystem that is functionally specialized for the acquisition, storage, and retrieval of trait self-knowledge. The ability to retrieve accurate information about one's own personality traits can be preserved despite damage to the systems that retrieve information from other content-based categories of semantic knowledge, including knowledge of other people's personality traits, knowledge of one's own personal history, knowledge of cultural history, and knowledge of facts about animals, foods, and objects. Neuropsychological case studies reveal dissociations not only of storage and retrieval but also of acquisition; personality knowledge may be acquired via learning mechanisms that are functionally distinct from those that cause the acquisition of knowledge about other domains. Taken together, the cognitive and neuropsychological evidence suggests that personality self-knowledge is acquired through domain-specific learning mechanisms, stored in proprietary databases, and retrieved via functionally specialized search engines.

Who is the I that knows the bodily me, who has an image of myself and a sense of identity over time, who knows that I have propriate strivings? I know all these things, and what is more, I know that I know them. But who is it who has this perspectival grasp? ... It is much easier to feel the self than to define the self.

—G. W. Allport, *Patterns and Growth in Personality*

What *is* the self? Philosophers and scientists pursuing an answer to this deep ontological question immediately find themselves immersed in a host of metaphysical questions about mind and body, subject and object, object and process, the homunculus, free will, self-awareness, and a variety of other puzzling matters (e.g., Williams, 1973; Cassam, 1994;

STANLEY B. KLEIN Department of Psychology, University of California, Santa Barbara, Santa Barbara, Calif.

Bermudez, 1998; Gallagher and Shear, 1999). The enduring nature of these problems has led some to question whether a conceptual understanding of the self is possible in practice (e.g., Olson, 1999; Uttal, 2000) or in principle (e.g., McGinn, 1991).

Although researchers are deeply interested in the complex questions and controversies raised by the problem of the ontology of self, that is not the focus of this chapter. Instead, the focus is on what can be thought of as *first-person epistemology*: how we come to know who and what we are (e.g., Crispin, Smith, and Macdonald, 1998). The cognitive architecture of an individual is able to learn about the individual it is situated in, and even experience itself as a knower. A cognitive account of the mechanisms, databases, and search engines that allow this information about the self to be acquired, stored, and retrieved should be possible, even if some of the more troubling ontological questions remain unanswered.

A brief history of social-cognitive explorations of self

In his book, *Consciousness Regained*, Nicholas Humphrey (1984) made a strong case for the proposition that our ability to reflect on the self—the capacity to experience ourselves as thinking, feeling, wanting, doing beings—is likely what gave rise to psychology in the first place. Indeed, no less an authority than William James (1890) proclaimed the self to be the fundamental unit of analysis for a science of mental life, the problem about which everything else revolves. Yet academic psychology, influenced by arguments from "black-box" behaviorism, largely ignored questions about the mental representation of self-knowledge until the late 1970s.

With the rise of the cognitive sciences, various components of the self began to be cashed out as computational systems and the databases they access. For example, research on theory of mind reframed Humphrey's "ability to reflect upon the self" as the ability to form metarepresentations: representations about other mental representations, whether one's own or others (e.g., Baron-Cohen, Leslie, and Frith, 1985; Leslie, 1987). In Leslie's (1987, 2000) account, these representations are data files with a particular format, including slots for an *agent* (e.g., "I," "you," "Lowell"),

that agent's *attitude* toward a proposition (e.g., "believes," "doubts," "hopes"), and an embedded proposition (e.g., "It is raining," "I will become anxious at the zoo"). Because the agent can be the self and the embedded proposition can itself be a metarepresentation, this data format allows the formation of self-reflective representations, such as "I believe that I will become anxious at the zoo."

The computational machinery that produces metarepresentations appears to come online at about 18 months (e.g., Leslie, 1987; Baron-Cohen, 1995), and it can be selectively impaired. For example, individuals with autism understand that photographs—physical representations of the world—can misrepresent the facts, but have difficulty understanding that beliefs—mental representations about the world—can do the same (Leslie and Thaiss, 1992). Autism, it has been argued, disrupts the development of metarepresentational machinery (Baron-Cohen, 2000; Leslie, 2000); schizophrenia may be a late-onset breakdown of the same system (Frith, 1992; Frith and Frith, 1992; Gallagher, 2000). If true, then the ability to reflect on one's own mental states should be impaired in both disorders, and this appears to be the case (e.g., Frith, 1992; Baron-Cohen, 1995).

But where does the information about the self that a metarepresentation represents come from? How is it derived, where is it stored, and how is it retrieved? Metarepresentational machinery may permit a cognitive architecture to reflect on the thoughts, desires, reactions, personality, and other properties of the individual in which it is situated. But it can do so only if that information is represented somewhere in the architecture. Recognizing this, social and personality psychologists began to examine how knowledge about the self is stored in memory. It was known that we store vast amounts of information about our own personality traits and those of others (e.g., Wiggins, 1973). The question is, in what data formats and storage systems is that knowledge represented, and what are the mechanisms whereby it is retrieved?

Investigations in social psychology centered on whether the representation of knowledge about the self differed from representations of knowledge about other social and nonsocial entities (for reviews, see Greenwald, 1981; Kihlstrom and Klein, 1994; Linville and Carlston, 1994). Speculation about the uniqueness of self-knowledge was fueled by theoretical and experimental work on the role of self in information processing. Particularly influential in this regard was the demonstration by Rogers, Kuiper, and Kirker (1977) that asking someone whether a trait adjective, such as *kind*, is self-descriptive leads to better recall of that adjective than asking the person to make other judgments about it (e.g., "Does the word *kind* describe you?" versus "What does *kind* mean?"; see also Bower and Gilligan, 1979; Klein and Kihlstrom, 1986; Klein and Loftus, 1988; for review, see Symons and Johnson, 1997). Given the recall superiority found for self-referential

judgments, it seemed to a number of investigators that self-knowledge might have properties that distinguish it from other structures in memory (e.g., Rogers, 1981; Greenwald and Pratkanis, 1984). Explaining these properties soon became the dominant focus of research exploring how self-knowledge is represented in and retrieved from memory (for review, see Higgins and Bargh, 1987; Kihlstrom and Klein, 1994; Linville and Carlston, 1994).

THE SELF AND MEMORY   How does a person know that he or she possesses some traits but not others? How is this knowledge represented in and retrieved from memory? These questions have been asked within the context of debates about multiple memory systems (for recent reviews of the memory systems debate, see Schacter and Tulving, 1994; Foster and Jelicic, 1999).

Psychologists generally agree that memory stores two basic types of information, procedural and declarative (e.g., Tulving, 1983, 1995; Cohen and Eichenbaum, 1993; Parkin, 1993; Schacter and Tulving, 1994). Procedural memory makes possible the acquisition and retention of motor, perceptual, and cognitive skills (e.g., knowing how to ride a bike); it consists in the nonconscious expression of previously acquired behavioral skills and cognitive procedures (e.g., Tulving, 1985; Tulving and Schacter, 1990; Parkin, 1993). Declarative memory consists in facts and beliefs about the world (e.g., knowing that canaries are yellow; knowing that Eli pitched a shutout yesterday). Conceptually, the difference between procedural and declarative memory coincides with Ryle's (1949) classic distinction between *knowing how* (operating on the environment in ways difficult to verbalize) and *knowing that* (stating knowledge in the form of propositions).

Tulving (1983, 1985, 1993a) distinguishes two types of declarative memory: episodic and semantic (see also Cermak, 1984; Wood, Brown, and Felton, 1989; Parkin, 1993; Moscovitch et al., 2000). Semantic memory is relatively generic, context-free knowledge about the world, such as *apples are edible, 2 + 2 = 4,* and *Sacramento is the capital of California.* Semantic memory usually lacks a source tag: it is experienced as knowledge without regard to where and when that knowledge was obtained (e.g., Tulving, 1983, 1993a, 1995; Perner and Ruffman, 1994; Gennaro, 1996; Wheeler, Stuss, and Tulving, 1997). Most semantic memory makes no reference to the self; it can, however, contain propositions expressing facts about the self (e.g., *Stan Klein was born in New York*), just as it can about other things in the world. But this information is known in the same way that one knows that apples are edible; it is not recalled or reexperienced.

In contrast to semantic memory, episodic memory records events as having been experienced by the self at a particular (and unique) point in space and time; when retrieved, these events are reexperienced in a quasi-

perceptual way, with conscious awareness that "this happened to *me*" (e.g., Tulving, 1983, 1993a; Suddendorf and Corballis, 1997; Wheeler, Stuss, and Tulving, 1997). Every episodic memory by definition entails a mental representation of the self as the agent or recipient of some action, or as the stimulus or experiencer of some state (Kihlstrom, 1997). Examples of episodic memory are *I remember attending a concert yesterday evening* and *I recall having met with my graduate student last Monday.*

Not surprisingly it is the episodic component of declarative memory that traditionally has been the focus of interest for psychologists studying the relation between self and memory. This is because retrieval from episodic memory is assumed to have a self-referential quality thought to be largely absent from other types of memorial experience (i.e., semantic and procedural; for discussion, see Klein, 2001; Kihlstrom, Beer, and Klein, 2002; Klein, Cosmides, Tooby, and Chance, 2002). By contrast, semantic memory is not accompanied by awareness of reexperiencing one's personal past (e.g., Tulving, 1993a, 1995; Perner and Ruffman, 1994; Wheeler, Stuss, and Tulving, 1997). I may *know* where I was born, but I do not know this by virtue of having recalled or reexperienced my birth. That is why this bit of personal history would be considered semantic knowledge, despite its being about oneself.

KNOWING ONESELF: SOURCES OF FIRST-PERSON DATA IN MEMORY Episodic memories of a personal past are clearly one source of information about the self (and others). It would be strange, however, if these were the only source of data about oneself. Many decisions require quick and accurate judgments of one's own personality traits and those of others (e.g., Will standing up to him make me anxious?). But making these judgments would be a slow process indeed if the only database with pertinent information was the episodic store: each time a judgment was needed, episodic memories would have to be retrieved, and the behavioral events they represent would have to be analyzed for evidence of the trait in question.

Better to have answers precomputed and available for whenever they are needed. Trait generalizations are precomputed summaries of the dispositions one manifested in various behavioral episodes. Research over the past 10 years has provided evidence that the semantic memory system contains a subsystem that stores information about one's own personality traits in the form of trait generalizations (e.g., *Self: Usually stubborn*). These trait summaries form a fast-access database that provides quick answers to decision processes that require trait judgments.

In the next section I review converging evidence that this trait summary database exists, based on studies of individuals with normal cognitive function and individuals with varying degrees of cognitive impairment (e.g., amnesia, autism, Alzheimer's dementia). From these studies, a tentative model is emerging of how representations from episodic and semantic memory interact to generate a conception of oneself (Klein, Cosmides, Tooby, and Chance, 2002). This model will be discussed after the evidence is reviewed.

Does the mind store trait summaries?

Two explanations have been offered for how personality trait judgments are made. The abstraction view proposes direct retrieval of precomputed trait summaries; the computational view eschews trait summaries and proposes instead that trait judgments are computed online, on the basis of retrieved episodes (for reviews, see Klein and Loftus, 1993a; Kihlstrom and Klein, 1994).

According to the computational view, there is a mechanism that makes trait judgments online by retrieving trait-relevant behaviors from episodic memory and computing their similarity to the trait being judged (e.g., Bower and Gilligan, 1979; Locksley and Lenauer, 1981; Smith and Zarate, 1992; Keenan, 1993). For example, if asked whether I am friendly, this mechanism would search the episodic memory store for trait-consistent episodes (in this case, records of events in which my behavior was friendly). The judgment would be computed from the episodes retrieved (based, e.g., on how diagnostic they were of friendliness or on how fast a given number could be retrieved).

According to the abstraction view, information about one's personality traits is abstracted from specific behaviors, either as they happen or on the basis of episodic memories of these behaviors. These abstractions are stored in the form of precomputed trait summaries (e.g., Buss and Craik, 1983; Klein, Loftus, Trafton, and Fuhrman, 1992; Klein and Loftus, 1993a; Lord, 1993). Trait judgments are made by direct retrieval from this store. When a trait summary is retrieved, trait-consistent episodes are not retrieved along with it (because the information they provide would be redundant). Trait-consistent episodes are consulted only when the search engine fails to retrieve a trait summary (e.g., when a summary does not exist yet for a particular trait).

Note that these two views carry very different predictions about the need to access episodic memories when making trait judgments. If the computational view is correct, then trait-consistent episodes must be retrieved to make a trait judgment. If the abstraction view is correct, then trait-consistent episodes will not be retrieved in making trait judgment, except under unusual circumstances (e.g., the absence of a summary). These predictions have been extensively tested through paradigms that take advantage of priming, encoding specificity, and encoding variability. Priming results are described in the next section (for converging results using the other methods, see Klein, Loftus, and Plog, 1992; Klein, Loftus, and Burton, 1989).

TESTING FOR TRAIT SUMMARIES IN COGNITIVELY NORMAL INDIVIDUALS The logic of tests using priming paradigms is straightforward. The computational view requires the retrieval of trait-consistent episodes; this means that being asked whether a trait describes you should activate trait-consistent behavioral episodes, allowing faster recall of them subsequently. No priming of trait-consistent episodes is predicted by the abstraction view (except in cases where summaries are absent).

Tests using priming paradigms support the abstraction view, not the computational view (Klein and Loftus, 1993a,b; Klein, Loftus, and Burton, 1989; Klein, Loftus, Trafton, and Fuhrman, 1992; Klein, Cosmides, Tooby, and Chance, 2002). Klein, Loftus, and colleagues presented each subject with many pairs of tasks; each pair involved a particular trait adjective (e.g., stubborn). Subjects were asked to retrieve a memory in which they displayed behavior relevant to the trait in question (e.g., "Recall a specific incident in which you behaved in a stubborn manner"). The dependent measure was the response latency for this recall task when it was the second task of the pair. The independent variable was the nature of the initial task, the prime.

The prime was either a describe task, a control task, or a filler task. The describe task asked subjects to judge whether the trait adjective was self-descriptive (e.g., "Does this describe you: Stubborn?"). The control task varied depending on the experiment; sometimes it was a define task (e.g., "Think of the definition of the word stubborn"), sometimes it was looking at a blank screen. Control tasks were ones that that do not elicit retrieval of trait-consistent behavioral episodes.

If the computational view is correct, then trait-consistent episodes will be activated and analyzed whenever one is asked to decide whether a trait describes oneself, that is, by performing the describe task. If trait-consistent episodic memories are activated by the describe task, then one should be able to retrieve those memories faster after performing a describe task than after performing a control task. This was not the case: when subjects were asked to recall a specific behavioral incident in which they manifested a particular trait (recall task), those who had first made a self-descriptiveness judgment were no faster than those who had not (e.g., Klein, Loftus, and Burton, 1989; Klein and Loftus, 1990, 1993a; Klein, Loftus, Trafton, and Fuhrman, 1992). Yet the procedure used is known to be sensitive enough to detect episodic priming when it occurs (e.g., Babey, Queller, and Klein, 1998; Klein, Loftus, Trafton, and Fuhrman, 1992; Sherman and Klein, 1994; Sherman et al., 1998). (For experiments showing that this result obtains regardless of how central a trait is to one's self-concept, see Klein, Cosmides, Tooby, and Chance, 2001; Klein, Loftus, Trafton, and Fuhrman, 1992.)

The fact that making a trait judgment did not prime episodic memories of trait-consistent behaviors is consistent with the abstraction view. That view holds that trait judgments can be made by directly retrieving trait summaries; no supporting evidence from episodic memory need be accessed. Klein and Loftus concluded from this series of results that people can answer questions about their own personality traits by accessing these trait summaries, without activating memories of episodes in which their behavior exemplified the trait.

Other research showed that the presence or absence of a trait summary is the variable that explains whether trait-consistent episodes are primed. When a trait summary is absent (as it is when trait-relevant behavioral experience is severely limited), trait-consistent episodes are indeed activated in the course of making trait judgments (e.g., Klein and Loftus, 1993a). The same holds when making judgments of others: trait-consistent episodes are not primed when summaries exist but are primed when they are absent (e.g., Babey, Queller, and Klein, 1998; Sherman and Klein, 1994).

. Additional support for the independence of episodic and semantic trait self-knowledge in brain-intact people recently was presented by Craik and colleagues (1999). Using positron emission tomography, these investigators discovered that requiring participants to judge trait adjectives for self-descriptiveness produced activation of cortical areas associated with semantic memory retrieval (left frontal regions) but not those associated with episodic memory retrieval (right frontal regions).[1] Similar findings have been reported by investigators employing functional magnetic resonance imaging (fMRI) technology (e.g., Kircher et al., 2000; Kelley et al., 2002; Zhang et al., 2002; Zysset et al., 2002; but see Keenan et al., 2001). Morin (2002), in a recent review of the literature on self-referential encoding and neuroimaging, concludes that the evidence points toward a left hemisphere involvement.[2]

These studies, all performed on individuals with normal cognitive function, converge on the following conclusion regarding "first-person epistemology": we can know what we are like by retrieving trait summaries from semantic memory. We do not have to compute the answer online based on information retrieved from episodic memory (e.g., Klein, Loftus, Trafton, and Fuhrman, 1992; Klein and Loftus, 1993a; Klein, Loftus, and Sherman, 1993, 1996; Schell, Klein, and Babey, 1996; Klein, Babey, and Sherman, 1997).

EVIDENCE FROM INDIVIDUALS WITH IMPAIRED COGNITIVE FUNCTION Given the automatic and flawless way in which different systems of memory normally interact, it is difficult to disentangle their respective contributions to knowledge about the self. However, because neuropsychological disorders of memory can be selective (i.e., patients may exhibit normal or near normal performance in some domains and profound impairments in others; e.g., Parkin and Leng, 1993; Parkin, 1996; Mayes, 2000), they can provide a

window into the operation of a component system in relative isolation, without the influence of other systems. By revealing differential patterns of impaired and preserved performance, the study of patients with neuropsychological impairments can illuminate aspects of a system's function and structure that are difficult to detect under normal operating conditions (e.g., Shallice, 1988; Tulving, 1983; Weiskrantz, 1997).

A SOCIAL NEUROPSYCHOLOGICAL APPROACH TO UNDERSTANDING THE SELF: FIVE CASE STUDIES   Klein and Loftus (1993a; see also Klein and Kihlstrom, 1998) proposed that the study of patients with amnesia—patients such as K.C., whose amnesia has been extensively studied by Tulving (1993b)—would prove to be a particularly effective method for examining the respective contributions of episodic and semantic memory to the creation of self-knowledge. This is because amnesic patients often experience highly selective memory loss, typically displaying intact semantic memory with impaired access to episodic memory (e.g., Tulving, 1983, 1995; Cermak, 1984; Parkin, 1987; Moscovitch et al., 2000). Amnesic patients therefore present a unique opportunity to test alternative models of self-knowledge: tests of trait knowledge can be conducted in amnesic patients with assurance that episodic memory for traits is not involved.

For example, if semantic memory contains a database of personality trait summaries, then an amnesic patient should be able to know what he or she is like despite being unable to recall the particular experiences from which that knowledge was derived. Neuropsychological data are now available from five patients. The dissociations found in these individuals speak strikingly to this and other issues involving how knowledge about oneself is acquired and represented in memory.

K.C.   Patient K.C. permanently lost his entire fund of episodic memory following a motorcycle accident (see Tulving, 1993b). He also underwent a marked personality change after the accident. Nevertheless, K.C. was able to describe his postmorbid personality with considerable accuracy (his mother's ratings served as the criterion; Tulving, 1993b). The fact that K.C. could accurately report his own personality traits supports the view that knowing oneself does not require retrieval of episodic memories. It is consistent with the hypothesis that personality information is stored independently from episodic memory, in the form of trait summaries.

It should be noted that K.C.'s self-knowledge reflected his postmorbid personality, not his premorbid personality. This means that K.C. not only had access to semantic knowledge of his own personality traits, he was also able to acquire new knowledge about his personality. Yet this updating occurred without his being able to recall any information about the behavioral episodes on which this updating was based. (It is unclear how K.C.'s updating was achieved; one possibility is that it occurred "online," as each new behavioral episode was unfolding.)

W.J.   W.J. suffered a concussive blow to the head shortly after completing her first quarter in college (Klein, Loftus, and Kihlstrom, 1996). Interviews conducted shortly after her accident revealed that W.J. had forgotten much of what had happened during the preceding 12 months, a period of time that included her first quarter at college. To document her deficit in episodic memory, Klein, Loftus, and Kihlstrom (1996) used the autobiographical memory cuing task originated by Galton (1879) and popularized by Crovitz (e.g., Crovitz and Schiffman, 1974) and Robinson (1976). W.J. was asked to try to recall specific personal events related to each of a list of cue words (e.g., car, sing, brave) and to provide for each recollection as precise a date as possible. On initial testing, she was unable to recollect personal events from recent years. Over the next month, however, her amnesia remitted completely, and when she was retested 4 weeks later, her performance had improved to the point that it was indistinguishable from that of neurologically healthy women who served as controls.

On two occasions, during her amnesia and after its resolution, W.J. was asked to provide personality ratings describing what she was like during her first quarter at college. While she was amnesic, W.J. was able to describe her personality; more important, the ratings she made during her amnesic period agreed with those she made afterward. Thus, while W.J. was amnesic she knew what she had been like in college, even though she could not episodically recollect any personal events or experiences from that time period.

Could W.J.'s judgments while amnesic be based on her continued access to episodic recollections of high school or earlier, periods not covered by her amnesia? Probably not. W.J., like many freshmen, manifested somewhat different personality traits in college than she did in high school. Yet her self-ratings during the amnesic period reflected her college personality (for data and analyses, see Klein, Loftus, and Kilstrom, 1996). This suggests that W.J.'s ratings were based on semantic knowledge of her personality during her time at college, not on recollections of episodes long past.

D.B.   The case of D.B. (like that of K.C.) shows that one can have accurate knowledge of one's own personality traits even with a total loss of episodic memory (Klein, Rozendal, and Cosmides, 2002). Patient D.B. was a 79-year-old man who became profoundly amnesic as a result of anoxia following cardiac arrest. On both informal questioning and psychological testing, D.B. was unable to consciously recollect a single thing he had ever done or experienced from any period of his life. In addition to his dense retrograde episodic amnesia, he also sustained severe anterograde episodic memory impairment, rendering him incapable of recollecting events that had transpired only minutes earlier.

To test D.B.'s semantic self-knowledge, we asked him on two separate occasions to judge a list of personality traits for self-descriptiveness. We also asked D.B.'s daughter (with whom he lived) to rate D.B. on the same traits. Our findings revealed that D.B.'s ratings were both reliable ($r = 0.69$ across sessions) and consistent with the way he is perceived by others ($r = 0.64$ between D.B. and his daughter). (Age-matched controls showed $rs = 0.74$ and 0.57 across sessions and raters, respectively.) D.B. thus appeared to have accurate and detailed knowledge about his personality even though he had no conscious access to any specific actions or experiences on which that knowledge was based. (For related findings, see Cermak and O'Connor, 1983; Evan et al., 1993; Tulving, 1993b; Starkstein, Sabe, and Dorrego, 1997; Kircher et al., 2000; for review, see Klein, Cosmides, Tooby, and Chance, 2002.)

D.B. manifested a clear dissociation between episodic and semantic self-knowledge. But can semantic knowledge of one's own personality traits dissociate from other types of semantic knowledge? Further testing of D.B. suggested that it can.

D.B.'s semantic memory was also affected by his illness, although this impairment was far less severe than that affecting his episodic memory (Klein, Rozendal, and Cosmides, 2002). For example, although he knew a variety of general facts about his life, he showed a number of striking gaps in his life story: he knew the name of the high school he attended and where he was born, but could not recall the names of any friends from his childhood or the year of his birth. He also showed spotty knowledge of facts in the public domain. For example, although he was able to accurately recount a number of details about certain historical events (e.g., the Civil War), his knowledge of other historical facts was seriously compromised (e.g., he claimed that America was discovered by the British in 1812). Despite these impairments in D.B.'s more general semantic knowledge, his knowledge of his own personality was intact. This result suggests a dissociation *within* semantic memory: between general semantic knowledge and semantic knowledge of one's own personality traits.

Additional testing revealed a dissociation between D.B.'s knowledge of his own personality traits and the traits of others. D.B. could not retrieve accurate knowledge of his daughter's personaity traits: the correlation between D.B.'s ratings of his daughter and her self-ratings was not reliable ($r = 0.23$), and was less than half that found between control parents' ratings of their child and the child's self-ratings ($r = 0.61$). Thus, although D.B.'s ability to retrieve accurate knowledge of his own personality was intact (no different from that of age-matched controls), he had lost the ability to retrieve accurate personality information about his adult daughter, with whom he lived.

In short, D.B.'s case goes well beyond the usual episodic/semantic distinction. It suggests category-specific dissociations within semantic memory. His ability to retrieve trait self-knowledge was intact; his ability to retrieve his daughter's traits was impaired; and his knowledge about the world at large (and specific facts about himself) was impaired. This pattern raises the possibility that the human cognitive architecture includes a subsystem of semantic memory that is functionally specialized for the storage and retrieval of trait self-knowledge. More data relevant to this claim come from the cases of R.J. and K.R.

R.J. Patients K.C., W.J., and D.B. lost access to episodic memory as a result of brain trauma. However, there also are cases of individuals in whom episodic memory failed to develop in the first place (e.g., Vargha-Khadem et al., 1997; Ahern, Wood, and McBrien, 1998). Such developmental dissociations are interesting because they permit inferences about the origins of self-knowledge that are not licensed by the discovery of dissociations caused by brain trauma in adults.

Consider, for example, the hypothesis that semantic self-knowledge, despite being functionally independent of episodic memory, initially is constructed from a database of episodic memories. This hypothesis cannot be ruled out by cases like D.B. and W.J.; their intact semantic self-knowledge could have been derived from episodic memories during the years prior to the brain trauma that caused their episodic loss as adults. But consider the implications of finding an individual who never developed the ability to access episodic memories, yet has intact semantic self-knowledge. This developmental dissociation would suggest that building a semantic database of trait self-knowledge does not require access to a database of episodic memories.

Autism is a developmental disorder that has been hypothesized to impair normal development of the cognitive machinery that supports metarepresentations (Baron-Cohen, Leslie, and Frith, 1985; Leslie, 1987; Baron-Cohen, 1995). It has been proposed that episodic memories are stored in and retrieved via metarepresentations (e.g., Perner, 1991; Cosmides and Tooby, 2000). If so, then autism should disrupt the normal development of episodic memory. To test this prediction, Klein, Chan, and Loftus (1999) assessed the episodic memory of R.J., a 21-year-old man with autism.

Compared with ability-matched controls, R.J. was found to be severely impaired on a variety of tests of recall, especially when memory for personally experienced events was tested (e.g., the Galton-Crovitz task). Although his impairment was developmental in origin, his episodic performance was similar to that found in classic amnesia caused by brain trauma (similar findings have been reported by Boucher and Warrington, 1976; Boucher, 1981; and Millward et al., 2000; but see Minshew and Goldstein, 1993).

Despite this deficit in episodic retrieval, R.J. demonstrated reliable and accurate knowledge of his personality traits. His test-retest correlations were high ($r = 0.86$; IQ-matched controls, $r = 0.78$). Moreover, the correlation between R.J.'s trait self-ratings and his mother's ratings of him was significant ($r = 0.56$) and did not differ reliably from that obtained from control mother-son pairs ($r = 0.50$). R.J.'s self-ratings also were compared with ratings of R.J. obtained from one of his teachers; the correlation again was reliable ($r = 0.49$) and comparable to those obtained between control mother-son pairs.

These findings suggest that R.J.'s knowledge of what he is like accurately reflects how he is perceived by people with whom he interacts. But how did he acquire this trait self-knowledge? His case suggests that conscious access to a database of episodic memories is unnecessary. R.J. cannot retrieve episodic memories now and, because his impairment is developmental in origin, he probably never developed this ability in the first place. All four cases—W.J., D.B., K.C., and R.J.—show that trait self-knowledge can exist independently of episodic access, but R.J.'s developmental dissociation suggests that the acquisition of trait self-knowledge does not require episodic access (as does K.C.'s ability to update).

As in the case of D.B., further tests of R.J. suggested content-specific dissociations within semantic memory. Klein, Cosmides, Costabile, and Mei (2002) asked R.J. to judge features of common objects (e.g., "Is a lemon sour?" "Is a balloon round?"). R.J.'s answers were reliable across sessions ($r = 0.77$). However, they did not correlate with those provided by others of the same mental age. There was high agreement among IQ-matched controls, with correlations among their answers ranging from 0.78 to 0.81. In contrast, correlations between R.J.'s answers and theirs ranged from 0.18 to 0.33.

R.J.'s atypical semantic knowledge is not due to a general inability to understand or answer questions; his ability to answer personality questions is fine. This pattern of consensually accurate personality knowledge coexisting with odd, nonconsensual knowledge of foods, animals, and objects is surprising. One would think the evidence of one's senses would allow the easy acquisition of knowledge about tastes, shapes, and colors. Indeed, words like *sweet, tall,* and *large* are more concrete and have more obvious referents than personality terms such as *kind, friendly,* and *ungrateful.* Nevertheless, an individual with autism was able to learn his own personality traits but was unable to acquire consensually held knowledge of foods, animals, and objects. Because R.J.'s condition is caused by a developmental disorder, this pattern raises the possibility that there may be mechanisms specialized for acquiring knowledge of one's personality that can remain intact even when the mechanisms for acquiring knowledge of other domains are quite impaired.

K.R. K.R., a patient diagnosed with late-stage Alzheimer's dementia, shows that reliable, accurate knowledge of one's own personality can exist without the ability to update that knowledge (Klein, Cosmides, and Costabile, 2003).

K.R.'s performance on standard tests of cognitive functioning (e.g., the Mini-Mental State Examination) indicated severe dementia. She was disoriented for time and place and experienced difficulties with word finding and object naming. K.R. could not, for example, name simple objects such as batteries and pencils or draw the face of a clock from memory. Her anterograde memory function was severely impaired, leaving her unable to recall events she had in mind moments before. Knowledge of her personal past was sketchy: for example, she sometimes believed her late husband was alive, and her estimates of how long she had lived in her current facility ranged from 2 months to 14 years.

Despite these profound deficits, K.R. had reliable knowledge of her own personality traits. We asked her on two occasions (separated by 2 weeks) to judge a list of personality traits for self-descriptiveness. We also asked K.R.'s daughter and her caregiver at the assisted living facility to rate K.R. on the same traits. The results showed that K.R.'s test-retest ratings were reliable ($r = 0.86$). However, her ratings did not agree with the ratings provided by either her daughter or her caregiver ($r = 0.31, -0.11$ for daughter and caregiver, respectively). This lack of consistency was not because the daughter and caregiver were poor judges of character; when asked to rate other individuals, their judgments correlated strongly with those of others.

How could K.R.'s ratings be so reliable, yet agree so little with those who know her best? According to her family, K.R.'s personality and behavior changed dramatically as the disease progressed, but she seemed unaware of her transformation (a situation fairly common among patients with Alzheimer's dementia; e.g., Siegler, Dawson, and Welsh, 1994; Mills, 1998; Clare, in press). This suggests that the disease may have impaired K.R.'s ability to update the mental records that stored information about her personality. If her self-knowledge was intact but not being updated, then K.R.'s ratings may have reflected her premorbid personality rather than her current one.

To test this hypothesis, we asked K.R.'s daughter to rate her mother on the same list of traits, only this time she was asked to base her ratings on her mother's personality prior to the onset of the disease. These ratings were strongly correlated with those provided by K.R. herself ($r = 0.59$), as were preonset trait ratings of K.R. provided by her son-in-law ($r = 0.79$). Taken together, these findings indicate that K.R.'s ratings were accurate, but reflected her pre-Alzheimer's personality.

K.R. also knew her daughter's personality traits: when asked to rate her daughter on the same list of traits, her ratings correlated strongly with her daughter's self-ratings (r = 0.65). This is expected if K.R.'s fund of personality knowledge was created premorbidly and remained intact. But if, as hypothesized, K.R. lost the ability to update her personality files, then her ratings should have been inaccurate for people whom she first met after the onset of her dementia. This was the case. On two occasions (again, 2 weeks apart), K.R. was asked to rate her caregiver, whom she had known for 2.5 years. When the subject was her caregiver, K.R.'s test-retest reliability was low (r = 0.34), in striking contrast to the reliablity of her self-ratings (r = 0.86). Moreover, K.R.'s ratings of the caregiver did not overlap reliably with the caregiver's ratings of his own personality (r = 0.18). This difference was not due to the caregiver having a skewed view of himself: His self-ratings were strongly correlated with those provided by two neurologically healthy women living in the same facility who were similar in age to K.R. and who had known him for about the same length of time (r = 0.73, 0.68). This also shows that K.R.'s inability to acquire new personality information was not a simple manifestation of the normal aging process, because the neurologically healthy age-matched controls were quite capable of acquiring accurate knowledge of the personality of someone they had recently met.

Thus, despite profound cognitive deficits, K.R. had intact knowledge of her own premorbid personality and that of her daughter. That her trait knowledge had been preserved and remained retrievable is remarkable, given the difficulties she had retrieving ordinary facts from semantic memory: the names of everyday objects, what a clock looks like, where she was. As in the cases of R.J. and D.B., K.R.'s preserved self-knowledge suggests a dissociation *within* semantic memory, suggesting the presence of a functionally specialized database for the storage and retrieval of information about her personality.

It would appear, however, that the computational machinery responsible for updating personality knowledge was impaired in K.R. by Alzheimer's disease. K.R. did not know her own current, postmorbid personality, nor was she able to learn the personality traits of her primary caregiver. In K.R., trait knowledge of self and other remained intact, but the ability to update that knowledge based on new experiences was no longer functional.

Conclusions

A SEMANTIC SUBSYSTEM SPECIALIZED FOR THE STORAGE AND RETRIEVAL OF PERSONALITY TRAIT KNOWLEDGE   The results of the neuropsychological case studies support the following inferences:

1. The human mind stores knowledge of its own personality in the form of trait summaries. Retrieving trait summaries from this database does not depend on accessing episodic memories. Accurate trait judgments can be made by amnesic individuals—people who cannot retrieve any episodic memories (K.C., W.J., D.B., R.J.).

2. That K.C. knew his postmorbid personality suggests that trait summaries can be updated without accessing episodic memories. (Perhaps they are updated online, as events unfold.)

3. Intact retrieval of personality trait summaries can occur despite Alzheimer's dementia so severe that it impairs access to knowledge about many semantic domains (K.R.).

4. D.B. and R.J. had intact knowledge of their own personality traits, yet they showed impairments in other domains of semantic knowledge (personal history; general history; facts about animals, foods, objects; etc.). This is a dissociation *between* domains of semantic memory. These dissociations suggest that trait self-knowledge is a functionally isolable subsystem of semantic memory.

5. That D.B. knew his own personality traits but not his daughter's suggests that knowledge about one's own traits is stored separately from knowledge of other people's traits.

6. Taken together, nos. 3–5 suggest that there is a subsystem of semantic memory that is functionally specialized for the storage and retrieval of trait self-knowledge.

The idea of a subsystem within semantic memory specialized for storage and retrieval of personality trait knowledge is consistent with recent findings suggesting that semantic memory can be fractionated into different components, each of which can be damaged independently (e.g., Hodges and Patterson, 1997; Mackenzie Ross, and Hodges, 1997; Cappa et al., 1998; Caramazza and Shelton, 1998). For example, there are cases in which brain damage creates very content-specific patterns of nonretrieval from semantic memory, as seen in patients who (for example) cannot retrieve information about animals but can retrieve information about inanimate objects, whereas others have the opposite pattern of impairment (e.g., Caramazza and Shelton, 1998; Caramazza, 2000), still others have a selective deficit in their ability to retrieve knowledge of types of food (e.g., Hillis and Caramazza, 1991; Hart and Gordon, 1992; Laiacona et al., 1993), and so on. In all cases, the information that is selectively spared or impaired is a type of general world knowledge. It is therefore argued that the inaccessible or missing information is drawn from a semantic memory system, and that category-specific impairments reflect subsystems within a more encompassing semantic system (e.g., Hodges and Patterson, 1997; Caramazza, 2000; but see Martin, Ungerleider, and Haxby, 2000). From this perspective, D.B.'s and R.J.'s normal performance on the trait self-knowledge questionnaire can be seen as reflecting the operation of a specialized subsystem within semantic memory that represents trait knowledge about the self and was not compromised by cortical damage.

A SPECIALIZED ACQUISITION SYSTEM? The neuropsychological case histories also permit drawing some tentative conclusions about how the cognitive architecture learns the personality traits of the individual in which it is situated.

1. Learning personality traits does not require conscious access to episodic memories. K.C. learned about his postmorbid personality despite having no ability to retrieve episodic memories. R.J. also knew his personality traits, yet he could not retrieve behavioral episodes from memory. Indeed, R.J.'s disorder is developmental in origin, suggesting that he had never been able to retrieve episodic memories.

2. Alzheimer's dementia can damage the mechanisms that allow one to learn about personality traits, whether one's own or others. Yet the inability to update personality knowledge need not interfere with the ability to retrieve information from a preexisting store of trait summaries (K.R.).

3. Any dissociation between semantic domains, whether due to brain trauma or to autism, suggests functionally isolable storage and retrieval systems (D.B., R.J., K.R.). But finding a developmental dissociation in R.J. suggests a functionally isolable acquisition system. His semantic dissociation suggests that trait self-knowledge is acquired via learning mechanisms that are functionally distinct from those that cause the acquisition of knowledge about animals, objects, and foods.

Domain-general learning theories, connectionist or otherwise, presume that the same learning mechanisms account for knowledge acquisition across content domains. But a developmental dissociation that impairs the acquisition of knowledge about animals, objects, and foods while having no effect on the acquisition of trait knowledge is difficult to reconcile with such theories. Such results are especially difficult for theories positing equipotential mechanisms that compute correlations between elementary perceptual or conceptual dimensions. Surely the evidence of one's senses is sufficient for R.J. and others to end up concurring that apples are sweet, lemons are not, rocks are hard, and giraffes are tall. Yet R.J. and others did not concur in their judgments of easily observable properties of food, animals, and objects. In stark contrast, R.J.'s judgments about his own personality were consistent with those of others who know him, even though R.J.'s judgments were those of an autistic individual with social deficits.

MIND DESIGN: WHY HAVE A PERSONALITY TRAIT DATABASE ALONGSIDE AN EPISODIC STORE? Trait summaries have a signal virtue: they provide fast answers to trait judgment questions. This is important, because social interaction often requires split-second decisions, and the best course of action may depend on assumptions about how you and others are likely to behave in various situations. Time can be saved if trait summaries—generalizations about how people are

likely to behave—are computed in advance and stored for later use (for discussion, see Klein, Cosmides, Tooby, and Chance, 2002). The alternative—retrieving and then evaluating a series of episodes online each and every time a trait judgment is needed—is more costly in both time and computation (e.g., Klein, Loftus, Trafton, and Fuhrman, 1992; Klein, Babey, and Sherman, 1997).

This view explains three sets of interlocking facts. First, it explains why trait summaries exist. Second, it explains why retrieving a trait summary fails to prime recall of trait-consistent episodes. Because summaries are precomputed answers to trait judgment questions, there is no additional advantage to retrieving trait-consistent episodes in tandem with them; the information that trait-consistent episodes provides is redundant with the summary (Babey, Queller, and Klein, 1998; Klein, Cosmides, Tooby, and Chance, 2001, 2002). Third, and less obviously, it explains why an episodic store is maintained despite trait summaries, and when trait judgments will access the episodic store.

Klein, Cosmides, Tooby, and Chance (2002) argue that an excellent package of speed plus accuracy can be engineered into a decision system by jointly activating a trait summary and episodic memories that are inconsistent with it. Trait summaries allow fast access to relevant information. But a trait summary (e.g., "I am usually friendly") gives information about behavior under average circumstances. It does not tell you under what circumstances a behavior deviates from average. In deciding how to behave, one is always facing a particular situation. Accordingly, a generalization is most useful when its scope is delimited, that is, when it is accompanied by information specifying those situations in which it does not apply. Episodic memories that are inconsistent with the generalization can serve this function, because they encode specific situations in which the generalization fails to predict the outcome. Thus, to render judgments that are both fast and accurate, judgment and decision procedures should be designed to search for summary information in semantic memory and, on retrieving it, also search for episodic memories that are inconsistent with that summary, ones that place boundary conditions on the summary's scope. Thus, there is a function to maintaining a store of episodic memories even after a trait summary has been formed: memories of behavioral episodes can provide boundary conditions on the scope of generalizations (Babey, Queller, and Klein, 1998; Cosmides and Tooby, 2000; Klein, Cosmides, Tooby, and Chance, 2001, 2002).

This *scope hypothesis* was tested in a recent series of experiments on trait self-judgments using the priming paradigm described above. Klein, Cosmides, Tooby, and Chance (2001, 2002) showed that when a trait summary is retrieved, trait-inconsistent behavioral episodes are retrieved along with it. More specifically, the time it took subjects to recall a trait-inconsistent episode was faster following a describe task than

following a control task. In other words, asking a subject whether he or she is *kind* coactivates memories of episodes in which that person did something *unkind*. Consistent with the scope hypothesis, inconsistent episodes are primed only when a trait summary has been retrieved. When a trait summary is absent, trait-consistent episodes are primed. This makes sense: in the absence of a trait summary, episodes are the only information one has on which to base a judgment.

Many view priming as a functionless by-product of neural activation. The results of these experiments support a quite different view: computational systems will be designed to prime representations when this solves an adaptive problem for the organism. The fact that retrieving trait summaries primes episodic memories that are inconsistent with the summary but not ones that are consistent with it cannot be explained as a by-product of neural activation (for discussion, see Klein, Cosmides, Tooby, and Chance, 2002). It is instead the signature of a functional and adaptive system: one that is designed to deliver just the right mix of information from memory to the right decision rules at the right time.

The human cognitive architecture does seem to contain a database that is functionally specialized for the storage and retrieval of personality trait information. Because social life is so adaptively important for our species, it should not be surprising to find that this database is resilient in the face of trauma and developmental disruptions. Nor should it be surprising to find mechanisms specialized for creating and refreshing this database; people change with time and experience, and the database needs to be constantly updated so that it accurately captures what a person is like at the moment that relevant decisions are made.

By isolating and elucidating the systems and databases that allow us to know ourselves, cognitive neuroscience is making "first-person epistemology" a topic of scientific inquiry rather than just philosophical speculation. As this process continues, the self may gradually reclaim the place William James originally carved out for it: as a central construct in psychology.

NOTES

1. These conclusions were based on results obtained using the analytic technique known as statistical parametric mapping. However, another technique, partial least-squares analysis, revealed activations of the right and left medial frontal lobes.

2. Although these pioneering studies are provocative and interesting, it is probably too early to conclude that the self is located in the left cerebral hemisphere. Although cognitive science generally has embraced a doctrine of modularity, the neural representation of individual items of declarative knowledge is distributed widely across the cortex. Accordingly, while self-referential processing may be performed by a specialized brain module or system, declarative knowledge of the self, whether episodic or semantic, is likely to be widely distributed over the same neural structures that represent knowledge of other people, as well as objects in the nonsocial domain.

REFERENCES

AHERN, C. A., F. B. WOOD, and C. M. McBRIEN, 1998. Preserved vocabulary and reading acquisition in an amnesic child. In *Brain and Values*, K. Pribram, ed. Mahwah, N.J.: Erlbaum, pp. 277–298.

ALLPORT, G. W., 1961. *Patterns and Growth in Personality*. New York: Holt, Rinehart & Winston.

BABEY, S. H., S. QUELLER, and S. B. KLEIN, 1998. The role of expectancy violating behaviors in the representation of trait-knowledge: A summary-plus-exception model of social memory. *Soc. Cogn.* 16:287–339.

BARON-COHEN, S., 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, Mass.: MIT Press.

BARON-COHEN, S., 2000. The cognitive neuroscience of autism: Evolutionary approaches. In *The New Cognitive Neurosciences*, M. Gazzaniga, ed. Cambridge, Mass.: MIT Press, pp. 1249–1257.

BARON-COHEN, S., A. M. LESLIE, and U. FRITH, 1985. Does the autistic child have a "theory of mind"? *Cognition* 21:37–46.

BERMUDEZ, J. L., 1998. *The Paradox of Self-Consciousness*. Cambridge, Mass.: MIT Press.

BOUCHER, J., 1981. Memory for recent events in autistic children. *J. Autism Dev. Disord.* 11:293–301.

BOUCHER, J., and E. K. WARRINGTON, 1976. Memory deficits in early infantile autism: Some similarities to the amnesic syndrome. *Br. J. Psychol.* 67:73–87.

BOWER, G. H., and S. G. GILLIGAN, 1979. Remembering information related to one's self. *J. Res. Pers.* 13:420–432.

BUSS, D. M., and K. H. CRAIK, 1983. The act frequency approach to personality. *Psychol. Rev.* 90:105–126.

CAPPA, S. F., M. FRUGONI, P. PASQUALI, D. PERANI, and F. ZORAT, 1998. Category-specific naming impairment for artifacts. *Neurocase* 4:391–397.

CARAMAZZA, A., 2000. The organization of conceptual knowledge in the brain. In *The New Cognitive Neurosciences*, M. S. Gazzaniga, ed. Cambridge, Mass.: MIT Press.

CARAMAZZA, A., and J. SHELTON, 1998. Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *J. Cogn. Neurosci.* 10:1–34.

CASSAM, Q., 1994. *Self-Knowledge*. Oxford, England: Oxford University Press.

CERMAK, L. S., 1984. The episodic-semantic memory distinction in amnesia. In *Neuropsychology of Memory*, L. R. Squire and N. Butters, eds. New York: Guilford Press, pp. 45–54.

CERMAK, L. S., and M. O'CONNOR, 1983. The anterograde and retrograde retrieval ability of a patient with amnesia due to encephalitis. *Neuropsychologia* 21:213–234.

CLARE, L., in press. Managing threats to self: Awareness in early stage Alzheimer's disease. *Soc. Sci. Med.*

COHEN, N. J., and H. B. EICHENBAUM, 1993. *Memory, Amnesia, and Hippocampal Function*, Cambridge, Mass.: MIT Press.

COSMIDES, L., and J. TOOBY, 2000. Consider the source: The evolution of adaptations for decoupling and metarepresentation. In *Metarepresentations: A Multidisciplinary Perspective. Vancouver Studies in Cognitive Science*, D. Sperber, ed. New York: Oxford University Press, pp. 53–115.

CRAIK, F. I. M., T. M. MOROZ, M. MOSCOVITCH, D. T. STUSS, G. WINOCUR, E. TULVING, and S. KAPUR, 1999. In search of the self: A PET investigation of self-referential information. *Psychol. Sci.* 10:26–34.

CRISPIN, W., B. C. SMITH, and C. MACDONALD, eds., 1998. *Knowing Our Own Minds*. Oxford, England: Oxford University Press.

CROVITZ, H. F., and H. SCHIFFMAN, 1974. Frequency of episodic memories as a function of their age. *Bull. Psychon. Soc.* 4(5B): 517–518.

EVAN, J., B. WILSON, E. P. WRAIGHT, and J. R. HODGES, 1993. Neuropsychological and SPECT scan findings during and after transient global amnesia: Evidence for the differential impairment of remote episodic memory. *J. Neurol. Neurosurg. Psychiatry* 56: 1227–1230.

FOSTER, J. K., and M. JELICIC, 1999. *Memory: Systems, Process, or Function?* New York: Oxford University Press.

FRITH, C. D., 1992. *The Cognitive Neuropsychology of Schizophrenia*. East Sussex, England: Erlbaum/Taylor and Francis.

FRITH, C. D., and U. FRITH, 1992. Elective affinities in schizophrenia and childhood autism. In *Social Psychiatry: Theory, Methodology, and Practice*, P. E. Bebbington, ed. New Brunswick, N.J.: Transactions Press.

GALLAGHER, S., 2000. Philosophical conceptions of the self: Implications for cognitive science. *Trends Cogn. Sci.* 4:14–21.

GALLAGHER, S., and J. SHEAR, 1999. *Models of the Self.* Thorverton, England: Academic.

GALTON, F., 1879. Psychometric experiments. *Brain* 2:149–162.

GENNARO, R. J., 1996. *Consciousness and Self-Consciousness*. Philadelphia: John Benjamins.

GREENWALD, A. G., 1981. Self and memory. In *The Psychology of Learning and Motivation*, vol. 15, G. H. Bower, ed. New York: Academic Press, pp. 201–236.

GREENWALD, A. G., and A. R. PRATKANIS, 1984. The self. In *Handbook of Social Cognition*, vol. 3, R. S. Wyer and T. K. Srull, eds. Hillsdale, N.J.: Erlbaum, pp. 129–178.

HART, J., and B. GORDON, 1992. Neural subsystems for object knowledge. *Nature* 6390:60–64.

Higgins, E. T., and J. A. Bargh, 1987. Social cognition and social perception. *Ann. Rev. Psychol.* 38:369–425.

HILLIS, A. E., and A. CARAMAZZA, 1991. Category-specific naming and comprehension impairment: A double dissociation. *Brain* 114:2081–2094.

HODGES, J. R., and K. PATTERSON, 1997. Semantic memory disorders. *Trends Cogn. Sci.* 1:68–72.

HUMPHREY, N., 1984. *Consciousness Regained: Chapters in the Development of Mind*. New York: Oxford University Press.

JAMES, W., 1890. *The Principles of Psychology*, vol. 1. New York: Holt.

Keenan, J. M., 1993. An exemplar model can explain Klein and Loftus' results. In *Advances in Social Cognition*, vol. 5, T. K. Srull and R. S. Wyer, eds. Hillsdale, N.J.: Erlbaum, pp. 69–77.

KEENAN, J. P., A. NELSON, M. O'CONNOR, and A. PASCUAL-LEONE, 2001. Self-recognition and the right hemisphere. *Nature* 409:305.

KELLEY, W. M., C. N. MACRAE, C. L. WYLAND, S. CAGLAR, S. INATI, and T. F. HEATHERTON, 2002. Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 15:785–794.

KIHLSTROM, J. F., 1997. Consciousness and me-ness. In *Scientific Approaches to Consciousness*, J. D. Cohen and J. W. Schooler, eds. Mahwah, N.J.: Erlbaum, pp. 451–468.

KIHLSTROM, J. F., J. S. BEER, and S. B. KLEIN, 2002. Self and identity as memory. In *Handbook of Self and Identity*, M. R. Leary and J. Tangney, eds. New York: Guilford Press, pp. 68–90.

KIHLSTROM, J. F., and S. B. KLEIN, 1994. The self as a knowledge system. In *Handbook of Social Cognition*, vol. 1, *Basic Processes*, R. S. Wyer and T. K. Srull, eds. Hillsdale, N.J.: Erlbaum, pp. 153–208.

KIRCHER, T. T. J., C. SENIOR, M. L. PHILLIPS, P. J. BENSON, E. T. BULLMORE, M. BRAMMER, A. SIMMONS, S. C. R. WILLIAMS, M. BARTELS, and A. S. DAVID, 2000. Towards a functional neuroanatomy of self processing: Effects of faces and words. *Cogn. Brain Res.* 10:133–144.

KLEIN, S. B., 2001. A self to remember: A cognitive neuropsychological perspective on how self creates memory and memory creates self. In *Individual Self, Relational Self, and Collective Self*, C. Sedikides and M. B. Brewer, eds. Philadelphia: Psychology Press, pp. 25–46.

KLEIN, S. B., S. H. BABEY, and J. W. SHERMAN, 1997. The functional independence of trait and behavioral self-knowledge: Methodological considerations and new empirical findings. *Soc. Cognit.* 15:183–203.

KLEIN, S. B., R. L. CHAN, and J. LOFTUS, 1999. Independence of episodic and semantic self-knowledge: The case from autism. *Soc. Cognit.* 17:413–436.

KLEIN, S. B., L. COSMIDES, and K. A. COSTABILE, 2003. Preserved knowledge of self in a case of Alzheimer's dementia. *Soc. Cognit.* 21:157–165.

KLEIN, S. B., L. COSMIDES, K. A. COSTABILE, and L. MEI, 2002. Is there something special about the self? A neuropsychological case study. *J. Res. Pers.* 36:490–506.

KLEIN, S. B., L. COSMIDES, J. TOOBY, and S. CHANCE, 2001. Priming exceptions: A test of the scope hypothesis in naturalistic trait judgments. *Soc. Cognit.* 19:443–468.

KLEIN, S. B., L. COSMIDES, J. TOOBY, and S. CHANCE, 2002. Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychol. Rev.* 109:306–329.

KLEIN, S. B., and J. F. KIHLSTROM, 1986. Elaboration, organization, and the self-reference effect in memory. *J. Exp. Psychol. Gen.* 115: 26–38.

KLEIN, S. B., and J. F. KIHLSTROM, 1998. On bridging the gap between social-personality psychology and neuropsychology. *Pers. Soc. Psychol. Rev.* 2:228–242.

KLEIN, S. B., and J. LOFTUS, 1988. The nature of self-referent encoding: The contributions of elaborative and organizational processes. *J. Pers. Soc. Psychol.* 55:5–11.

KLEIN, S. B., and J. LOFTUS, 1990. The role of abstract and exemplar-based knowledge in self-judgments: Implications for a cognitive model of the self. In *Advances in Social Cognition*, vol. 3, T. K. Srull and R. S. Wyer, eds. Hillsdale, N.J.: Erlbaum, pp. 131–139.

KLEIN, S. B., and J. LOFTUS, 1993a. The mental representation of trait and autobiographical knowledge about the self. In *Advances in Social Cognition*, vol. 5, T. K. Srull and R. S. Wyer, eds. Hillsdale, N.J.: Erlbaum.

KLEIN, S. B., and J. LOFTUS, 1993b. Behavioral experience and trait judgments about the self. *Pers. Soc. Psychol. Bull.* 19:740–745.

KLEIN, S. B., J. LOFTUS, and H. A. BURTON, 1989. Two self-reference effects: The importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-referent encoding. *J. Pers. Soc. Psychol.* 56:853–865.

KLEIN, S. B., J. LOFTUS, and J. F. KIHLSTROM, 1996. Self-knowledge of an amnesic patient: Toward a neuropsychology of personality and social psychology. *J. Exp. Psychol. Gen.* 125:250–260.

KLEIN, S. B., J. LOFTUS, and A. E. PLOG, 1992. Trait judgments about the self: Evidence from the encoding specificity paradigm. *Pers. Soc. Psychol. Bull.* 18:730–735.

KLEIN, S. B., J. LOFTUS, and J. W. SHERMAN, 1993. The role of summary and specific behavioral memories in trait judgments about the self. *Pers. Soc. Psychol. Bull.* 19:305–311.

KLEIN, S. B., J. LOFTUS, R. G. TRAFTON, and R. W. FUHRMAN, 1992. The use of exemplars and abstractions in trait judgments: A model of trait knowledge about the self and others. *J. Pers. Soc. Psychol.* 63:739–753.

KLEIN, S. B., K. ROZENDAL, and L. COSMIDES, 2002. A social-cognitive neuroscience analysis of the self. *Soc. Cognit.* 20:105–135.

KLEIN, S. B., R. W. SHERMAN, and J. LOFTUS, 1996. The role of episodic and semantic memory in the development of trait self-knowledge. *Soc. Cognit.* 14:277–291.

LAIACONA, M., R. BARBAROTTO, K. TRIVELLI, and E. CAPITANI, 1993. Perceptual and associative knowledge in category specific impairment of semantic memory. *Cortex* 4:727–740.

LESLIE, A. M., 1987. Pretense and representation: The origins of "theory of mind." *Psychol. Rev.* 94:412–426.

LESLIE, A. M., 2000. "Theory of mind" as a mechanism of selective attention. In *The New Cognitive Neurosciences*, M. S. Gazzaniga, ed. Cambridge, Mass.: MIT Press, pp. 1235–1247.

LESLIE, A. M., and L. THAISS, 1992. Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition* 43:225–251.

LINVILLE, P., and D. E. CARLSTON, 1994. Social cognition of the self. In *Social Cognition: Impact on Social Psychology*, P. G. Devine, D. L. Hamilton, and T. M. Ostrom, eds. San Diego, Calif.: Academic Press, pp. 143–193.

LOCKSLEY, A., and M. LENAUER, 1981. Considerations for a theory of self-inference processes. In *Personality, Cognition, and Social Interaction*, N. Cantor and J. F. Kihlstrom, eds. Hillsdale, N.J.: Erlbaum, pp. 263–277.

LORD, C. G., 1993. The "social self" component of trait knowledge about the self. In *Advances in Social Cognition*, vol. 5, T. K. Srull and R. S Wyer, eds. Hillsdale, N.J.: Erlbaum, pp. 91–100.

MACKENZIE ROSS, S. J., and J. R. HODGES, 1997. Preservation of famous person knowledge in a patient with severe postanoxic amnesia. *Cortex* 33:733–742.

MARTIN, A., L. UNGERLEIDER, and J. HAXBY, 2000. Category specificity and the brain: The sensory/motor model of semantic representations of objects. In *The New Cognitive Neurosciences*, M. S. Gazzaniga, ed. Cambridge, Mass.: MIT Press, pp. 1023–1036.

MAYES, A. R., 2000. Selective memory disorders. In *The Oxford Handbook of Memory*, E. Tulving and F. I. M. Craik, eds. New York: Oxford University Press, pp. 427–440.

MCGINN, C., 1991. *The Problem of Consciousness: Essays Toward a Resolution.* Oxford, England: Blackwell.

MILLS, M. A., 1998. *Narrative Identity and Dementia.* Aldershot, England: Ashgate Publishing.

MILLWARD, C., S. POWELL, D. MESSER, and R. JORDAN, 2000. Recall for self and other in autism: Children's memory for events experienced by themselves and their peers. *J. Autism Dev. Disord.* 30: 15–28.

MINSHEW, N. J., and G. GOLDSTEIN, 1993. Is autism an amnesic disorder? Evidence from the California Verbal Learning Test. *Neuropsychology* 7:209–216.

MORIN, A., 2002. Right hemisphere self-awareness: A critical assessment. *Consciousness Cognit.* 11:396–401.

MOSCOVITCH, M., T. YASCHYSHYN, M. ZIEGLER, and L. NADEL, 2000. Remote episodic memory and retrograde amnesia: Was Tulving right all along? In *Memory, Consciousness, and the Brain: The Tallinn Conference*, E. Tulving, ed. Philadelphia: Psychology Press, pp. 331–345.

OLSON, E. T., 1999. There is no problem of the self. In *Models of the Self*, S. Gallagher and J. Shear, eds. Thorverton, England: Academic, pp. 49–61.

PARKIN, A. J., 1987. *Memory and Amnesia.* New York: Basil Blackwell.

PARKIN, A. J., 1993. *Memory: Phenomena, Experiment and Theory.* Cambridge, England: Blackwell.

PARKIN, A. J., 1996. *Explorations in Cognitive Neuropsychology.* Malden, Mass.: Blackwell.

PARKIN, A. J., and N. R. C. LENG, 1993. *Neuropsychology of the Amnesic Syndrome.* Hillsdale, N.J.: Erlbaum.

PERNER, J., 1991. *Understanding the Representational Mind.* Cambridge, Mass.: MIT Press.

PERNER, J., and T. RUFFMAN, 1994. Episodic memory and auto-noetic consciousness: Developmental evidence and a theory of childhood amnesia. *J. Exp. Child Psychol.* 59:516–548.

ROBINSON, J. A., 1976. Sampling autobiographical memory. *Cogn. Psychol.* 8:578–595.

ROGERS, T. B., 1981. A model of the self as an aspect of the human information processing system. In *Personality, Cognition, and Social Interaction*, N. Cantor and J. F. Kihlstrom, eds. Hillsdale, N.J.: Erlbaum, pp. 193–214.

ROGERS, T. B., N. A. KUIPER, and W. S. KIRKER, 1977. Self-reference and the encoding of personal information. *J. Pers. Soc. Psychol.* 35:677–688.

RYLE, G., 1949. *The Concept of Mind.* New York: Barnes and Noble.

SCHACTER, D. L., and E. TULVING, eds., 1994. *Memory Systems 1994.* Cambridge, Mass.: MIT Press.

SCHELL, T. L., S. B. KLEIN, and S. H. BABEY, 1996. Testing a hierarchical model of self-knowledge. *Psychol. Sci.* 7:170–173.

SIEGLER, I. C., D. V. DAWSON, and K. A. WELSH, 1994. Caregiver ratings of personality change in Alzheimer's disease patients: A replication. *Psychol. Aging* 9:464–466.

SHALLICE, T., 1988. *From Neuropsychology to Mental Structure.* New York: Cambridge University Press.

SHERMAN, J. W., 1996. Development and mental representation of stereotypes. *J. Pers. Soc. Psychol.* 70:1126–1141.

SHERMAN, J. W., and S. B. KLEIN, 1994. Development and representation of personality impressions. *J. Pers. Soc. Psychol.* 67: 972–983.

SHERMAN, J. W., S. B. KLEIN, A. LASKEY, and N. A. WYER, 1998. Intergroup bias in group judgment processes: The role of behavioral memories. *J. Exp. Soc. Psychol.* 34:51–65.

SMITH, E. R., and M. A. ZARATE, 1992. Exemplar-based models of social judgment. *Psychol. Rev.* 99:3–21.

STARKSTEIN, S. E., L. SABE, and M. F. DORREGO, 1997. Severe retrograde amnesia after a mild closed head injury. *Neurocase* 3: 105–109.

SUDDENDORF, T., and M. C. CORBALLIS, 1997. Mental time travel and the evolution of the human mind. *Genet. Soc. Gen. Psychol. Monogr.* 123:133–167.

SYMONS, C. S., and B. T. JOHNSON, 1997. The self-reference effect in memory: A meta-analysis. *Psychol. Bull.* 121:371–394.

TULVING, E., 1983. *Elements of Episodic Memory.* New York: Oxford University Press.

TULVING, E., 1985. Memory and consciousness. *Can. Psychol.* 26:1–12.

TULVING, E., 1993a. What is episodic memory? *Curr. Direct. Psychol. Sci.* 2:67–70.

TULVING, E., 1993b. Self-knowledge of an amnesic individual is represented abstractly. In *Advances in Social Cognition*, vol. 5,

T. K. Srull and R. S. Wyer, eds. Hillsdale, N.J.: Erlbaum, pp. 147–156.

TULVING, E., 1995. Organization of memory: Quo vadis? In *The Cognitive Neurosciences*, M. S. Gazzaniga, ed. Cambridge, Mass.: MIT Press, pp. 839–847.

TULVING, E., and D. L. SCHACTER, 1990. Priming and human memory systems. *Science* 247:301–306.

UTTAL, W. R., 2000. *The War Between Mentalism and Behaviorism: On the Accessibility of Mental Processes.* Mahwah, N.J.: Erlbaum.

VARGHA-KHADEM, F., D. G. GADIAN, K. E. WATKINS, A. CONNELLY, W. VAN PAESSCHEN, and M. MISHKIN, 1997. Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* 277:376–380.

WEISKRANTZ, L., 1997. *Consciousness Lost and Found.* New York: Oxford University Press.

WHEELER, M. A., D. T. STUSS, and E. TULVING, 1997. Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychol. Bull.* 121:331–354.

WIGGINS, J. S., 1973. *Personality and Prediction: Principles of Personality Assessment.* Reading, Mass.: Addison-Wesley.

WILLIAMS, B., 1973. *Problems of the Self.* Cambridge, England: Cambridge University Press.

WOOD, F. B., I. S. BROWN, and R. H. FELTON, 1989. Long-term follow-up of a childhood amnesic syndrome. *Brain Cogn.* 10: 76–86.

ZHANG, L., J. ZHANG, Z. LIU, T. ZHOU, J. FAN, and Y. ZHU, 2002. *In search of the Chinese self: An fMRI study.* Unpublished manuscript.

ZYSSET, S., O. HUBER, E. FERSTL, and D. Y. VON CRAMON, 2002. The anterior frontomedian cortex and evaluative judgment: An fMRI study. *NeuroImage* 15:983–991.