Dissociations and Dissociation Theory in Hypnosis: Comment on Kirsch and Lynn (1998)

John F. Kihlstrom
University of California, Berkeley

I. Kirsch and S. J. Lynn's (1998) critique of the neodissociation theory of divided consciousness fails to consider evidence of dissociations between explicit and implicit memory and perception in hypnosis. Contrary to their conclusions, evidence that the rate of hidden observer response (like other hypnotic responses) varies with the wording of instructions does not contradict neodissociation theory; rather, it underscores the fact that hypnosis entails social interaction as well as alterations in conscious awareness. Neodissociation and socio-cognitive theories of hypnosis complement each other. Each draws attention to aspects of the experience of hypnosis that the other neglects.

For more than 25 years, Hilgard’s (1973, 1977/1986, 1991, 1994) neodissociation theory of divided consciousness has been an influential analysis of what happens when a highly hypnotizable person responds to suggestions for alterations in perception and memory. But the impact of neodissociation theory has not been confined to hypnosis researchers. Hilgard’s theory, reviving a point of view first put forth in the late 19th and early 20th centuries by Pierre Janet (1889, 1907), William James (1890/1980; Taylor, 1983), Boris Sidis (1902), and Morton Prince (1896, 1914), helped break the death grip in which Freudian psychoanalysis had long held psychologists' notions of unconscious mental life. It laid much of the foundation for the renewed interest in the psychological unconscious among cognitive and social psychologists (e.g., Cohen & Schooler, 1997; Kihlstrom, 1987, 1990, in press-a; Uleman & Bargh, 1989; Underwood, 1996; Velma, 1996). For better or for worse, it supported renewed interest among personality and clinical psychologists in what used to be called “hysteria” and is now known as the dissociative and conversion disorders (e.g., Kihlstrom, 1992a, 1994; Kihlstrom, Tataryn, & Hoyt, 1993; Lynn & Rhue, 1994; Michelson & Ray, 1996; Singer, 1990; Spiegel, 1994).

Given its influence outside hypnosis, it may seem somewhat paradoxical that neodissociation theory has been vigorously challenged from within the field of hypnosis itself. In their article, Kirsch and Lynn (1998) concluded that “the evidence supporting . . . dissociation theory is slim . . . and beset with serious conceptual difficulties” (p. 11) and implied that their own social cognitive theory—an approach which is distinct from the “socio-cognitive” approach of Spanos (1991)—does a better job at explaining what happens in hypnosis. On the contrary, there is a great deal of evidence that dissociations of various sorts occur in hypnosis, every reason to think that something like neodissociation theory is needed to account for them, and little reason to think that approaches of the kind preferred by Kirsch and Lynn (1995, 1998, in press; Kirsch, 1991) provide a viable alternative explanation.1

Dissociations in Hypnosis

Hilgard’s (1977/1986) neodissociation theory begins with the assumption that the core of hypnosis is not merely response to suggestions, as Kirsch and Lynn (1995, 1998) seem to assume, but rather a division of awareness such that hypnotized subjects are unaware of percepts and memories of which they should ordinarily be aware. For Hilgard, this division of awareness was exemplified by hypnotic analgesia, in which the subject feels little or no pain in spite of receiving stimulation that would ordinarily be painful. It is also exemplified by posthypnotic amnesia, in which the subject cannot remember the events and experiences that transpired while he or she was hypnotized—events that ordinarily would be quite memorable (Kihlstrom, 1985b, in press-b; Kihlstrom & Evans, 1979).

Both hypnotic analgesia and posthypnotic amnesia occur following suggestions made by a hypnotist, but the essence of hypnosis is not that a subject responds behaviorally to suggestions by claiming to feel no pain or to remember nothing. Rather, it is that the subject’s conscious experience has been altered, so that he or she really feels no pain or remembers nothing. To account for these alterations in conscious experience, neodissociation theory postulates a cognitive state of affairs in which one stream of mental activity (e.g., perception, memory, planning) proceeds outside of phenomenal awareness and apparently out-

---

1 See Kihlstrom (1997) for a critique of other aspects of Kirsch and Lynn’s (1995) point of view on hypnosis.

2 The word subject is used throughout this article, rather than the word participant preferred by the American Psychological Association, because in hypnosis research the word subject is a technical term, referring to that participant in a hypnotic interaction whose role is to respond to suggestions offered by the other participant, known as the hypnotist. Strictly speaking, both the hypnotist and the subject are participants in the social interaction known as hypnosis.
side of voluntary control. This is what *dissociation* means for Hilgard (1977/1986): a division between two streams of consciousness—the image is deliberately borrowed from James (1890/1980)—with only one of these streams accessible to phenomenal awareness. Janet (1889, 1907), Sidis (1902), and Prince (1906, 1914) characterized the stream that is inaccessible to phenomenal awareness as *subconscious*, while James characterized it as *coconscious* (for a review of this terminology, see Kihlstrom, 1984).

### Explicit and Implicit Memory

Dissociations of this sort are ubiquitous in hypnosis. For example, it is easy to demonstrate in appropriately selected highly hypnotizable subjects a dissociation between explicit and implicit expressions of memory in posthypnotic amnesia. *Explicit* memory entails conscious recollection of a past event, such as a list of words studied during hypnosis; *implicit* memory reflects the influence of that event on a task that does not require conscious recollection (Schafer, 1987). For example, neurological patients who have been rendered amnestic by damage to the hippocampus and medial temporal lobe cannot recall words that have been presented to them, but they show an increased likelihood of producing those items when they are asked to complete a stem or fragment with the first word that comes to mind—an effect now known as *priming* (e.g., Warrington & Weiskrantz, 1968). Similarly, beginning with the work of Hull (1933), a wealth of evidence has accumulated that suggests for posthypnotic amnesia impair performance on free- and cued-recall tasks, which require conscious recollection, and spare performance on tasks such as free association, category-instance generation, savings in relearning, and proactive and retroactive interference, which do not (for a comprehensive review of this literature, which was written before the explicit--implicit distinction was formally articulated, see Kihlstrom, 1985b; for a later and briefer account, see Kihlstrom, in press-b). For example, Dorfman and Kihlstrom (1994) found intact semantic priming on a word-association test, despite severely impaired performance on a matched cued-recall test.

Posthypnotic suggestion, in which the subject responds after the termination of hypnosis to a suggestion administered while he or she was hypnotized, may also reflect a dissociation between explicit and implicit memory (Kihlstrom, 1984; for reviews, see Barnier, 1997; and Sheehan & Orne, 1968). This is because, in at least some instances, the subject is unaware (by virtue of posthypnotic amnesia) that the posthypnotic suggestion had been given or that he or she is carrying it out. Thus, the subject's failure to consciously recollect the suggestion may be taken as an impairment of explicit memory, while the subject's acting on the suggestion, nevertheless, outside of awareness indicates that implicit memory has been spared.

Of course, it might also be that the subject remembers the suggestion perfectly well and consciously carries it out when the cue to do so has been given. However, an experiment by Bowers (1975) suggests otherwise. In this study, subjects received a posthypnotic suggestion to choose paintings in a style (portraits or landscapes) different from that which they had preferred on baseline testing. Some of these subjects also received verbal reinforcement for their choices. When the posthypnotic suggestion was canceled and verbal reinforcement discontinued, the subjects who received both the suggestion and the reinforcement reverted to their original preferences, while those who received only the suggestion maintained their new behavior. Bowers concluded, in line with attribution theory, that the subjects in the suggestion-only group interpreted their overt behavior as expressions of genuine internal preferences rather than a response to the hypnotist's suggestion.

### Explicit and Implicit Perception

More recently, the explicit--implicit distinction has been extended beyond the domain of memory to perception (Kihlstrom, Barnhardt, & Tataryn, 1992). The phenomenon of *blindsight*, observed in neurological patients with damage to the striate cortex, provides a dramatic example (Weiskrantz, 1986, 1997). Within hypnosis, salient examples occur in the context of suggested analgesia and were cited by Hilgard (1973, 1977/1986) as part of the motivation for the neodissociation theory of divided consciousness. Thus, it has been known since the time of Sears (1932) and Sutcliffe (1960, 1961) that hypnotic analgesia produces a subjectively compelling reduction in both sensory pain and suffering but does not necessarily alter electrodermal or cardiovascular responses to pain stimuli (for a review, see Hilgard & Hilgard, 1983). This paradox has sometimes been interpreted as indicating that the hypnotic subject merely denies feeling pain (Wagstaff, 1982), but an interpretation in terms of a dissociation between explicit and implicit perception of pain seems more plausible.

Similar sorts of dissociations occur in other facets of perception, as they are affected by hypnotic suggestion. For example, hypnotic suggestions for deafness produce subjectively compelling impairments in auditory acuity but have little effect on speech dysfluencies produced by delayed auditory feedback (Sutcliffe, 1961) or on intrusions in dichotic listening (Spanos, Jones, & Malfara, 1982). Similarly, hypnotic suggestions for blindness result in subjectively compelling reductions in visual acuity but have little negative impact on the use of visual cues in guiding choice behavior (Bryant & McConkey, 1989a) or on priming effects of visual cues in a homophone-spelling task (Bryant & McConkey, 1989b). Hypnotic ablation of the background does not affect perception of the Ponzo illusion (R. J. Miller, Hennessy, & Leibowitz, 1973) or the perception of a slant (Jansen, Blum, & Loomis, 1982). Hypnotic suggestions for tubular blindness have no effect on subjects' automatic compensation for distance cues in size perception (Leibowitz, Lundy, & Guez, 1980) or the illusory feelings of egomotion induced when a stimulus is rotated in the periphery of one's vision (Leibowitz, Post, Rodemer, Wadlington, & Lundy, 1981). Although some theorists might wish to conclude that these subjects see and hear perfectly well, despite their claims to the contrary (e.g., Spanos et al., 1982), an alternative interpretation is that response to such suggestions is characterized by a dissociation between explicit and implicit expressions of perception.

To introduce dissociations between explicit and implicit memory and perception into an article on the neodissociation theory of divided consciousness is not to play on words. Kirsch and Lynn (1998) did make glancing reference to the fact that posthypnotic amnesia impairs recall but spares priming, but they...
puzzlingly concluded that "the hypothesis that suggested amnesia is produced by a division of consciousness . . . is without an evidential base" (p. 108). On the contrary, although the explicit–implicit distinction had not been formally articulated when Hilgard (1977/1986) developed his theory, such dissociations are precisely the kinds of phenomena that neodissociation theory is trying to understand. This evidential base was completely ignored by Kirsch and Lynn.

The Theory of Dissociated Experience

Kirsch and Lynn (1998) followed Bowers (1992; M. E. Miller & Bowers, 1986; Woody & Bowers, 1994) in dividing neodissociation theory into two variants. One, the theory of *dissociated experience*, places primary emphasis on the development of an amnesia-like mental barrier that impairs hypnotic subjects' awareness of themselves and their environment (Hilgard, 1977/1986; Kihlstrom, 1984, 1992b)—thus selectively affecting explicit expressions of perception, memory, and the like but leaving implicit expressions largely intact. The other, the theory of *dissociated control*, places primary emphasis on the apparent automaticity of hypnotic subjects' cognitive and behavioral responses to suggestion (Woody & Bowers, 1994). Unfortunately, their evaluation of the evidence bearing on neodissociation theory often confuses these variants, so that a reader unfamiliar with the area might well take evidence contrary to one view as also contrary to the other. For example, studies of the hidden observer in analgesia, deafness, and amnesia bear on dissociated experience but not dissociated control, whereas evidence of interference by ostensibly dissociated streams of mental activity is critical to dissociated control theory but irrelevant to the theory of dissociated experience.

Moreover, Kirsch and Lynn's (1998) review is highly selective. As noted earlier, they failed to consider dissociations between explicit and implicit memory in posthypnotic amnesia and posthypnotic suggestion and between explicit and implicit perception in hypnotic analgesia, blindness, and deafness. Their limited treatment of posthypnotic amnesia is especially puzzling, given Hilgard's (1977/1986) characterization of dissociation as an amnesia-like barrier. Instead, their coverage of posthypnotic amnesia focuses on tangential issues not critical for either variant, such as whether posthypnotic amnesia occurs spontaneously and the extent to which it can be breached. Similarly, Kirsch and Lynn's analysis of the hidden observer echoes Spanos's earlier critique of this research (e.g., Spanos, 1986; Spanos & Hewitt, 1980) but fails to consider criticisms of Spanos's position (e.g., Laurence, Perry, & Kihlstrom, 1983; for a reply, see Spanos, 1983).

The Hidden Observer

With respect to the hidden observer studies, the primary thrust of Kirsch and Lynn's (1998) argument seems to be that the rate of hidden observer responses varies as a function of the detailed cues present in the experimental situation, indicating that the hidden observer is not "independent of the instructions used to communicate with it" (p. 105). But it is not clear why this fact should bother anyone. After all, hypnosis occurs in the context of a social interaction, in which the subject acts on suggestions offered by the hypnotist (Kihlstrom, 1985a). In interpreting these suggestions, the subject necessarily goes beyond the hypnotist's literal words to consider the totality of cues available in the hypnotic situation (Orne, 1959, 1962). In fact, it is precisely the interplay between internal, cognitive processes and external, social processes that makes hypnosis so fascinating (and, admittedly, somewhat difficult to study). In the absence of suggestions, nothing happens in hypnosis at all. It would be strange indeed if a social interaction were unaffected by the details of the participants' attitudes, motivations, and expectancies and if one's response to suggestion were not affected by his or her interpretation of that verbal communication.

The fact of the matter is that if anything, the hidden observer responses of hypnotized subjects appear to be less responsive to experimental demands than are the parallel responses of subjects who are simulating hypnosis. Consider the evidence from hypnotic analgesia. Confronted with very strong demands, hypnotic subjects produced hidden observers about 79% of the time, compared with 86% for simulators (Spanos, deGroot, Tiller, Weckes, & Bertrand, 1985). Under conditions of moderate demands, the incidence of hidden observers falls to about 50% among hypnotic subjects but remains at about 75% for simulators (Hilgard, Hilgard, Macdonald, Morgan, & Johnson, 1978). In response to weak demands, the rate for hypnotized subjects falls slightly further to 42% (5 of 12 highly hypnotizable subjects) but drops to 0% for simulators (Nogrady, McConkey, Laurence, & Perry, 1983). Given such findings, it is unclear why Kirsch and Lynn (1998) concluded that "although the majority of hypnotized participants do not appear to be faking, their [hidden observer] responses vary in keeping with situational cues in a manner that was indistinguishable from both simulating and alert-imagining participants" (p. 105). In fact, they later conceded that "although hidden observer response rates for both simulating and nonsimulating participants vary as a function of instructions, they may not vary to the same degree" (p. 105). This latter conclusion, which is the more accurate summary of the research results, is the same as that offered by Laurence et al. (1983). It is exactly the pattern of results one would expect when consciousness is divided in the context of suggestive social interaction.

The Interference Bugaboo

Kirsch and Lynn (1998) cited evidence dating back to the time of Hull (1933), indicating that ostensibly dissociated mental processes nevertheless interfere with other ongoing mental activity. It is true that some of the early authorities on dissociation asserted that subconscious processes did not interfere with conscious ones and demonstrations to the contrary (e.g., White & Shevack, 1942) were devastating to this notion. However, Hilgard (1973, 1977/1986) has always treated the issue

---

4 Even here, the review is highly selective: For example, they cited Coe (1996) on the breaching of amnesia but ignored the studies that set Coe's (1978, 1996) program of research in motion (Kihlstrom, 1978; Kihlstrom, Easton, & Shor, 1983; Kihlstrom, Evans, Orne, & Orne, 1980).
of interference as an empirical question—in part because his view was informed by new cognitive theories of attention that emerged in the 1960s and 1970s (e.g., Deutsch & Deutsch, 1973; Kahneman, 1973; Norman, 1968; Norman & Bobrow, 1975). In fact, this difference with the earlier formulation was the primary reason that Hilgard called his a neodissociation theory: Again, neodissociation theory is about awareness, not about interference.

It may well be, as Hilgard (1977/1986) concluded, that keeping a dissociated stream of mental activity out of conscious awareness requires cognitive resources that must be diverted from other ongoing tasks, thus generating interference. Alternately, it may be that the interference is produced because the conscious and subconscious streams of activity draw on the same fund of attentional resources or compete for expression through a single output system (Kihlstrom, 1984, 1992b). In the final analysis, the degree of interference observed between conscious and subconscious processes is likely due to the attentional requirements of these individual processes. Where the dissociated activity demands little attention, as in the posthypnotic nose touching of Bowers and Brenneman (1981), little interference is observed; but where the dissociated activity demands considerable attention, as in the rhythmic key pressing of Knox, Crutchfield, and Hilgard (1975), considerable interference is apparent. This is exactly the pattern that one would expect on the basis of nonhypnotic studies of attention and automaticity (e.g., Kahneman & Triesman, 1984): Nothing about hypnosis changes the way the mind works.

The Domain of Neodissociation Theory

Kirsch and Lynn (1998) faulted neodissociation theory because it provides an incomplete explanation of hypnosis. Admittedly, it does not account well for individual differences in hypnotizability, but what theory does? Wide individual differences in hypnotizability are a fact of life: At best, only about 15% of an unselected sample of college students classify as hypnotic virtuosos (Hilgard, 1965; Register & Kihlstrom, 1986). When it comes to predictors and correlates, hypnosis researchers—regardless of their theoretical orientation—would sell their souls to obtain the “personality coefficients” disparaged by Mischel (1968). Nor need neodissociation theorists bear any responsibility for the excesses of those psychotherapists who think dissociative disorder is pandemic (or a Central Intelligence Agency plot to create Manchurian candidates) and who discover (or create) dozens or hundreds of multiple personalities in their patients. In fact, Hilgard and some other advocates of neodissociation theory have been vigorously critical of these trends (e.g., Bowers, 1991; Bowers & Farvolden, 1996; Kihlstrom et al., 1993).

Like Woody and Bowers (1994), Kirsch and Lynn (1998) are worried that neodissociation theory invokes a relatively rare process, dissociation, to explain a fairly common behavior, response to hypnotic suggestions. But this concern is misplaced because neodissociation theory does not attempt this sort of explanation. Dissociation does not cause response to hypnotic suggestions; rather, it is a quality of mental life that can be observed during such responses. Like the term altered state (Hilgard, 1969), with which neodissociation theory is associated, dissociation is a descriptive construct, not a causal one. Neodissociation theory seeks to understand that particular alteration of consciousness that characterizes response to hypnotic suggestions, especially on the part of highly hypnotizable subjects—that is, subjects who do not perceive and remember things they ought to be able to perceive and remember and who feel actions that ordinarily would require deliberate initiative and effort as involuntary instead. Subjective conviction bordering on delusion and experienced involuntariness bordering on compulsion are, in fact, relatively rare even in hypnosis. But the fact that they are relatively rare does not mean that they are uninteresting. Neodissociation theory assumes that studying such phenomena provides a new perspective on mental architecture.

Finally, Kirsch and Lynn (1998) demanded too much when they faulted neodissociation theory for its inability to explicate the mechanism by which dissociations occur. Do we criticize cognitive theories of memory because we do not have the foggiest idea how memories are encoded in neural tissue? Should the theory of evolution by natural selection have been dismissed because Darwin knew nothing about DNA? Kirsch and Lynn themselves cited as plausible (if incomplete) a dissociative account of posthypnotic amnesia framed in terms of associative network models of memory (Kihlstrom, 1985b): Is that not a mechanistic account? In any case, it should be noted that the sociocognitive theory preferred by Kirsch and Lynn suffers from the same problem. Put bluntly, exactly what is the mechanism by which behavioral responses are automatically triggered by suggestions and expectations?

The mechanisms by which dissociations occur will become clarified as we learn more about how the mind works. But one will never learn how dissociations occur, in hypnosis or elsewhere, if one ignores the available evidence for their occurrence. Neodissociation theory, first and foremost, is intended to draw attention to these interesting phenomena of mental life and to show how the dissociations observed in hypnosis might be related to dissociations that occur elsewhere. It does not pretend to be a comprehensive theory of hypnosis because certain aspects of hypnosis, such as the interpersonal relation between subject and hypnotist, lie outside of its domain. Similarly, the social—cognitive theories preferred by Spanos (1991) and Kirsch and Lynn (1998) draw our attention to the important role played in hypnosis by attitudes, motivations, beliefs, and expectancies—variables that lie outside the domain of neodissociation theory. If one puts them together, one has the beginnings of a comprehensive theory of hypnosis. If one eliminates either one, one dooms oneself to an incomplete understanding of a phenomenon that has fascinated psychologists for more than 100 years.

5 Kirsch and Lynn (1998) complained that neodissociation theory does not explain positive hallucinations. On the contrary, neodissociation theory holds that hallucinations occur when an amnesia-like barrier prevents subjects from being aware of their own role in generating mental images; in other words, hallucinations are images that are experienced as percepts (Kihlstrom, 1992a). A similar account can be given of the experience of involuntariness in the case of direct and challenge ideomotor suggestions (Kihlstrom, 1992a).
References


Received August 10, 1997

Accepted October 9, 1997 •