

Conscious, Subconscious, Unconscious:  
A Neodissociation Perspective

John F. Kihlstrom  
University of Wisconsin

What gives us the impression that we are conscious? What kind of evidence would convince us that a machine such as a computer, or a lower animal such as a dolphin or a chimpanzee, or -- for that matter -- another human being, was conscious? Cognitive scientists of all stripes, especially those who specialize in psychology, philosophy, and artificial intelligence, disagree violently on the answers, and even on whether these are sensible questions. But nobody doubts that we humans, at least, possess consciousness. The facts that erase any doubt about ourselves are the facts of experience. As James put it in the Principles, "the first fact for us, then ... is that thinking of some sort goes on" (p. 224). Introspectively, the experience of consciousness seems to have to do with two things: monitoring ourselves and our environment, such that certain perceptual events and memories come to be accurately represented in phenomenal awareness; and controlling ourselves and our environment, such that we are able to voluntarily initiate and terminate behavioral and cognitive activities.

Cognitive science has been vexed by the problem of consciousness since its prehistory. It has had a checkered past, for example, in psychology: almost the whole of the field for James, but a virtual nonentity with the onslaught of the behaviorist movement. Interest in the topic persisted in the hands of the psychoanalysts, and was revived within mainstream psychology with the cognitive revolution and its emphasis on attention and the span of apprehension. Neurologists commonly encounter disorders of consciousness of various types, and those associated with the "split-brain" syndrome have recently received much notice. Ethologists and behavioral biologists have considered whether lower animals possess the capacity for awareness and voluntary control over their actions -- though this concern within comparative psychology has been supplanted to some degree by a series of similar questions having to do with the capacity for language. Parallel concerns have sometimes caught the fancy of those in the artificial intelligence movement, who must deal with the question of whether computers will ever possess consciousness in the sense of awareness and voluntary control over what they are doing. The problems posed by the experience of consciousness for contemporary cognitive science boil down to questions like these: What is the nature of consciousness? What is it good for? Are there unconscious mental processes, and if so what are they like and what are they good for? Finally, who cares? That is, would cognitive science proceed any differently if its practitioners did not ask questions like these? Let us get some perspective on these questions by turning to some early authorities, before examining some more recent theoretical and empirical developments.

William James devoted the better part of four chapters of the Principles to the topic of consciousness. At the same time, he argued vigorously against the notion of unconscious thought, although he did agree that there were brain processes associated with mental activity of which we might not be aware. As if in warning to Freud and the other psychoanalysts who were to follow, James asserted that the concept of unconscious states of mind "is the sovereign means of believing what one

likes in psychology, and of turning what might become a science into a tumbling-ground for whimsies" (p. 163). But the Freudian psychology which was yet to come shared the force of James' critique with other trends in the psychology of his time, such as those which implicated unconscious inference in perception and judgment. To the contrary, he argued that either the allegedly unconscious thought was rapidly forgotten; or that it represented a revision of an earlier (and conscious) thought; or that it was not a thought at all, but merely an innate or habitual brain process. For James, thought and consciousness were identical. It was as difficult for him to contemplate unconscious thought as it was for Hume to contemplate a round square copula on Berkeley College.

Nevertheless, James did admit that under some circumstances "the total possible consciousness may be split into parts which coexist but mutually ignore each other, and share the objects of knowledge between them" (p. 206). Following Janet and Prince, from whom he drew most of his examples, he referred to this phenomenon as representing "secondary" consciousness, rather than "unconsciousness". In order to understand what James had in mind, it is necessary to consider an important but almost-forgotten school of thought within psychiatry and psychology at the turn of the century.

It is commonly thought that the concept of unconscious mental processes traces its origin to Freud and the theory of psychoanalysis. To the contrary, as Ekkenberger has shown, the idea has a long history before Freud. In 1775, with the appearance of Mesmer on the European medical scene, speculation about the unconscious combined with rationalized, materialistic versions of primitive psychotherapeutic procedures to form what is known as the First Dynamic Psychiatry, whose leader was the French neurologist and psychiatrist J.-M. Charcot. This psychiatry was concerned with demonstrable "functional" as opposed to "organic" mental illnesses -- that is, those pathological syndromes which appeared not to be associated with brain insult, injury, or disease. It attempted to account for a wide range of phenomena, including hysteria, fugue (then called ambulatory automatism), and multiple personality; the "magnetic diseases" of catalepsy, lethargy, and somnambulism (so named because of their resemblance to certain phenomena of animal magnetism, a precursor of hypnosis); spiritistic practices such as automatic writing and crystal-gazing; hypnosis; and suggestibility in the normal waking state. Each of these phenomena, the school held, represented the power of ideas to turn into action (one of the meanings of "dynamic" in the psychological sense); and each seemed to reflect a change in consciousness, as thought and actions occurred outside phenomenal awareness and voluntary control.

The First Dynamic Psychiatry, with its emphasis on unconscious mental contents and processes, invoked one or another of two explicit models of the mind. The point of view known as dipsychism (e.g., Dessoir) held that the mind consisted of two layers, each of which in turn consisted of chains of associations. The "upper consciousness" was active in the normal waking state, while the "lower consciousness" was active in such phenomena as dreams, hysteria, and hypnosis. According to the "closed" version of dipsychism, the lower consciousness contained mental contents which passed into it through the upper consciousness: unattended stimuli, forgotten memories, and various daydreams and fantasies. This point of view contrasts with the less materialistic "open" version, in which the lower

consciousness was held to be in direct communication with other minds. According to polypsychism (e.g., Durand de Groa), each segment of the anatomy was served by its own mental structures, called egos, each of which was capable of perception, memory, and thought. These structures, in turn, were subject to the control of a superordinate structure which was identified with normal consciousness. When the link between subordinate and superordinate egos was broken, certain aspects of cognition and action were carried out subconsciously. Clearly, the concepts of dipsychism and polypsychism are at the root of Freud's first (conscious-preconscious-unconscious) and second (id-ego-superego) models of the mind.

The issues confronted by the First Dynamic Psychiatry were subsequently taken up by another French psychiatrist, Pierre Janet. Following the principle of analysis-then-synthesis familiar in physiology, Janet began by considering the elementary parts of the mental system. Instead of following the lead of the earlier faculty psychology, or the chemical analogies of the structuralists, he argued that the elementary structures of the mind were psychological automatisms: complex acts, tuned to environmental and personal circumstances, preceded by an idea and accompanied by an emotion. Each of these psychological automatisms, by combining cognition, conation, and emotion with action, represented a rudimentary consciousness. According to Janet, all of these elementary automatisms ordinarily were bound together into a single, united stream of consciousness, and operated in awareness and under voluntary control. Under certain circumstances, however, one or more of these automatisms could be split off -- Janet's term was disaggregation -- from the rest, functioning either outside awareness, or voluntary control, or both.

This dissociation view of the unconscious, as distinct from the repression view elaborated by Freud and his followers, was further developed by the American psychologist and psychiatrist Morton Prince. Prince, following the practice of his day as exemplified by James' ten arguments against the existence of unconscious thoughts, reserved the term "unconscious" for the dormant traces of forgotten memories and unattended perceptual inputs, as well as the strictly neurophysiological processes associated with mental activity. Instead, he offered the term coconscious, referring to mental activity which takes place outside phenomenal awareness. Prince preferred this term because it connoted mental activity rather than the lack of mentation (as in the ordinary-language conception of unconsciousness associated with concussion or coma); and because it permitted the division of consciousness into parallel streams without one or more of these being outside awareness. Coconscious mental activities performed outside awareness, together with unconscious mental contents and brain processes, formed the subconscious.

This conceptualization of consciousness was very popular on both sides of the Atlantic, featured prominently in the pages of the then-new Journal of Abnormal and Social Psychology (founded and edited by Prince), and was the chief alternative within dynamic psychiatry to Freudian psychoanalysis. However, it was a conceptualization which was short-lived. The eventual dominance of psychoanalysis in clinical psychology and scientific personology led investigators to be interested in different syndromes and phenomena, a different model of the mind, and the eventual replacement of

dissociation by repression as the hypothetical mechanism for blocking mental contents from consciousness. At the same time, the behaviorist revolution in academic psychology removed consciousness (not to mention the unconscious) from the vocabulary of the science. At fault as well were the dissociation theorists themselves, who often made extravagant claims for the centrality of their phenomenon and whose investigations were often methodologically flawed. The final blow to the concept stemmed from the interpretation that dissociated streams of consciousness, because they were ignorant (Janet's term) of each other, should not influence each other. Numerous demonstrations of mutual interference between ostensibly dissociated tasks showed the contrary, and reference to dissociation gradually disappeared.

In part, the insistence of both early and late dissociation theorists on non-interference between dissociated mental activities seems to stem from a misunderstanding of James' metaphor of the stream of consciousness. Following the metaphor, it is sometimes held that two streams of water, running parallel but separated by tall banks, should not affect each other. However, if the two streams originate from the same source, each will certainly draw some of the flow from the other. Given a model of attention such as Kahneman's, in which a single source of attentional capacity may be deployed in multiple directions, James' metaphor would certainly lead one to predict some degree of mutual interference between simultaneous, thought dissociated, tasks. In fact, the available evidence indicates that simultaneous tasks performed outside of awareness (for example, in hypnosis) do interfere with each other, with the extent of interference a function of the attentional demands of the tasks in question. Where the tasks are easy, there is little or no interference; where one or both are difficult, interference increases proportionately. Awareness and control are the defining feature of dissociation, while noninterference is an open, empirical question.

Viewed in these terms, a number of phenomena -- observed in the laboratory, the clinic, and in the ordinary course of everyday living -- seem to invite a notion such as dissociation. Some of the observations are dramatic, some mundane; the quality of some of the research is impeccable; some demonstrations are marred by poor methodology or contaminated by extraneous social-psychological variables. Some of the results are open to alternative interpretations, and the possibility of performing a definitive experiment seems slim. Some of the claims, in fact, may turn out on close investigation to be false. But not all of them are false. To deny some of them is to deny the facts or our everyday experience. In each of these instances, some aspect of past or present experience cannot be brought into phenomenal awareness, or voluntary control has been lost over thought and action.

Consider, first, the observations of cerebral commissurotomy patients (and intact subjects run under special laboratory conditions), whose right hand literally does not know what the left one is doing: Here is a division in consciousness associated with a literal division in brain structures. Or consider Korsakoff's syndrome, whose dominant feature is an extremely dense anterograde amnesia: recent experiments have revealed, somewhat surprisingly, that these patients can acquire new information, and that this new learning can have an impact on subsequent cognition and

action -- even though the patients have no recollection of the learning experience, and cannot voluntarily retrieve the critical memories. Turning from neurology to psychiatry, there are the very syndromes that caught the attention of the practitioners of the First Dynamic Psychiatry: hysterical anesthetics, paralyses, and amnesias, in which a person complains that he or she cannot remember certain events from the past, perceive stimuli in certain modalities, or voluntarily move certain portions of the body -- all in the absence of any demonstrable organic brain syndrome; fugue states, in which a person loses his or her identity as well as the whole of the autobiographical record, relocates, and takes up a new life under a new name; and multiple personality, where separate personalities, each with its own identity, characteristic features, and personal history, seem to inhabit the same body, separated by amnesic barriers and alternating control over overt action and phenomenal awareness.

In the laboratory, phenomena phenotypically similar to the symptoms of hysteria -- analgesia and other negative hallucinations, spanning all the perceptual modalities; paralyses; compulsive automatism in the form of posthypnotic suggestions; and posthypnotic amnesia for events and experiences transpiring during the state -- can be induced in normal subjects simply by the hypnotist's spoken word -- provided that the subjects are hypnotizable to begin with. Under more familiar conditions, we have numerous experiments on divided attention in which information in the unattended channel influences performance outside awareness; and experiments on multiple simultaneous tasks in which complex activities, executed at an acceptable level of performance, are unrecalled afterwards. Then there are all the experiments on perceptual defense and subliminal perception. In the domain of memory, there are of course the phenomena of state-dependent retention, context-dependent retention, and other manifestations of the encoding specificity principle. There are also compelling demonstrations that unremembered experiences can influence perceptual recognition, and of significant savings in relearning material which appears, even after sensitive testing, to have been completely forgotten.

Examples of disassociation can also be found in abundance outside the clinic and the laboratory. One such experience is familiar to all of us: the dream of REM sleep, in which vivid images are constructed without our intending to do so, and in which complex plots are played out five or more times a night (on average), only to be completely forgotten in the morning. Similarly, there is the pavor nocturnus (night terror) common in children, which scares the daylights out of their parents even though the episodes are never remembered by the children themselves. The sleepwalker carries out complex motor activities while deeply in NREM sleep, and remembers nothing of it in the morning. (Sleeptalking, by the way, which also occurs in NREM sleep, is a doubtful case of disassociation, because the speech does not seem to be intelligent or goal-directed in most cases.) Harkening back to the literature on state-dependent retention, there have been demonstrations that some individuals can respond to hypnotic-like suggestions during (REM) sleep, and continue responding on subsequent nights even though they are amnesic for their actions, and the suggestions, during intervening periods of wakefulness.

Given observations such as these, Hilgard has recently revived the concerns of the First Dynamic Psychiatry by proposing a "neodissociation" theory of divided consciousness. He begins with the assumption that the cognitive apparatus is organized hierarchically, with various subsystems monitoring and controlling thought and action in various domains. Under ordinary circumstances, each subsystem is in communication with each of the others, and with a superordinate central executive structure. It is this central executive which is the source of our subjective feelings of awareness and intentionality. Under certain circumstances, Hilgard holds, a subsystem (or more than one) can lose contact with the central executive. In this case, percepts, memories, and actions represented in one of the subsystems fail to be represented in phenomenal awareness; or perceptual exploration, memorial reconstruction, and overt action occur outside the control of the central executive. Despite this loss of communication with the central executive, the dissociated subsystems can, in principle, continue to interact with each other. This continued interaction is the source of the facilitation and interference effects which formed the basis of the empirical critique of the initial versions of dissociation theory.

It should be clear that the subconscious of neodissociation theory is rather different from the unconscious as it is conceptualized by other schools within psychology. Neodissociation theory differs from psychoanalysis, for example, because the subconscious is not restricted to primitive sexual and aggressive impulses, and those memories and ideas associated with them. Nor do subconscious mental processes operate according to the irrational "primary process" principles associated with the Freudian unconscious (as opposed to the rational, "secondary process" of the ego). Dissociated percepts and memories can be closely tied to objective reality; and dissociated ideas can be rational and even creative. Equally important, rendering something subconscious is not necessarily motivated by defense against anxiety, as is the case with Freudian repression. It can simply happen, as in the case of hysteria, fugue, or multiple personality; or it can be done for entirely adaptive purposes, as in the case of the subjects who voluntarily enter hypnosis or go to a movie precisely so they will become totally absorbed in the action on the screen, forgetting for awhile their everyday concerns (and even who they are).

The subconscious of neodissociation theory also differs in important ways from the manner in which unconscious mental contents and processes are construed, at least implicitly, in classical theories of human information processing. Here four major trends can be discerned: an identification of consciousness with attention, short-term memory, or working memory -- in other words, what we are aware of apprehending at any particular moment; with complex as opposed to simple, or difficult as opposed to routine, information-processing procedures; with the availability of linguistic representations for ideas and experiences; and with declarative, as opposed to procedural, knowledge. But the subconscious of neodissociation theory is not restricted to the procedural knowledge by which we detect features in perceptual stimuli, decode and encode language, retrieve memories, make judgments, perform routine motor tasks, and the like. It can also involve complex factual knowledge, both semantic and episodic in nature, concerning the presence of certain stimuli or the occurrence of certain past events. Nor is it restricted to the simple, automatic, and routine: complex cognitive and behavioral activities apparently can be performed outside

awareness. Linguistic contents can be rendered subconscious, and percepts and memories can be subconscious even though the person's linguistic abilities remain intact. Nor, within the realm of declarative knowledge, is the subconscious simply the repository of unattended perceptual inputs, weak memory traces, and the products of early, simple, and automatic cognitive operations.

Neodissociation theory links a diverse set of real-world and laboratory phenomena under a unified descriptive rubric, and challenges cognitive science to account for them. It comes as no surprise that attention can be divided, though that fact in itself poses problems for those information-processing theories which are predicated on the existence of limited-capacity channels or storage structures. But if attention can be divided with one stream of complex, deliberate, cognitive activity proceeding outside awareness, this seems to cause some problems for the way we usually think about things. The empirical base for the theory is sometimes problematic, but the phenomena of dissociation are trying to tell us something about the nature of conscious, subconscious, and unconscious mental processing. If we do not take these phenomena seriously, and consider their implications for our understanding of the cognitive system, our models of the mind may be led seriously astray. This seems reason enough to continue to pursue neodissociation theory, and to incorporate its insights into larger theories, to produce a comprehensive view of the mind in order and disorder.

#### Acknowledgments

Paper presented at the 4th annual conference of the Cognitive Science Society, Ann Arbor, August 1982. The point of view presented in this essay developed in part from research supported by Grant #MH-35856 from the National Institute of Mental Health, United States Public Health Service. I thank Patricia A. Register and Leanne Wilson for their comments during the preparation of this paper, and Ernest R. Hilgard for promoting the concept of dissociation. An expanded version of this paper, with references, is forthcoming in K.S. Bowers & D. Meichenbaum (Eds.), The unconscious: Several perspectives (Wiley).