On Bridging the Gap Between Social-Personality Psychology and Neuropsychology

Stanley B. Klein
Department of Psychology
University of California, Santa Barbara

John F. Kihlstrom
Department of Psychology
University of California, Berkeley

Although cognitive psychology has learned much from the study of patients with neuropsychological impairments, social and personality psychologists have been slow to do the same. In this article we argue that the domain of clinical neuropsychology holds considerable untapped potential for formulating and testing models within social and personality psychology and describe some of the ways in which questions of interest to social and personality psychologists can be addressed with neuropsychological data. Examples are drawn from a variety of neuropsychological syndromes, including amnesia, autism, anosognosia, commissurotomy, frontal lobe damage, and prosopagnosia. We conclude that consideration of the personal and social lives of patients with neuropsychological impairments ultimately will lead to a richer understanding of the person, one that bridges the gap between social and cognitive levels of analysis.

For a very long time psychology thought it could get along without looking at the brain. Skinner and other functional behaviorists treated the organism as a "black box" that connected stimuli with responses but whose internal workings could safely be ignored. Classic cognitive psychology and artificial intelligence also endorsed a version of the doctrine of "empty organism" by focusing on the analogy between mind and software and embracing the notion of a Turing machine that could be made out of neurons, silicon chips, or even old radio parts—thus making the biological substrates of mind irrelevant (for a review, see Gardner, 1985).

All that began to change in the mid-1950s, when theory oriented psychologists began to take notice of patients being seen in the neurological clinic and realized that experimental studies of such cases might provide evidence for theories about how the mind is organized (for comprehensive coverage of the neuropsychological syndromes, see Ellis & Young, 1988; Gazzaniga, Ivry, & Mangun, 1998; Heilman & Valenstein, 1993; Kolb & Whishaw, 1996; McCarthy & Warrington, 1990). The most famous case, of course, is the patient known as H.M., who underwent a bilateral resection of the medial portion of the temporal lobes, including the hippocampus and mammillary bodies, in a desperate attempt to ameliorate intractable epilepsy (e.g., Milner, Corkin, & Teuber, 1968; Scoville & Milner, 1957). H.M. emerged from surgery greatly relieved of his epileptic symptoms; the down side was that he now suffered a profound anterograde amnesia, which prevented him from remembering anything that happened to him from the day of surgery until the present.

Studies of H.M., and patients like him, have provided evidence for a medial temporal lobe memory system (e.g., Squire & Zola-Morgan, 1991) that seems to be critical for encoding lasting representations of new experiences. That much is clear, but what exactly

The point of view represented by this article is based in part on research supported by an Academic Senate Research Grant from the University of California, Santa Barbara, and National Institute of Mental Health Grant MH35856.

Preliminary versions of this article were presented at the annual meeting of the Society for Experimental Social Psychology (SESP), Santa Barbara, CA, October 1993 and the 1st Annual SELF preconference at the annual meeting of the SESP, Sturbridge, MA, October 1996.

We thank Jennifer Beer, Lawrence Couture, Elizabeth Gilskey, Martha Gilskey, Judith Lofhus, Sheilagh Mulvaney, Michael Polster, Victor Shamas, Michael Valdisseri, and Susan Valdisseri for their comments.

Requests for reprints should be sent to either Stanley B. Klein, Department of Psychology, University of California, Santa Barbara, CA 93106, E-mail: klein@psych.ucsb.edu, or John F. Kihlstrom, Department of Psychology, MC 1650, Tolman Hall #3210, University of California, Berkeley, Berkeley, CA 94720-1650.
do these neural structures do? This, of course, is a question for psychological theory, and over the years various theories about memory structure and processing have been proposed to explain the behavior of H. M. and others like him.

Initially, amnesic patients were thought to lack a capacity for transferring information from short-term to long-term memory (e.g., Baddeley & Warrington, 1970; Cermak, 1972; Milner, 1966; Wickelgren, 1973). This view, however, soon ran into problems. Consider, for example, what happens when amnesic patient K. C. is tutored in the basics of computer operation (e.g., Glisky, Schacter, & Tulving, 1986b). A few minutes following completion of a lesson, K. C. has no conscious recollections of what he was taught or even that he had a lesson. Nonetheless, he shows clear evidence of having acquired complex knowledge about the programming and operation of a computer; he can understand computer related vocabulary (e.g., software, modem, save, print), can perform disk storage and retrieval operations, and even can be taught to write simple programs (e.g., Glisky, 1995; Glisky, Schacter, & Tulving, 1986a; Glisky et al., 1986b). Yet, if asked how he knows the procedure for downloading a file or writing a program, K. C. is likely to respond that these are just ordinary facts about the world that everyone knows (cf. Tulving, Hayman, & Macdonald, 1991).

Observations such as these (for related findings, see Brooks & Baddeley, 1976; Graf, Squire, & Mandler, 1984; Warrington & Weiskrantz, 1970) suggest that amnesic patients show a dissociation between two forms of long-term memory: Episodic memory, which enables people to become consciously aware of specific past events from their life, is impaired, whereas semantic memory, which enables people to retrieve knowledge abstracted from events but does not entail recollection of the events themselves, is intact (e.g., Cermak, 1984; Evans, Wilson, Wraight, & Hodges, 1993; Kinsbourne & Wood, 1975; Klein, Loftus, & Kihlstrom, 1996; Schacter & Tulving, 1982; Tulving, 1983, 1993; Van der Linden, Bredart, Depoorter, & Coyette, 1996). Characterizing the difference between types or systems of memories is a major growth industry within contemporary cognitive psychology (see, for example, Foster & Jelicic, in press; Schacter & Tulving, 1994), but cognitive psychologists didn't really start asking questions about multiple memory systems until they started contemplating evidence from brain damaged patients (for a review, see Polster, Nadel, & Schacter, 1991). In this way, a great advance in psychological theory began with data from the neurological clinic.

Cognitive psychologists now agree on the value of neurological evidence. For example, psycholinguists are interested in syndromes like Broca's and Wernicke's aphasia for the insights they can provide into the nature of language processing (e.g., Berndt & Caramazza, 1980; Brown, 1972; Goodglass, 1993; Pinker, 1994). Vision scientists are interested in phenomena like blindsight, prosopagnosia, and visual neglect for what they can tell us about perceptual processes (e.g., Coslett, 1997; Prigatano & Schacter, 1991; Weiskrantz, 1997).

The central question addressed in this article is whether the study of patients with neuropsychological impairments should interest personality and social psychologists. We believe the answer to this question is "yes" and hope the arguments we present will challenge our colleagues to join us in considering neuropsychological evidence in theorizing about personality and social processes. Brain damage isn't just for cognitive psychologists anymore; it has a great deal to tell personality and social psychologists about the things we're interested in. However, before that happens, we have to ask the appropriate questions. In the following sections we suggest ways in which neuropsychological evidence can provide important new insights into the role of cognition in personality and social interaction.

H. M., Amnesias, and Knowledge of Self

Reading a case like H. M. can be extremely frustrating to personality and social psychologists because there is so much we want to know, yet so few answers. For example, H. M.'s surgery was in 1953, when he was 27 years old. So what happens now, 45 years later, when H. M. goes into the bathroom to shave in the morning: Does he look in the mirror and say: "Who the hell are you?" What can the self-concept of a person who lacks episodic memory for the past 45 years be like? Can a person preserve a sense of identity, including changes in identity over a long period of time, without also preserving an autobiography? More generally, to what extent is our knowledge of what we are like dependent on our ability to remember the behavioral evidence on which that knowledge is based? Unfortunately, with rare exceptions (e.g., Klein, Loftus, et al., 1996; O'Connor, Cermak, & Seidman, 1995; Tulving, 1993), neuropsychological investigations of the amnesic syndrome...
seldom have considered the impact of catastrophic memory loss on the patient's personal identity.

In the last few years, however, this situation has begun to change as psychologists come to appreciate the ways in which theoretical issues surrounding the self can be addressed with neurological data. Consider, as an example, the case of W. J. (Klein, Loftus, et al., 1996), an 18-year-old undergraduate who suffered a concussive blow to the head shortly after completing her first quarter in college. Brain scans revealed no neurological abnormalities, but she complained of memory and concentration difficulties and testing revealed that she had, in fact, forgotten much of what had happened in her life during the preceding 6 to 7 months—a period of time covering approximately her first quarter at college. Over the next month, W. J.'s amnesia remitted completely.

W. J.'s amnesic deficit in episodic memory was documented by the Galton (1879) memory cueing procedure popularized by J. A. Robinson (1976) and Crovitz (e.g., Crovitz & Quina-Holland, 1976; Crovitz & Schiffman, 1974). In this task, participants are read cue words (representing affects, objects, and activities) one at a time and for each are asked to recall a specific personal event from any time in the past and provide as precise a date as possible for that event. When tested 5 days after her accident, W. J. showed little episodic memory for personal events from recent years. Four weeks later her performance had improved considerably and was indistinguishable from that of three neurologically healthy women who served as controls.

W. J. was also asked both during her amnesia and after its resolution to provide personality ratings describing what she was like during her first term at college. In contrast to the change in her episodic memory performance over the month following her accident, W. J.'s personality ratings of herself at college did not change at all over the same period of time; her trait ratings made during her amnesic period agreed with those she made afterward. Thus, although she was amnesic, W. J. knew what she had been like in college despite the fact that she couldn't recall anything from her time in college.

Of course, it is conceivable that W. J.'s personality didn't change much between high school and college. If so, then she could have achieved reasonably reliable ratings of her personality simply on the basis of her memories from high school, without accessing any information from her college years. To check this possibility, W. J. was asked during her amnesia to rate how she saw herself during high school. Statistical analyses revealed that the correlation between her ratings of herself at high school and ratings of herself at college was reliable ($r = .53$), meaning that some degree of reliability in W. J.'s ratings of her college self could have been achieved by reliance on her memories of her precollege behaviors and experiences. However, this figure was significantly lower than the correlation obtained between W. J.'s two ratings of herself at college, taken during and after her amnesia ($r = .74; p < .05$). Therefore, there is reliable variability in her college self which is not accounted for by her high school self. Put another way, although she was amnesic, W. J. knew something about what she was like in college, which was different from what she was like in high school; she knew this despite the fact that she could not recall any personally experienced events from her time in college.

To explain these findings, Klein, Loftus, et al. (1996; see also Kihlstrom & Klein, 1994, 1997; Klein, 1993, in press; Klein, Babey, & Sherman, 1997; Klein & Loftus, 1993; Klein, Sherman, & Loftus, 1996) proposed that knowledge of personality traits and recollections of specific personal events involving those traits reflect the operations of two distinct, neurally dissociable types of personal memory: semantic personal memory and episodic personal memory (see also Brewer, 1986; Cermak & O'Connor, 1983; Kihlstrom et al., 1988; Tulving, 1993; Wheeler, Stuss, & Tulving, 1996). Episodic personal memory stores the specific details of personally experienced events, whereas semantic personal memory stores generalizations about the self abstracted from those experiences. The fact that during her amnesia W. J. had access to trait abstractions about herself, but not the particular episodes on which that knowledge was based, was taken as evidence that these two types of self-knowledge are served by different neural systems, one of which has become dysfunctional as a result of her concussion, whereas the other remained unimpaired (e.g., Kihlstrom & Klein, 1994, 1997; Klein & Loftus, 1993; Klein et al., 1997).³

³The finding that one function is impaired and another one is spared reveals the basic methodology of cognitive neuropsychology: the functional dissociation (e.g., Shallice, 1988; Teuber, 1955; Weiskrantz, 1989). This term is rather confusing, because to most psychologists the term dissociation refers to the isolation of some percepts, memories, thoughts, or actions from conscious awareness (Kihlstrom, 1993). What neuropsychologists mean by dissociation, social psychologists recognize as the interaction: An independent variable (reflecting a state, condition, or experimental manipulation) affects one dependent variable but not another.

The dissociations that interest neuropsychologists come in four types (e.g., Dunn & Kirchner, 1988; Kelley & Lindsay, 1996; Neely, 1989). In the case of a single dissociation, a single independent variable, A, selectively affects performance on one task, X, but not on another, Y. In the double dissociation, one independent variable, A, affects dependent variable X but not dependent variable Y, whereas another independent variable, B, affects Y but not X. The double dissociation can be uncrossed or crossed: Crossed double interactions are especially interesting to neuropsychologists, because they are especially good evidence that two different processes are involved in the two tasks. Otherwise one would worry about artifacts like differential task difficulty. Even better evidence is provided by the reversed association, in which there is a positive correlation between dependent variables X and Y under conditions of independent variable A, but a negative correlation between these same variables under conditions of independent variable B. Reversed associations are particularly difficult to account for in terms of task difficulty (e.g., Dunn & Kirchner, 1988; Klein et al., 1997; Neely, 1989).
Admittedly, Klein, Loftus, et al.'s (1996) conclusion could be questioned on the basis of W. J.'s continued access to episodic memories that were not covered by her amnesia and the possibility that she drew on those memories, not her semantic personal knowledge, for her ratings of self-at-college. However, there is other evidence indicating that accurate self-description can occur even with total episodic memory loss. Tulving (1993), for example, found that patient K. C., who lost his entire fund of episodic memory (and underwent a marked personality change) following a motorcycle accident, was able to describe his postmorbid personality with considerable accuracy. Tulving asked K. C. to judge a list of trait adjectives for self-descriptiveness. Tulving also asked K. C.'s mother on two separate occasions to rate K. C. on the same traits, the first time rating K. C. as he currently was and the second time rating him as he was before his accident. K. C.'s choices were highly correlated with his mother's judgments of his postmorbid personality, but not with her judgments of his premorbid personality. Thus, K. C. was able to acquire accurate knowledge of his new personality (with his mother's ratings serving as the criterion) without being able to retain any episodic knowledge of the specific actions and experiences on which that knowledge was based.

Although theorists differ concerning the precise interpretation of the findings just discussed (e.g., Schneider, Roediger, & Kahn, 1993), this much is clear: Neurologically impaired individuals who have lost the ability to recall personal experiences show no obvious impairment in the ability to make accurate personality judgments about themselves, and (in the case of K. C.) even maintain the ability to revise those judgments based on new episodes that they cannot remember. Apparently you do not need to remember how you behaved in the past to know what you are like.

Additional support for this conclusion recently was presented by Craik et al. (in press). Using positron emission tomography (PET), 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).4

The dissociations between episodic and semantic self-knowledge have made several things clear. First, contrary to long-held beliefs about the memorial basis of self (e.g., Grice, 1941; James, 1890; Keenan, 1993; Locke, 1690/1731; Quinton, 1962; Tulving, 1984), episodic memory is not the sole repository of self-knowledge. The fact that a loss of episodic memory does not lead to a complete loss of self-knowledge has led theorists to expand the basis of self-knowledge to include both episodic and semantic memory (e.g., Cermak & O'Connor, 1983; Conway, 1992; Evans et al., 1993; Klein & Loftus, 1993; Klein, Loftus, et al., 1996; Tulving, 1993; Tulving, Schacter, McLachlan, & Moscovitch, 1988). Second, the finding that individuals can have accurate and detailed knowledge of their personalities despite having no conscious access to behavioral episodes suggests these two types of self-knowledge are represented independently in memory and perhaps mediated by separate cognitive systems.

Over and above these specific issues, the analysis of cases like W. J. and K. C. shows a little of what is possible when neurological disorders are approached with personality and social theories in mind. We hope that these studies stimulate other self-theorists to consider the theoretical promise of patients with neuropsychological impairments, for it would seem there is much such patients can teach us about the representation and function of knowledge about the self.

Autism, Theory of Mind, and Theory of Self

An interesting implication of the proposal that episodic and semantic self-knowledge are served by different cognitive systems is that a person could, in principle, have complete access to his or her episodic self-knowledge yet be unable to know whether a particular trait adjective was descriptive of self. Although this question has not been addressed empirically, some intriguing hints at an answer are found from a rather unusual source—the study of patients with autism (Klein, 1996).

In a series of publications, Baron-Cohen, Leslie, U. Frith, and colleagues (e.g., Baron-Cohen, 1989, 1990, 1991, 1995; Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, Tager-Flusberg, & Cohen, 1993; U. Frith, 1989; Leslie, 1987, 1991; Leslie & Frith, 1988; Leslie & Thaiss, 1992) have argued that a defining feature of the autistic syndrome is the failure of autistic individuals to develop what Premack and Woodruff (1978) termed a theory of mind—a capacity to attribute mental states (e.g., intentions, desires, thoughts, beliefs) to other persons in order to make sense of their behavior (see also, Flavell, Green, & Flavell, 1995; Gopnik & Metzloff, 1997; Wellman, 1990).

Leslie (1987), in a pioneering paper on the topic, suggested that the failure of autistic individuals to explain behavior in terms of mental states (i.e., to mentalize) stemmed from their inability to form "sec-

4Craik et al. (in press) also concluded that the cognitive processes involved in self-reference were no different from those involved in referring to other individuals or in performing nonsocial semantic analyses. Thus, the neuroscience method of brain imaging confirmed conclusions that already had been reached on the basis of traditional experimental procedures employing behavioral measures (e.g., Kihlstrom et al., 1988; Klein & Kihlstrom, 1986).
ond order representations." By this account, autistic individuals are able to form "first order representations" of people, things, and events based directly on perceptual experience (e.g., "Robert smiled when he got the candy bar"). They are, however, deficient in forming second order representations—that is, representations of first order representations (e.g., "Robert smiled because he thought [or knew, or hoped, or believed] he would get the candy bar"). The capacity to represent representations, Leslie argued, is the essence of a theory of mind and is a capacity that fails to fully develop in autistic individuals (a recent review can be found in Baron-Cohen, 1995).

What about the autistic individual's awareness of his or her own mental states? If autism involves a dysfunction of the neural structures necessary for forming a theory of other minds, it is reasonable to wonder whether these individuals might also show an impaired ability to reflect on their own mental states—to know about their own knowing.3

Surprisingly, the question of whether the problems autistic patients experience in understanding and recognizing mental states in others extend to their understanding of their own mental states has been largely overlooked (for a recent discussion, see Carruthers, 1996). However, the few empirical findings that are available do suggest that autistic patients have trouble reflecting on their own mental states (e.g., Baron-Cohen, 1989, 1991; Baron-Cohen, Ring, Moriarty, Schmitz, Costa, & Ell, 1994; Jordan, 1989; Tager-Flusberg, 1992). For example, several recent studies reported that compared to normally developing children, autistic children have problems in acquiring a normal grasp of the personal pronouns / and me (e.g., Fay, 1979; Jordan, 1989; Lee, Hobson, & Chiat, 1994). Tager-Flusberg (1992) showed that autistic individuals use significantly less spontaneous speech than matched controls when referring to their own cognitive mental states (e.g., beliefs, desires, traits). Along similar line, Baron-Cohen (1991) found that autistic individuals have as much trouble attributing beliefs to themselves as they do in attributing beliefs to others. Finally, clinical descriptions of autistic individuals often make reference to their inability to self-reflect or to self-monitor (e.g., Baron-Cohen, 1989; Bishop, 1993).

Admittedly, the evidence that autistic individuals may be lacking in awareness of their own mental states is small, indirect, and often anecdotal. Nonetheless, if this hypothesis is correct, it suggests the interesting possibility that an autistic individual, although capable of recalling trait-relevant personal behaviors (e.g., "I remember getting a high score on a math test").), may be unable to make trait-based generalizations about the self on the basis of those behaviors (e.g., "I know [or think, or hope, or believe] that I am an intelligent person"). If such an outcome were obtained, it would provide strong converging evidence in support of the proposed independence between episodic and semantic self-knowledge.

Self-Awareness and the Brain: Locating the Jamesian Self-as-Knower

In light of the previous discussion, it is interesting to wonder whether we know enough about the neural correlates of mentalizing to identify where in the brain such capacity resides. Although a definitive answer is not yet available, some fascinating clues can be found. For example, neuroimaging studies conducted on individuals engaged in theory of mind tasks (e.g., tasks requiring inferences about mental states) report evidence for selective activation of the frontal lobes during task performance, suggesting a role for these structures in the capacity to mentalize (e.g., Baron-Cohen et al., 1994; Fletcher et al., 1995; Goel, Grafman, Sadato, & Hallett, 1995). This possibility receives support from two additional sources. First, there is some evidence that patients with frontal lobe damage show deficits on theory of mind tasks (e.g., Price, Daffner, Stowe, & Mesulam, 1990; Stone, Baron-Cohen, & Knight, 1996). Second, a number of investigators have noted strong parallels between the behavior of autistic individuals and that of patients suffering frontal lobe damage (e.g., Bishop, 1993; Damasio & Maurer, 1987; C. D. Frith & U. Frith, 1991; Ozonoff, Pennington, & Rogers, 1991; Prior & Hoffman, 1990). Specifically, both groups show (a) limited ability to plan for the future, or to anticipate the long-term consequences of their behavior, (b) deficits in the capacity to self-reflect or self-monitor, and (c) difficulties learning from mistakes, persevering with maladaptive strategies even when repeatedly made aware of their errors (for comprehensive reviews, see Damasio, 1985, and Fuster, 1997).

3Nicholas Humphrey's (1984, 1986, 1990) recent writings on the evolution of self-awareness in humans are suggestive of such a possibility. According to Humphrey, self-awareness, having been designed by natural selection, must contribute to our biological success. However, what selective advantage is provided by an ability to reflect on one's own mental states?

Humphrey proposed that the answer is to be found in the social challenges faced by our ancestors. From their initial appearance approximately 150,000 years ago (e.g., Dunbar, 1996), modern humans lived in a highly complex, interpersonal milieu; accordingly, their survival depended on their being able to explain, predict, and manipulate the behavior of others. Self-reflective awareness served this function: By showing us how our own mind works, it provided us, by analogy, with a tool for understanding the minds of others like ourselves (for a related view, see Sedikides & Skowronska, 1997). Thus, a necessary precondition for developing a theory of other minds is the possession of a theory of one's own mind. By implication, the absence of a theory of other minds may be diagnostic of a failure to develop a theory of self.
Interestingly, the psychological processes compromised in patients with frontal lobe dysfunction—the capacity to monitor and reflect on one's mental states—are defining features of James's (1890) *self-as-knower*, the subjective experience of self as a thinking, feeling, wanting, doing being. Although there is much we do not understand about this self-reflective aspect of self (for discussions, see Greenwald & Pratkanis, 1984; Kihlstrom & Klein, 1994; Stuss, 1991), we are perhaps a step closer to knowing where in the brain such a capacity resides. By capitalizing on what we know about frontal lobe function in both normal and brain damaged individuals, we may come to a better understanding of the structure and function of this most elusive of Jamesian concepts.

Anosognosia and Attribution Theory

H. M. is aware of his memory deficit (he describes it as "like waking from a dream"), and he knows that there are things that he can't remember. However, there are other patients suffering from a variety of problems with memory, language, perception, or voluntary movement who appear to have no awareness of their deficits. This lack of awareness of a mental deficit was named anosognosia by Babinski (1914, 1918; for recent reviews, see McGlynn & Schacter, 1989; Prigatano & Schacter, 1991). Anosognosic patients may acknowledge some difficulty in their impaired domains, but they attribute their problems to something besides their own deficits. It should be understood that these patients' behavior is not mere denial of deficit or indifference to it (when a patient acknowledges deficit but seems unconcerned about it, the syndrome is called anosodiaphoria).

Most of the classical descriptions of anosognosia are in cases of acute hemiplegia, hemianesthesia, and hemianopia (Bisiach & Geminiani, 1991). In hemiplegia, the person is paralyzed on one side of the body, due to damage to the contralateral hemisphere; in hemianopia, the person has a loss of sight in one side of the visual field. Interestingly, anosognosia is more likely to occur when the loss is localized on the left side of the body, implying that it is caused by damage to structures in the right cerebral hemisphere. However, the syndrome also occurs in cases of left-hemisphere damage, as for example in aphasia (e.g., Rubens & Garrett, 1991). Many aphasics, both expressive and receptive, attempt to correct their faulty speech production; by virtue of hesitations, pauses, and self-corrections they show clearly that they know that what they have intended to say hasn't come out as planned. However, many do not realize this, a failure that is common in cases of jargon aphasia, a special form of receptive aphasia in which the patient's speech is freely littered with meaningless utterances and phonemic and semantic paraphasias (using the wrong sounds or words). Such patients do not seem to realize that they are not communicating with their listeners, and, furthermore, they do not seem to realize that they don't understand what is being said to them. Interestingly, jargon aphasia seems to be more common in cases of bilateral damage; again, this implies that the right hemisphere plays a special role in awareness of deficits.

Anosognosia is a real danger to the patient, of course. People who don't realize that they are paralyzed on one side are headed for disaster if they should try to get up; those who don't realize they are blind on one side are unlikely to take special steps to avoid obstacles and oncoming objects on the affected side. In the dememting disorders, such as Alzheimer's disease and even schizophrenia, anosognosia is particularly insidious because it occurs in the late stages of illness (e.g., McGlynn & Kaszniak, 1991), when the patient is most impaired. Interestingly, however, anosognosics sometimes implicitly acknowledge their difficulties. The hemiplegic may not complain about being confined to a hospital bed or attempt some task that must be performed with both hands, and the hemianopic may actively ignore the affected portion of the environmental field. Neurological patients who are unaware of their deficits are poorly motivated for rehabilitation.

From a social-psychological view one wants to know what these patients make of their own behavior, given that they don't acknowledge their deficits. Some patients attribute their inability to move to arthritis or rheumatism rather than paralysis; others, when asked to move the affected limb, appear distracted or move the unaffected limb or respond that they have moved the affected limb, when in fact they have not (this even happens when patients look at the affected limb during the examination). The explanations can sometimes become bizarre or delusional.

For example, the patient may claim that the affected limb is not his or her own, but rather belongs to someone else—forgotten by a previous patient or belonging to someone else lying at their side (often doing something naughty). One woman studied by Bisiach and Geminiani (1991) was anosognosic for her hemiplegia. She claimed that her left hand did not belong to her, but rather had been forgotten in the ambulance by another patient. She acknowledged that her left shoulder was her own and agreed with the inference that her left arm and elbow were also her own, because they were attached to her shoulder, but this inference did not extend to her left hand (she could not explain why that hand carried her wedding ring). Another hemiplegic patient stated that his own left arm was the examiner's. When the examiner placed the patient's left hand between his own two hands, the patient continued to deny that his arm hand was his own and attributed three arms and three hands to the examiner.
What we’re seeing here, of course, are phenomena of attribution; the patients are trying to make sense of their experiences, given their beliefs about themselves and the world at large. These attributions may be convenient laboratory models for other kinds of beliefs, including those that are frankly delusional (for a review of attributional accounts of delusions, see Kihlstrom & Hoyt, 1988). Consider the following scenario: A hemiplegic patient is unaware of the loss of function on his left side and denies that his left arm and hand are his. Then what’s he doing in bed? Why is someone else wearing his wedding ring? Where is the rest of that person, anyway? If he’s forgotten his left hand, doesn’t he miss it? Why doesn’t the patient retrieve his wedding ring and put it back on his own left hand? Anomalous perceptual experiences arouse anxiety until they are satisfactorily explained, and in the course of formulating acceptable explanations, the patient must go through the sorts of processes studied by attribution theory. Accordingly, cases of anosognosia can provide an interesting proving ground for testing and refining theories about causal attributions, self-other differences, and other aspects of social judgment and inference.

Split Brains and Self-Perception

Few neuropsychological syndromes have generated greater interest among neuroscientists (e.g., Gazzaniga, 1970; Sperry, 1968, 1974; Springer & Deutsch, 1998) and philosophers (e.g., Marks, 1981; Nagel, 1971; Puccetti, 1973) than that of the commissurotomized (colloquially referred to as split-brain) patient. These patients have suffered from severe and uncontrollable epileptic seizures, much like those experienced by H. M., but their treatment is quite different. In an effort to alleviate the effects of otherwise intractable epilepsy, a procedure known as a complete cerebral commissurotomy is performed (e.g., Bogen, Fisher, & Vogel, 1965; Bogen & Vogel, 1962) in which the corpus callosum, a large transverse band of approximately 200 million nerve fibers that directly connect the left and right cerebral hemispheres, is surgically cut. Because epileptic seizures, which originate as electrical outbursts at a particular cortical site, tend to spread from one cerebral hemisphere to the other (thereby increasing the magnitude of the disturbance), cutting the corpus callosum is seen as a way of limiting the disturbance to one hemisphere, thereby decreasing its magnitude (e.g., Gazzaniga & LeDoux, 1978; Kolb & Whishaw, 1996; Sperry, 1974).

Medically, complete cerebral commissurotomy proved quite successful; confined to a single hemisphere, patients’ epileptic seizures became less frequent or disappeared entirely (e.g., Kolb & Whishaw, 1996; Springer & Deutsch, 1998). Moreover, initial reports revealed no obvious postsurgical changes in their perceptual, cognitive, or everyday behavior (e.g., Akelaïtis, 1941a, 1941b, 1944; Bogen, 1985).

However, extensive psychological testing by Roger Sperry and his colleagues (e.g., Franco & Sperry, 1977; Levy-Agresti & Sperry, 1968; Sperry, 1968) eventually uncovered some peculiar psychological consequences of hemispheric disconnection. Sperry’s approach to testing split-brain patients depended on two key assumptions. First, that under suitable experimental control, it is possible to confine input presented to a split-brain patient to a single hemisphere (e.g., Sperry, 1968, 1974). Second, that in the vast majority of people, verbal reports issue from the left cerebral hemisphere. By contrast, the right hemisphere, although capable of limited linguistic analyses, lacks access to the speech mechanisms of the left hemisphere and thus is unable to initiate speech (e.g., Corballis, 1991; Kolb & Whishaw, 1996; Springer & Deutsch, 1998).

Using several subtle experimental techniques, Sperry and his colleagues were able to direct input exclusively to a single hemisphere and request a response of it (e.g., Franco & Sperry, 1977; Gordon & Sperry, 1969; Levy, Trevarthen, & Sperry, 1972; Levy-Agresti & Sperry, 1968; Sperry, 1968, 1974; Sperry, Gazzaniga, & Bogen, 1969; Zaidel, 1975). For example, when an object was visually presented to the left hemisphere, split-brain patients reported seeing it and could identify it verbally. However, when the same object was presented to the nonspeaking right hemisphere, patients claimed they saw nothing at all. Nevertheless, the right hemisphere could demonstrate nonverbally what it had seen by pointing at the correct object with the left hand (which is controlled by the right hemisphere). Similar findings were obtained using olfactory stimuli. When a clove of garlic was presented to a split-brain patient’s right nostril (which stimulates the right hemisphere), he verbally denied smelling anything. However, when asked to point with his left hand to the object corresponding to the odor he smelled, he correctly selected the clove from among a set of smell related objects, at the same time verbally protesting that he didn’t smell anything!

Findings such as these led Sperry to propose that surgery had left split-brain patients with two separate minds, each with its own separate sphere of consciousness (e.g., Sperry, 1966, 1968, 1974). In Sperry’s (1968) words:

---

4In addition to sectioning the corpus callosum, the Bogen and Vogel (1962) procedure also involved complete sectioning of the anterior and hippocampal commissures. It is via these three links that direct interhemispheric communication and integration takes place.

5Strictly speaking, the designation split-brain surgery is a somewhat of a misnomer—although the corpus callosum and minor commissures are surgically severed, the subcortical regions linking the two hemispheres are left untouched by the surgery.
Each hemisphere seems to have its own separate and private sensations; its own perceptions; its own concepts; and its own impulses to act, with related volitional, cognitive, and learning experiences. Following surgery, each hemisphere also has thereafter its own separate chain of memories that are rendered inaccessible to the recall processes of the other. (p. 724)

A particularly intriguing case is that of patient P. S. (e.g., Gazzaniga & LeDoux, 1978; LeDoux, 1985; LeDoux, Wilson, & Gazzaniga, 1977). P. S. is unique among split-brain patients in that his right hemisphere, although unable to generate speech, has extensive linguistic abilities, enabling it to respond to a wide variety of verbal commands. For example, when the experimenters asked his right hemisphere to "laugh," it did as told and P. S. laughed aloud. Interestingly, however, when asked why he was laughing, the left hemisphere replied "Oh you guys are really something" (Gazzaniga & LeDoux, 1978, p. 146). In another study the experimenters simultaneously flashed an image of a snow scene to P. S.'s right hemisphere and an image of a chicken claw to his left hemisphere. Each hemisphere then was shown a set of pictures and instructed to select the one most closely associated with the image it had seen. The right responded by choosing (with his left hand) a picture of a shovel, and the left selected (with the right hand) a picture of a chicken to match the claw. When asked why he chose these particular pictures, his left hemisphere responded "I saw a claw and I picked a chicken, and you have to clean out the chicken shed with a shovel" (Gazzaniga & LeDoux, 1978, p. 148).

In each of these examples, P. S.'s left hemisphere was faced with a problem—it had observed a response but did not know why the response was performed. When asked "Why are you doing that?", the talking left hemisphere had to come up with a plausible explanation for a behavior performed in response to a command directed to the mute right hemisphere. As Gazzaniga and LeDoux (1978) noted, the left hemisphere proved quite adept at this task, interpreting the actions of the right as though it had insight into the cause of the behavior (when in fact it did not). On the basis of these findings, Gazzaniga and LeDoux concluded that the left hemisphere acts as the interpreter of action, attempting to provide as plausible an account as possible for the individual's behavior (for related views, see Jaynes, 1976; Popper & Eccles, 1977; Sperry, 1974).

Gazzaniga and LeDoux (1978; see also LeDoux, 1985) go on to suggest that the left hemisphere plays a similar role in individuals with intact brains. A considerable body of evidence suggests that we are not consciously aware of the causes of all the behaviors we produce or feelings we experience (for reviews, see Gazzaniga, 1998; Kihlstrom, 1987, in press-a, in press-b; Nisbett & Wilson, 1977; Oakley & Eames, 1985; Velmans, 1996). When an activity is initiated by a neural system whose motives are not consciously accessible, the verbal left hemisphere finds itself confronted with behavior carried out for unknown reasons. Under these circumstances, it attempts to attribute a cause to the action, thereby integrating the action into a coherent personal narrative (e.g., Gazzaniga, 1998; LeDoux, 1985). As Gazzaniga and LeDoux (1978) remarked: "It is as if the verbal self [i.e., left hemisphere] looks out to see what the person is doing, and from that knowledge it interprets reality" (p. 150).8

What Gazzaniga and LeDoux have provided us with, of course, is a neuropsychological model of Bern's (1967, 1972) influential theory of self-perception—the idea that people "come to know their own attitudes, emotions, and other internal states partially by inferring them from observations of their own overt behavior and/or the circumstances in which this behavior occurs" (Bern, 1972, p. 5). We believe that such a model can contribute in important ways to our understanding of the process involved in self-perception. The relation between lateralization and hemispheric specialization is becoming increasingly well-mapped experimentally (for a recent review, see Springer & Deutsch, 1998). By drawing on that knowledge, self-perception theorists may gain an understanding of the functional properties of the neural system responsible for drawing inferences about behavior whose origins are outside conscious awareness—an understanding that ultimately may lead to a better appreciation of the ways in which individuals attempt to construct a coherent story of self. And, by learning which of a person's behaviors are likely to be initiated by the nonverbal right hemisphere, theorists may be better able to identify the types of behaviors whose explanation require the inference-making capacities of the left hemisphere. Although the contributions of these particular neuropsychological perspectives on self-perception theory remain to be determined, it seems clear to us that, in the long run, research and theory both will benefit from a greater understanding of the neuropsychological mechanisms that make self-perception possible.

Phineas Gage and the Question of Cognition and Emotion

The relevance of neuropsychology for personality and social psychology is also illustrated by the classic case of Phineas Gage (e.g., Macmillan, 1986). In 1848,
this young railway worker was preparing some explosive charges for use in an excavation. In so doing, he accidentally set off a spark that exploded the gunpowder, driving his custom-made tamping iron right through his skull—entering under his left eye socket, traveling behind his eye (severing the optic nerve), and emerging from the top of his head. Gage lived for a dozen more years, which is extraordinary in itself, but he also showed a marked change in personality. Whereas before the accident he had been described as shrewd, smart, energetic, and persistent, he now was described as fitful, irreverent, grossly profane, lacking in deference, impatient, obstinate, capricious and vacillating, childlike in his intellectual capacity, and with strong animal passions—in short, as the physician who attended his wounds put it, he was “no longer Gage” (Harlow, 1868).

The significance of the Gage case was not lost on the phrenologists. Nelson Sizer (1882), an American disciple of Gall and Spurzheim, concluded that the injury had obviously destroyed brain tissue “in the neighborhood of Benevolence and the front part of Veneration” (pp. 193–194). Even after the abandonment of phrenology, the Gage case was used, along with Broca’s and Wernicke’s cases of expressive and receptive aphasia, as a primary example of specialization of function in the cerebral cortex—in particular, for the localization of faculties relating to personality, social relationships, and emotion in the frontal lobe. Modern neuropsychology has generally confirmed this conclusion, although we now know that the frontal lobes support cognitive as well as emotional and interpersonal functions (for a review of other cases of frontal lobe damage, see Damasio, 1985).

Neuropsychological evidence also can be brought to bear on the vexatious question of the relation between emotion and cognition. Is emotion a cognitive construction or an independent mental faculty? Although cognitive processing undoubtedly plays a role in emotion (e.g., Clark & Fiske, 1982), neuropsychological evidence does seem to show that some brain structures are specialized for emotion and for the processing of emotional as opposed to nonemotional memories. Consider, for example, the Klüver-Bucy syndrome (Kluver & Bucy, 1939), resulting from bilateral destruction of the amygdala and associated inferior portions of the temporal cortex. Humans and nonhuman animals with such lesions show a loss of fear and other emotional responses and increased and inappropriate sexual activity, among other symptoms. Such outcomes suggest that the amygdala plays a special role in emotion, a hypothesis that has been supported by LeDoux’s (1987, 1996) finding that bilateral amygdaloectomy impairs classical fear conditioning. Of course, cortical structures also are involved in emotion; patients with lesions in the right hemisphere, and in particular the temporal-parietal regions of the right hemisphere, have special difficulties in judging the mood of others from vocal or facial cues, choosing which uncaptioned cartoons are funny, selecting the correct punchlines to joke set-ups, and matching scenes for emotional valence. Cognition and emotion are certainly related, but the neuropsychological evidence seems to indicate that there are certain brain systems that are specialized for emotional processing, suggesting that cognition and emotion are also different mental faculties.

Prosopagnosia and Face Recognition

Neuropsychological evidence would seem to be especially relevant to understanding a basic social-cognitive process: how one person recognizes the face of another. The face is the fundamental social stimulus. It is the point of contact in the infant’s very earliest social interactions; the smiles exchanged between infant and caregiver are the beginnings of lifelong social bonds. Perceiving, identifying, and comprehending faces is absolutely basic to social interaction. We have to know who we are dealing with, what they are like, and how we relate to them, before we can interpret their behavior or plan our own. Even when dealing with strangers, the face provides cues to the emotional state of the other person, as well as hints of other things, like deception, that are important in negotiating an interaction. If we want to understand how we come to know another person, we have to understand how we read the face.

As it happens, neuropsychology has been very interested in the face, and, in fact, there is a specific form of visual agnosia involving the face. In general, visual agnosia—a term coined by Sigmund Freud (1891/1953) before he turned from neurology to psychoanalysis—refers to the inability to recognize objects (or pictures of objects). A person with visual object agnosia can describe an object, but cannot name it, recognize it as familiar, or demonstrate how it is used. Visual agnosia specific to the face is called prosopagnosia, a term coined by Bodamer (1947; for a recent review,
see Damasio, Damasio, & Van Hoesen, 1982), and refers to the inability to recognize faces. Prospagnosic patients can describe the physical features of faces, but they cannot name the individuals to whom they belong; interestingly, they are able to identify people from such characteristics as their voice, dress, posture, or gait. However, given the face alone, these patients have no idea who the person is or what to expect from them. This deficit is linked to bilateral damage in the occipital lobe, especially those areas adjacent to the temporal lobe. As social animals, we seem to have been built by evolution with brain structures specifically tuned to that most social of stimuli, the face (e.g., Brothers, 1997).

Prospagnosia has been taken as evidence that there is a particular brain system specialized for the identification of faces (e.g., Farah, 1990; Farah, Wilson, Drain, & Tanaka, 1998). This proposal is not unreasonable, given our status as social animals and the obvious evolutionary advantages of being able to identify faces and discriminate among them, quickly and reliably. However, there is an interesting controversy here. It has been suggested that prosopagnosics have difficulty identifying any particular visual stimulus, not just faces. Unfortunately, the clinical evidence is equivocal. One prosopagnosic farmer was unable to recognize his own cows, as well as members of his own family (Bornstein, Sroka, & Munitz, 1969), whereas another prosopagnosic farmer—What’s the chance of that?—lost the ability to recognize both his family members and his cows, but eventually recovered the former but not the latter (Assal, Favre, & Anderes, 1984). However, case studies are always difficult to interpret, and recent experimental and neuroimaging studies of both prosopagnosic patients and intact participants strongly suggests that the “face area” damaged in prosopagnosia is actually specialized for expert recognition of objects at subordinate levels of categorization—objects which include, but are not limited to, faces (Gauthier, 1998).

Although prosopagnosia is dramatic, it turns out that there are many different forms of facial agnosia, each reflecting the selective impairment of some functions and the sparing of others. In general, the finding that two functions are dissociable from each other supports the hypothesis that the functions in question are qualitatively different. For example, some prosopagnosics are able to interpret the emotional meaning of facial expressions without being able to recognize the faces themselves, and others recover the ability to identify familiar faces but not the ability to interpret facial expressions. Interestingly, single-unit analyses of face perception in monkeys finds separate neurons (or, more likely, separate clusters of neurons) that are responsive to identity and expression.

In an attempt to summarize the neuropsychological evidence, Bruce and Young (1986) proposed that facial perception involves several different processes that are carried out in parallel. According to their view, input from a facial stimulus is first processed by a structural encoding system that creates two different descriptions of the face—one which is view-centered (e.g., full-face or profile) and one that is independent of the particular expression on the face. Output from this structural encoding system then passes to other systems specialized for analysis of facial expressions, speech (actually, lip-reading), sameness or difference (as between full-face or profile views), and facial recognition. These functions are performed by separate systems, as indicated by the fact that they are dissociable. For example, prosopagnosic patients can identify facial expressions even though they do not recognize the faces as familiar, and there is at least one patient who has lost the ability to analyze facial expressions of emotion, but who retains the ability to lip-read.

In addition, among brain-damaged patients performance on a test of memory for unfamiliar faces is essentially unrelated to the ability to recognize famous faces. All of these results indicate that remembering unfamiliar faces and recognizing familiar faces are mediated by separate systems. The face recognition system is a sort of visual lexicon containing template representations of familiar faces. Information processed by the face recognition system then contacts associated information pertaining to the identity of the person whose face has been recognized and by this route retrieves the name associated with the familiar face. Thereafter, other information about the person is retrieved through the generic cognitive system. Note that the general cognitive system can influence some facial processing systems (e.g., expression analysis, facial speech analysis, and directed visual processing), but it cannot directly influence facial recognition. That influence must be mediated by cognitive activation of the person-identity nodes.

A model like this makes some interesting predictions about face processing that should interest social psychologists. For example, priming with the name attached to a face should influence face recognition, but not expression analysis; priming with the label of an emotional state should influence expression analysis, but not face recognition. We don’t know yet whether this is true. A prediction that has been tested, however, is that familiarity should influence identity matching but not expression analysis. In an experiment performed by Young, Ellis, and their colleagues (e.g., Young, McWeeny, Hay, & Ellis, 1986), intact participants were asked whether two photographs showed the same type of facial expression or whether they showed the same person. Half the photographs were of individuals who were familiar from the news or entertainment media, the other half were mere mortals. In terms of response latencies, familiarity affected identity matching but not expression matching. This is especially interesting, insofar as other research indicates that facial ex-
pression analysis and face recognition rely on the same facial features. Although these features may be analyzed by a single structural encoding system, the output from this module appears to be passed to different task-specific systems operating in parallel.

We don't mean to imply that the Bruce–Ellis model has been tested and proven in every respect; it hasn't, and it might be wrong in significant ways. The point is only to show how neuropsychological evidence can contribute to social-psychological theory: First, by providing empirical evidence of a sort that would be difficult or impossible to obtain in laboratory studies of college sophomores or interviews of people in airports and laundromats; second, by providing specific theoretical models of cognitive processes that can be tested in laboratory studies of the sort that we do.

**Toward a Social Neuropsychology**

One of the most exciting trends in cognitive psychology over the past 20 years has been the increasing application of data and conceptual tools derived from the study of patients with neuropsychological syndromes. To date, however, social and personality psychologists have rarely considered neuropsychological case material when developing theories about social and personality processes. We hope this situation changes, for we believe the domain of clinical neuropsychology holds considerable untapped potential for formulating and testing models within personality and social psychology.

In this article we have described some of the ways in which questions of interest to social and personality psychologists can be addressed with neuropsychological data. For example, we have shown (a) how studying both the preserved and impaired capacities of patients suffering amnesia and autism can provide important new insights into the mental representation of self, (b) how understanding the ways in which anosognosic patients attempt to make sense of their disabilities can shed new light on the process of causal attribution, and (c) how consideration of the data from frontal lobe patients can help address questions concerning the relation between cognition and emotion. As we hope our review shows, there clearly is much social and personality psychologists can learn from the study of patients with neuropsychological syndromes.

Although our focus in this article has been on ways in which personality and social psychology can benefit from a consideration of neuropsychological case material, we also are convinced that neuropsychological theory and research can benefit from insights derived from personality and social psychology. To date, almost all of the work on patients with neuropsychological impairments has been done within the confines of cognitive psychology and cognitive neuropsychology, with relatively little attention paid to the interpersonal, emotional, and motivational lives of these individuals. Yet, the syndromes described in this article invariably are accompanied by profound changes in the individual's personal, social, and professional life (e.g., Blumer & Benson, 1975; Damasio, 1994; Hits, 1995; Luria, 1972; O'Connor et al., 1995; M. F. Robinson & Freeman, 1954; Sacks, 1985), changes that have important implications for the way we approach treatment, conduct research, and formulate theory. Thus, it would seem that an important agenda item for the near future would be the adoption by cognitive neuropsychologists of the concepts and principles that have served their social and personality colleagues so well and the systematic extension of research on neuropsychological impairment beyond the purely cognitive to include the personal and social.

The study of the interpersonal and emotional lives of patients with neuropsychological syndromes promises to provide new perspectives on the relation between cognitive neuropsychology and social-personality psychology. However, this will not happen until psychologists interested in social and personality issues start considering neuropsychological case material and psychologists interested in neuropsychology begin to inquire into their patients’ personal and social lives. Such an interdisciplinary approach is exciting because it would represent the beginning of a collaboration that ultimately might bridge a gap between social-personality psychology and cognitive neuropsychology.

**References**


Akelaitis, A. J. (1941b). Studies on the corpus callosum: II. The higher visual functions in each homonymous field following complete section of the corpus callosum. *Archives of Neurology and Psychiatry, 45*, 788–796.


239


