Consciousness, the Unconscious, and the Self

John F. Kihlstrom

Department of Psychology, University of California Berkeley

This article examines the various ways in which the self is (or is not) involved in conscious and unconscious mental life. The self may be construed as a cognitive structure representing a person's knowledge of him or herself. This knowledge structure may take the form of a concept, image, or a node in an associative network of memories. Conscious states are not just represented in working memory (e.g., the "global workspace"), but must be linked to a mental representation of the self (as agent or patient, stimulus or experiencer), also represented in working memory. In unconscious mental life, as exemplified by automatic processing or explicit–implicit dissociations, this aspect of self-reference is missing, giving rise to effects that occur outside of phenomenal awareness. At the biophysical level of analysis, the self may be represented by a single "grandmother" neuron, a sparse network of neurons, or it may be widely distributed across the cortex. Viewed phylogenetically, ontogenetically, or culturally, the development of consciousness may be intimately tied to the development of the sense of self.

Keywords: consciousness, the self, automaticity, explicit–implicit distinction, theory of mind

Consciousness and Me-Ness (Reprise)

> Think about it: There is no experience you've had that you were not at the absolute center of. (David Foster Wallace, 2005)

Consciousness has to do with two things: *Monitoring ourselves and our environment*, so that we become aware of the world outside our minds (including our bodies) and our relation to it, and *controlling ourselves and our environment*, so that we voluntarily initiate and terminate various mental and behavioral activities. The monitoring aspect enables percepts, memories, thoughts, feelings, and desires to be represented in phenomenal awareness. It is through the controlling aspect that we exercise free will or agency. Conscious awareness is the prerequisite

for conscious control: Logically, we cannot consciously control things of which we are not consciously aware. It is by means of consciousness that we become aware of events, their meanings and implications, and plan and execute strategies for dealing with them.

All conscious experiences refer somehow to the self as the stimulus or experiencer of some mental state, or the agent or patient of some action (Kihlstrom, 1993a, 1997a). This insight dates back at least to William James who wrote, in his classic introspective analysis of the "Five Characters in Thought," that "Thought tends to Personal Form" (James, 1890/1980a, p. 221; italics and apostrophes original):

> Every thought is part of a personal consciousness.... In this room ... there are a multitude of thoughts, yours and mine, some of which cohere mutually and some not.... Whether anywhere in the room there be a mere thought, which is nobody's thought, we have no means of ascertaining, for we have no experience of its like. The only states of consciousness that we naturally deal with are found in personal consciousnesses, minds, selves, concrete particular I's and you's.... It seems as if the elementary psychic fact were not *thought* or *this thought* or *that thought*, but *my thought*, every thought being *owned*.... On these terms the personal self rather than the thought might be treated as the immediate

datum in psychology. The universal conscious fact is not "feelings and thoughts exist", but "*I* think: and "*I* feel."

Based on his clinical observations of cases of hysteria, Pierre Janet (Janet, 1907, pp. 304–305) had a similar insight:

> The complete consciousness which is expressed by the words, "I see, I feel a movement", is not completely represented by this little elementary phenomenon [i.e., of a sensation of vision or motion]. It contains a new term, the word "I", which designates something very complicated. The question here is of the idea of personality, of my whole person.... There are then in the "I feel", two things in presence of each other: a small, new psychological fact, a little flame lighting up – "feel" – and an enormous mass of thoughts already constituted into a system – "I". These two things mingle, combine; and to say "I feel" is to say that the already enormous personality has seized upon and absorbed that little, new sensation which has just been produced.

And so did Claparede (1911/1951, p. 71), based on one of the earliest observations of implicit memory and source amnesia in the amnesic syndrome (see also Claparede, 1911/1995; Kihlstrom, 1995a):

> If one examines the behavior of such a patient, one finds that everything happens as though the various events of life, however well associated with each other in the mind, were incapable of integration with *the me* [i.e., the self] itself.

The point of these quotations is that consciousness and self are inextricably intertwined. All conscious mental states involve some form of self-reference: *I* see the painting, *I* hear the music; *I* remember where I put my car keys; *I* feel angry at the President, *I* want a hamburger. Without self-reference, there would be just a painting, music, car keys, a president, and a hamburger, but no experience of sensing, perceiving, or remembering; without self-reference, the President and the hamburger would still be there, but the feeling and desire would not. As Searle (1992) has rightly insisted, conscious mental states have a first-person ontology: They exist only insofar as someone experiences them. This is in contrast to other entities, such as molecules and mountains, which have a third-person ontology: They exist regardless of whether there is anyone (or anything) to observe them. For Searle, first-person subjectivity is an irreducible quality of consciousness—which means that reductionist explanations of consciousness cannot succeed because they leave its essence unexplained.

The Self as a Knowledge Structure in Declarative Memory

If consciousness entails self-reference, then the next question is: What is the self? Much ink has been spilled over this question, such as the distinction between self-as-object and self-as-subject (Allport, 1961; Mead, 1934); whether the self-as-subject is an illusion (Kunzendorf, 1988, 2015; Kunzendorf, 2022) whether an individual has *a* core self as opposed to a multiplicity of selves (Markus & Nurius, 1986); and even whether the self can be understood using current scientific methodologies (Klein, 2012). From a cognitive point of view, however, we can simply define the self as one's mental representation of oneself—recording a person's fund of knowledge concerning him- or herself (for comprehensive overviews and relevant references, see Kihlstrom, 1993b, 2012a; Kihlstrom & Cantor, 1984; Kihlstrom & Cunningham, 1991; Kihlstrom et al., 1988, 1997).

As such, the self is not qualitatively different from the mental representations of other entities that we carry around in our heads. Cognitive psychologists (e.g., Anderson, 1995) generally identify two broad forms of mental representation. *Perception-based* representations provide information about the physical appearance of an object or event; these analog representations generally take the form of mental images. *Meaning-based* representations are abstracted from perceptual details and provide information about the meaning of an object or event; they generally take the form of sentence-like propositions.

From such a starting position, we can apply the theoretical apparatus of cognitive psychology to explicate such folk-psychological concepts as *self-concept* and *self-image*. Given what we know about other natural categories (Murphy, 2002; Smith & Medin, 1981), it seems unlikely that the self-concept is structured as a proper set of features that are singly necessary and jointly sufficient to distinguish oneself from all others. However, there are other models of conceptual structure that seem more viable. The self may be structured as a prototype, whose features are only imperfectly correlated with the self; or it may be structured as a collection of exemplars, each relevant to a broad class of situations; or it may be structured as a theory,

which not only lists one's characteristic features but explains (at least to the individual him- or herself) how he or she got that way. In similar fashion, far from being identified solely with self-esteem, the self-image is better construed as a perception-based mental representation of oneself—how one's face and the rest of one's body appears to oneself (Schilder, 1938). Mostly, the self-image is a reflection of what we see in a mirror (Mita et al., 1977), but our mental image of ourselves can be markedly different from what we (or others) actually see (Fallon & Rozin, 1985).

For present purposes, however, it is most convenient to think of the self as a knowledge structure stored in memory. Cognitive psychology commonly distinguishes between *declarative* knowledge, which consists of factual statements about the world represented as sentence-like propositions, and *procedural* knowledge, which consists of a motor and mental skills represented by "if–then" *productions* consisting of a goal, a condition, and an action which will achieve the goal under the condition stated (Anderson, 1995; Winograd, 1972). In theory, productions are executed automatically provided that the relevant goals and conditions are represented in working memory. Declarative knowledge, in turn, comes in two forms: *Episodic* memory for events and experiences associated with a unique spatiotemporal context, and *semantic* memory for abstract, context-free knowledge (Tulving, 1972). In terms of a generic associative-network model of memory such as the various iterations of Anderson's Adaptive Control of Thought theory of cognition (ACT; Anderson, 1976; Anderson & Bower, 1973; Anderson et al., 2004) or Mandler (1980) activation–integration model of retrieval, the self can be thought of as a node in semantic memory representing oneself, linked to other nodes representing one's own characteristic physical and psychosocial features.

It should be clear, though, that the term "self-node" is a kind of euphemism. As with any other element of semantic memory, the node representing the self is linked to nodes representing the various abstract characteristics of the self-physical (male/female, tall/short, fat/thin), demographic (race, ethnic, national origins, socioeconomic status, educational attainment), personality (neurotic, extravert, agreeable, conscientious, open), social (liberal/conservative), and the like, as well as to nodes representing various personal experiences encoded in episodic memory. The "self-node,"

then, exists at the apex of a whole network of nodes, each representing various components of semantic and episodic self-knowledge.

Autobiographical memories are nodes in episodic memory representing specific events and experiences; these are linked to other nodes representing the spatiotemporal context in which the event occurred, and to each other by associations representing temporal and other relations (Kihlstrom, 2009). Each event-node is also linked to the self-node by associations representing the self in one of the semantic roles that a person can take: As the agent or patient of some action, or the stimulus or experiencer of some state (Brown & Fish, 1983; Fillmore, 1968). In a conscious individual, the self-node is activated in working memory, where it can become linked to other activated nodes representing ongoing experiences, thoughts, and actions. Under various circumstances, such as repetition, the links to the self and context are weakened or dissolved, such that what once was a specific episodic memory becomes a more generic semantic one. But before that happens, any representation activated in working memory alongside the mental representation of the self will be represented in phenomenal awareness. Note that it is not activation in working memory that makes some mental state conscious; what is required is the activation of the further link to the self.

James on Unconscious Mental Life

It is common for people to identify consciousness with all mental life. Even William James did so, as when his original definition of psychology as "the science of mental life" (James, 1890/1980a, p. 1) changed to "the description and explanation of states of consciousness as such" (James, 1892/1980b, p. 1). Because James identified consciousness with thought (by which he meant all forms of mental life), the notion of unconscious mental life (as opposed to unconscious brain processes) struck him as a contradiction in terms. Adopting the doctrine of *esse est sentiri* ("to be is to be felt"), he argued that the essence of consciousness (its "to be") is to be sensed (p. 172). Mental states are felt; therefore, they cannot be unconscious. In Chapter 6 of the *Principles*, James went on to consider and rebut ten "proofs" of the existence of unconscious mental states (James, 1890/1980a, pp. 162–176). A major target of James's critique was Hartmann (1868/1931), a Romantic philosopher who had

argued in a best-selling treatise that The Unconscious (the initial capitals were his) pervaded the physical universe—a foreshadowing of modern panpsychist. A minor target was Helmholtz (1878/1968), who—reacting in part to Hartmann's excesses—had begun to hedge on his notion of unconscious inference in perception.

At the same time, however, James cited studies of hysteria and hypnosis by Janet (1889), Binet and Fere (1888), and others that seemed to provide clear evidence of unconscious mental life (e.g., James, 1890/1980a, pp. 202–213). For example, he noted that, in cases of hysterical anesthesia, patients who claimed to have no sense of touch would nonetheless adapt their hands to a grasped object. And in posthypnotic suggestion, subjects executed commands without knowing that they were doing so, or why. James resolved this apparent contradiction by concluding that consciousness could be divided: "*the total possible consciousness may be split into parts which coexist but mutually ignore each other*, and share the objects of knowledge between them" (James, 1890/1980a, p. 206).

Unconscious mental life may have seemed a contradiction in terms for James, but he did not hesitate to write about "secondary or subconscious" (James, 1890/1980a, p. 210) selves. Commenting on Ansel Bourne, a case of fugue which he himself had studied, James noted that, even after recovery, "Mr. Bourne's skull to-day still covers two distinct personal selves" (James, 1890/1980a, p. 392). In discussing fugue and multiple personalities, he wrote of multiple selves, each with its own stream of consciousness and its own fund of autobiographical memories, taking turns monitoring and controlling experience, thought, and action (see also Taylor, 1983, Chapter 4).

Similarly, James's colleague Morton Prince (1905/1978, 1908) used the term *co-conscious* interchangeably with *subconscious* in his studies of Miss Beauchamp and other cases of multiple personalities. Later, Hilgard (1977) revived James's approach in his neodissociation theory of divided consciousness (see also Kihlstrom, 1984, 1992). Hilgard proposed that, under some circumstances (like hypnotic analgesia; Hilgard, 1973), consciousness could be divided so that a subject would be unaware of a painful stimulus, even as that stimulus elicited undiminished physiological responses.

These subconscious streams of consciousness operated independently of each other, in the sense that information processed by one stream was not necessarily processed by the other(s). These "subconscious" or "co-conscious" streams of mental activity are, for all intents and purposes, unconscious. There is no contradiction, however, provided that we do not equate "consciousness" with "mental life" in the first place. Note, however, James's insistence that each stream of consciousness is associated with its own *self*—whether conscious or not.

The Unconscious and the Psychologist's Fallacy

When James wrote that "The distinction . . . *between the unconscious and the conscious being of the mental state* . . . is the sovereign means for believing what one likes in psychology, and of turning what might become a science into a tumbling-ground for whimsies" (James, 1890/1980a, pp. 163–164), he was not rejecting the idea of unconscious (for him, *sub*-conscious) mental life. He was only asking for good standards for demonstrating the existence of unconscious percepts, memories, thoughts, and the like. Otherwise, we risk falling victim to a version of the psychologist's fallacy (James, 1890/1980a, p. 196): Imputing to others unconscious thoughts, memories, feelings, and desires which are not really theirs—but which they cannot deny, precisely because they are held to be unconscious.

Freud, himself made this mistake with his case of Dora (Freud, 1905/1953), whose neurotic symptoms were more likely caused by family pathology than by her unconscious attraction to a family friend (and his wife)—never mind her unconscious incestuous sexual desire for her father (for comprehensive critical coverage of the "Dora" case, see Crews, 2017; Macmillan, 1991/1997). This variant on the psychologist's fallacy became a serious problem during the "memory wars" (Crews, 1995) of the late 20th century, when some psychologists interpreted certain symptoms as unconscious expressions of repressed memories of childhood sexual abuse and other traumata (Kihlstrom, 1995c, 1996b, 1997b, 1998).

The Automatic and the Implicit

Fast-forward 100 years, and psychology has embraced not one but two somewhat different conceptions of unconscious mental life, each associated with operational definitions that can

be used to distinguish between "the conscious and the unconscious being of the mental state" (Kihlstrom, 1987, 1995b).

Automaticity

The revival of scientific interest in unconscious mental life began in the mid-1970s with the introduction of a distinction between *controlled and automatic processing* (LaBerge & Samuels, 1974; Posner & Snyder, 1975; Schneider & Shiffrin, 1977). Conventionally, automatic processes have four canonical features: Inevitable evocation by a specific class of stimuli; incorrigible completion, meaning that they cannot be stopped once initiated; efficient execution, in that they consume little or nothing by way of attentional resources; and parallel processing, meaning that they create little or no interference with other, ongoing cognitive processes (additional features have also been proposed, but appreciable consensus has developed around these four). Automatic processes are unconscious in the strict sense of the term, in that they are executed outside of phenomenal awareness and conscious control (Shiffrin, 1997).

The distinction between automatic and controlled processes stimulated the development of a number of "dual-process" theories in cognitive and social psychology (Chaiken & Trope, 1999; Kahneman, 2011; Sherman et al., 2014; Smith & DeCoster, 2000). According to these theories, many cognitive tasks can be performed either automatically or under cognitive control—although because automatic processes are characteristically fast, unconscious automaticity generally dominates conscious control. Automaticity has become so closely identified with unconscious processing that an anthology of papers exploring various aspects of automatic processing was titled *The New Unconscious* (Hassin et al., 2005).

Even though the concept of automaticity, and its identification with unconscious processing, has gained wide acceptance, there remains a vigorous debate over how automatic processes should be defined (Moors, 2013, 2016). The four canonical features are not perfectly correlated with each other, raising the question of whether any of them are necessary, or any subset of features is sufficient, to define a process as automatic. And while it is convenient to think of the canonical features, and thus automaticity itself, as

qualitative in nature, either present or absent in an all-or-none fashion, it seems more likely that automaticity and control anchor a continuum. At the same time, more recent research has sometimes entailed a degradation in the definition of automaticity—so that, for example, a process is defined as automatic simply because it occurs in the absence of any specific instruction. In the final analysis, it seems best to consider the four canonical features as constituting a prototype of automaticity. The more features such a process displays, the more certain we can be that it operates automatically and unconsciously.

The differential contributions of automatic and controlled processing to task performance can be estimated by means of techniques such as the process-dissociation procedure (PDP), which employs a "Method of Opposition" to pit conscious and unconscious processes against each other (Jacoby, 1991; Yonelinas & Jacoby, 2012). Consider, for example, an experiment in which subjects are presented with a list of words containing the item *ashcan*, followed by a stem-completion test. In an "Inclusion" condition, subjects are instructed to complete each stem, if possible, with a word from the study list—in this case, completing *ash_* with *ashcan*. In an "Exclusion" condition, they are instructed to complete the stems with any word *but* one from the study list—perhaps *ashtray*, *ashamed*, *ashen*, *ashore*, or *ashram*. If the subjects do not consciously recognize *ash_* as related to a studied item, the relatively unlikely response of *ashcan* may nonetheless slip into their stem-completion performance by virtue of automatic priming effects. The PDP, thus, defines automatic, unconscious processing in terms of the lack of both awareness and control. As such, the PDP provides an operational definition of automatic processing to supplement the four prototypical features of inevitable evocation, incorrigible completion, efficient execution, and parallel processing. To the extent to which automatic processing dominates controlled processing, we can say that a process is being executed unconsciously.

Explicit and Implicit Cognition

The revival of interest in unconscious mental life received another boost from the discovery of dissociations between explicit and implicit memory. Specifically, neurological patients with the amnesic syndrome, who could not remember

previously studied pictures or words, nevertheless showed priming effects when asked to identify fragmented versions of those items (Warrington & Weiskrantz, 1968). Conceptually similar effects can also be observed in posthypnotic amnesia (Kihlstrom, 2020a), and in the functional amnesias of the dissociative disorders, such as dissociative identity disorder—formerly known as multiple personality disorder (Eich et al., 1997; Kihlstrom, 2005). They can also be observed in neurologically intact subjects with normal memory functions, such as savings in relearning (Nelson, 1978) and various priming effects (Jacoby & Dallas, 1981). Even if subjects remember seeing some of the primes on the study list, priming is relatively independent of the level of processing and other factors that exert substantial effects on conscious recollection. Following Schacter (1987), we can define *implicit memory* as any effect of a past event on experience, thought, or action, in the absence of, or independent of, the conscious recollection of that event.

Based on the model of implicit memory, the implicit–explicit distinction has been extended to other domains of mental life (Kihlstrom, 1987, 2012b). Thus, *implicit learning* can be defined as the effect of semantic or procedural knowledge on experience, thought, or action, in the absence of conscious awareness of that knowledge (Kihlstrom, 1996a). Reber (1967, 1992), who initially coined the term, showed that subjects could learn to identify letter strings that conformed to an artificial grammar, without being able to specify the grammatical rules themselves. The difference between implicit learning and implicit memory is that the former term applies to semantic and procedural knowledge acquired through experience, while implicit memory applies to episodic knowledge of specific events. In implicit memory, the subject does not consciously remember a particular experience. In implicit learning, subjects may consciously remember the learning experience, but they are not consciously aware of what they have learned.

Similarly, *implicit perception* can be defined as the effect on experience, thought, and action of a stimulus in the present environment, in the absence of conscious perception of that stimulus (Kihlstrom, 1996a; Kihlstrom et al., 1992). Implicit perception includes the processing of "subliminal" stimuli which are presented at too low an intensity or for too short a duration to be consciously perceived (Dixon, 1971, 1981).

But implicit perception is a broader term because it also includes the perception of stimuli that are not technically subliminal but have been masked in such a manner as to prevent conscious perception (Cheesman & Merikle, 1984; Greenwald et al., 1996; Marcel, 1983). It also covers examples of "preconscious" or "preattentive" processing where the stimuli in question are in no sense subliminal, such as parafoveal vision, dichotic listening, inattentional blindness, and what might be called attentional blindness (repetition blindness, the attentional blink, and change blindness). Dissociations between explicit and implicit perception can also be seen in hypnotic analgesia, blindness, and deafness, as well as in neuropsychological syndromes such as blindsight, prosopagnosia, and hemispatial neglect. Implicit perception differs from implicit memory in terms of the interval between the priming stimulus event and the test (stimulus-onset asynchrony, or SOA).

In implicit memory, the SOA is relatively long, and frequently includes some form of distraction; in implicit perception, the interval is typically short, with no distraction, such that both prime and test occur in what James (1890/1980a, p. 609) called "the specious present" (see also White, 2020). In implicit memory, the subject was aware of the priming event at the time it occurred; in implicit perception, the subject was unaware of the prime. For this reason, priming for events occurring during general anesthesia (Kihlstrom & Cork, 2017; Kihlstrom et al., 1990), which by virtue of the SOAs involved might be counted as an instance of implicit memory, better counts as implicit perception instead.

Continuing the analogies, *implicit thought* can be defined as the effect of an internally generated idea, or image, on experience, thought, and action in the absence of conscious awareness of that idea or image (Dorfman et al., 1996; Kihlstrom et al., 1996). Unconscious processing has played a role in theories of thinking and problem-solving at least since Wallas (1926) identified five stages of thought (Sadler-Smith, 2015): Preparation (conscious acquisition of the knowledge required to solve the problem at hand), incubation (the time-dependent emergence of a solution into consciousness), intimation (the intuitive feeling that a solution is at hand), illumination (the breaking through of an insight into consciousness), and verification (consciously checking the solution). The role of unconscious processing in incubation remains a controversial subject

(Sio & Ormerod, 2009), while intuition has come to be identified with automatized, and somewhat unreliable, processes in judgment and decision-making (Kahneman, 2011; Ross, 1977; Tversky & Kahneman, 1974).

Intuition began to come back into psychology's good graces, however, with research by Bowers and his colleagues (Bowers et al., 1990) showing that subjects could identify which of two difficult verbal or pictorial problems was soluble, without also being able to say what the solutions were (see also Dijksterhuis & Strick, 2016; Fleck et al., 2013; Topolinski, 2018). These intuitions seem to be mediated by priming, similar to that encountered in implicit memory, implicit perception, and the "feeling of knowing" (Hart, 1965; Nelson & Narens, 1990). The difference is that what is primed is neither a memory (i.e., a representation of some past event) nor a percept (i.e., a representation of some object in the current stimulus environment). Rather, what is primed is preexisting knowledge, in either verbal or imagistic form.

Explicit and Implicit Motivation

Beyond cognition, the explicit-implicit distinction can be applied in the domains of emotion and motivation. Following the example of implicit memory, we can define implicit motivation as the influence of a motive on experience, thought, and action in the absence of conscious awareness (McClelland et al., 1989). Implicit motivation is commonly measured with the Picture-Story Exercise, a variant on the Thematic Apperception Test (Murray, 1943), in which subjects write stories in response to pictorial stimuli which are then coded for themes and imagery related to such social motives as achievement, affiliation or intimacy, and power (Schultheiss & Brunstein, 2010). As with implicit emotion, the evidence for implicit motivation is somewhat controversial (Kihlstrom, 2019). With few exceptions (e.g., Schultheiss et al., 2009), the means by which explicit and implicit motivation are assessed are so different that any dissociations between them may reflect little more than method variance. Still, given the evidence for various aspects of implicit cognition, the hypothesis that feelings and desires can be unconscious, as well as percepts, memories, and thoughts, cannot be dismissed out of hand.

Explicit and Implicit Emotion

Compared to implicit motivation, the concept of unconscious emotion is much more controversial (Winkielman & Berridge, 2004; Winkielman et al., 2011). While agreeing that emotional states can be generated automatically and unconsciously, even in the absence of conscious perception of the evoking stimulus (Feldman Barrett et al., 2005; Ohman, 1999; Ohman et al., 2002), some theorists have insisted that unconscious emotion is a contradiction in terms—not least because feelings must be felt (e.g., Clore, 1994; Clore et al., 2005; LeDoux, 1994). On the other hand, we already know that percepts, memories, thoughts, and motives can operate outside of phenomenal awareness, and there seem to be no a priori reasons to rule out the possibility of unconscious emotional states as well. Accordingly, implicit emotion can be defined as the influence on experience, thought, or action of a positive or negative affective state, in the absence of conscious awareness of that state (Kihlstrom et al., 2000).

The dissociation between explicit and implicit emotion is anticipated in Lang (1968) multiple-systems theory of emotion, and his tripartite distinction between the subjective, behavioral, and psychophysiological components of fear, and Rachman and Hodgson (1974) concept of desynchrony among them. For example, patients successfully treated for phobia may report feeling little or no fear in the presence of a phobic object, but nevertheless show substantial levels of physiological arousal or facial and other nonverbal behavioral expressions of fear. To some extent, the argument from desynchrony depends on our ability to infer an individual's emotional state from various psychophysiological and behavioral markers—which remains a matter of some controversy (Barrett et al., 2019; Lang, 2014). However, a simple discrepancy between the subjective and the behavioral or physiological components is not definitive evidence for a dissociation between explicit and implicit emotion. Even under ordinary circumstances, the coherence among the three components of emotion is too low to justify such a conclusion (Brown et al., 2020; Mauss & Robinson, 2009). But when coherence is low, we can at least suspect that the subjective feeling state is genuinely unconscious.

While denying that emotional states can be unconscious, LeDoux (1995, 2000; LeDoux &

Brown, 2017) has offered a model of the brain circuits involved in fear which actually allows for the critical desynchrony between the subjective experience of emotion on the one hand, and the physiological and behavioral expressions of emotion on the other. Based on his studies of fear conditioning in rats and other nonhuman animals, but supported by studies of human patients who have suffered damage to the amygdala and surrounding brain tissue, LeDoux proposes that fear stimuli are processed by the amygdala, which in turn generates appropriate behavioral, autonomic, and endocrine responses. Cortical arousal, feedback of somatic and visceral information, and information about the stimulus situation are then integrated into working memory to generate the subjective experience of being afraid. As with the simpler systems described by earlier Papez and MacLean, a communication failure between the amygdala and the cortical substrate of working memory can produce a dissociation between explicit and implicit emotion: The person will respond in a fearful manner without feeling fear or anxiety. In this respect, then implicit emotion can be a symptom of a disconnection syndrome (Geschwind, 1965a, 1965b).

The particular form of emotional desynchrony at issue here may be illustrated by hypnotic analgesia (Hilgard, 1969; Hilgard & Hilgard, 1975), because pain includes an affective component, sometimes known as "suffering." Hypnotic suggestions diminish the subjective experience of pain, both sensory pain and suffering; however, they generally leave the physiological responses to the pain stimulus unchanged. Interestingly, hypnotically suggested emotional numbing dampens both the subjective experience and facial expressions of distress (Bryant & Fearns, 2007; Bryant & Kapur, 2006). With respect to the hypothesis of desynchrony, this pattern of results is what might be expected, if facial expressions are at least in part a reflection of consciously felt emotion.

Another approach to implicit emotion may be the unconscious attitudes ostensibly tapped by the Implicit Association Test (IAT; Greenwald & Banaji, 2017; Greenwald et al., 1998; Greenwald & Lai, 2020). The IAT employs reaction-time measures to reveal associations between, for example, racial identity and judgments of "good" or "bad"; these associations then serve as proxies for subjects' attitudes toward race. The assumption is controversial (Kihlstrom, 2004):

The associations revealed by the IAT may simply reflect subjects' knowledge of social stereotypes, but it is widely shared by social psychologists. Attitudes have an affective component to them, and when IAT measures of attitude are uncorrelated with such traditional measures as self-ratings, it is possible to speak of a dissociation between implicit and explicit attitudes—and, by extension, between explicit and implicit emotion.

The Problem of Self-Reports

In principle at least, automatic, unconscious processing can be identified by applying aspects of the fourfold operational definition of automaticity described earlier. For example, to the extent that a process consumes little or no cognitive capacity, or creates little or no interference with other, ongoing cognitive operations, we can consider it to be automatic and unconscious. Alternatively, Jacoby's (1991) PDP can be used to test whether automatic processing dominates the performance of a particular task.

Unfortunately, dissociations between explicit and implicit cognition, emotion, and motivation offer no such easy identification, because they all ultimately depend on self-report (e.g., Peters et al., 2017). For example, we construe priming as an expression of implicit memory when it occurs in the absence of conscious recall or recognition of the prime, and the only way we know what subjects consciously remembers is by what they tell us. As an aspect of psychology's "conscious shyness" (Flanagan, 1992) psychologists have long distrusted self-reports, and introspection in general, as a window onto the mind (Baumeister et al., 2007; Nisbett & Wilson, 1977). Jacoby's PDP is of some help in this respect because the inability to exclude target items depends on the lack of conscious awareness: The priming effects that constitute the bulk of evidence for implicit cognition, emotion, and motivation are themselves the products of automatic processing. But even in this case, successful performance on the Exclusion task depends on self-report, and thus on the criterion adopted for conscious perception or memory (Fisk & Haase, 2007; Sandberg et al., 2014).

The problem of self-reports was resolved to the satisfaction of some theorists by studying explicit-implicit dissociations in neurological patients with the amnesic syndrome, blindsight, and

the like. But that tack only raised what I think of as *The Paradox of H.M.*: Some psychologists only believe the self-reports of subjects who are brain-damaged. But even brain-damage does not escape the problem of setting the criterion for conscious perception or memory. Perhaps the only way to solve the problem of self-reports is to create conditions in which subjects will feel comfortable reporting accurately on what they perceive, remember, know, think, feel, and desire (Harris, 1988; Kihlstrom, 2002, 2021; Orne, 1962, 1973).

The Automatic, the Implicit, and the Self

What does this have to do with the self? It appears that, in both categories of unconscious mental life (automaticity and explicit–implicit dissociations), processing does not make contact with the mental representation of the self currently activated in working memory. And because conscious awareness requires contact with some mental representation of the self, both automatic processing and expressions of implicit memory (perception, thought, emotion, and motivation) occur outside conscious awareness.

In automatic processing, nodes representing the goals and conditions of actions are activated in working memory, where they can make contact with the activated mental representation of the self. But, at least in Anderson (1992) ACT model, the action itself is executed outside working memory. Thus, skilled drivers of standard-shift automobiles will be aware of where they are (in a car with the shift in first gear) and what they want to do (move the car forward), but they will only be dimly aware, if at all, of exactly what they are doing to accomplish this goal (ease up on the clutch slowly while pressing down gently on the accelerator). Novices consciously follow a set of steps to get the car moving, and they can tell you what they're doing while they're doing it; experts just do it automatically, and they find it hard to describe what they are doing. For a skilled driver, manual shifting is experienced as automatic, not just because it is overlearned, but because the action itself is not represented in working memory. And because it is not represented in working memory, it cannot contact the mental representation of the self which *is* active in working memory. As a result, the person is consciously aware of where s/he is, and what

s/he wants to do, but is not consciously aware of what he or she is doing.

The self plays a somewhat different role in explicit–implicit dissociations. Consider first the example of memory: Subjects study a list containing a word like *ashtray*, and later complete various tests of explicit memory. The encoding process will activate a representation of that word in semantic memory, which we can call the item-node, and link it to nodes representing the self (e.g., "I studied this item") and the spatiotemporal context in which the experience occurred (e.g., "This item was on the second list"). In a standard test of free recall, subjects are simply asked what items they remember from the list they just studied. Such a query will activate nodes representing the self and the spatiotemporal context in which the study phase occurred. If these nodes are linked to the item-node for *ashcan*, then the subject will recall having studied that word.

The process is somewhat reversed for recognition: Presentation of the cue *ashcan* will activate the corresponding item-node in memory; if it is linked to the self and context nodes, the word will be recognized. If for any reason the link to the self-node is absent or degraded, perhaps because the word was poorly encoded to begin with, conscious recollection will fail. Nevertheless, recognition may not fail entirely. If the retrieval process contacts just the activated item-node, successful recognition may be mediated by a priming-based feeling of familiarity (Yonelinas, 2002). If it also contacts the context-node, recognition may be mediated by the abstract knowledge that the item was on the list (Kihlstrom, 2020b). But without the link between the item-node and the self-node, successful recognition will not be accompanied by the conscious recollection of the study episode.

Explicit episodic memory—what Tulving (1985) would call "autonoetic" conscious recollection because it involves conscious awareness of oneself in the past—requires a link between the representation of an event and a representation of the self as the agent, patient, stimulus, or experiencer of that event. On the other hand, the usual tests of implicit memory require no contact with the self-node at all. In stem-completion, the subject may be asked to complete the stem *ash_* with a legal English word. A search of semantic memory will find nodes representing candidate words, including *ashtray* and *ashram*, but the residual

activation attached to *ashcan* will facilitate that word as a response, yielding a repetition priming effect. Similarly for a fragment-completion task, in which the subject is presented with a partial word such as *a_h_a_*; or a perceptual identification test, in which the whole word *ashcan* is presented in a form so degraded as to be virtually unreadable. The subjects do not have to say: "I remember that I saw the word *ashcan*." All they have to say is "That word could be *ashcan*." In a semantic priming test, a subject might be asked to name an object found in the back yard or even an American artistic movement: "The answer is *ashcan*"—again, a response worded in the third rather than first person (and in neither case the response most likely to come to mind first). The subject may have a conscious experience ("That word looks like it could be *ashcan*"), but it is not a conscious recollection of the past. Similarly, on a test of savings in relearning (Nelson, 1978), subjects only have to remember the *second* series of acquisition trials; they do not have to remember anything about the first series. In Claparede's (1911/1951) famous case of source amnesia, the patient did not say "I remember that you pricked my hand"—based on the conscious recollection of a specific episode; instead, she said "Sometimes people hide pins in their hands," based on generic knowledge acquired from past experience (see also Kihlstrom, 1995a).

Succeeding on tests of explicit memory requires that subjects answer in the first person, about their own personal pasts; succeeding on tests of implicit memory does not. The same comparison holds true for other dissociations between explicit and implicit cognition, emotion, and motivation. On a lexical decision test of implicit perception (Marcel, 1983), subjects do not say "I saw the word *doctor*"; instead, they say that "*Nurse* is a legal English word" more quickly than they would in the absence of the prime. In implicit learning (Reber, 1967), they do not say "I know that if the first letter in the string is an *S*, it must be followed by a *T*"; instead, they say "This string looks grammatical, but that one does not." When presented with a Dyads of Triads problem (Bowers et al., 1990), subjects say, "I don't know what the solution is, but *this* problem is soluble while the other one is not." In the particular form of desynchrony that dissociates explicit and implicit emotion (Hodgson & Rachman, 1974), a phobic patient successfully treated with systematic desensitization does not say "I'm afraid of snakes," but when he sees one his heart rate still increases and he only begrudgingly takes his children to the reptile house at the zoo. In implicit motivation, the subject does not report, on a questionnaire, that she is highly motivated to achieve; instead, she tells a story about *someone else* who accomplishes a goal or surmounts some challenge with a combination of skill and grit. In all cases, the answer posed by the explicit test must be answered in the first person, because it requires a report of conscious experience; the response on the implicit test is formulated in the third person because the basis for the response is unconscious. The subject is aware of telling a story, but not aware of projecting her own motives on the protagonist. Whatever implicit memory (etc.), is, it lacks the self-reference that is essential to James's "universal conscious fact" of "*I* see, *I* feel."

References

Allport, G. W. (1961). *Pattern and growth in personality*. Holt, Rinehart, & Winston.

Anderson, J. R. (1976). *Language, memory, and thought*. Erlbaum.

Anderson, J. R. (1992). Automaticity and the ACT* theory. *The American Journal of Psychology*, *105*(2), 165–180. https://doi.org/10.2307/1423026

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). W. H. Freeman.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060. https://doi.org/10.1037/0033-295X .111.4.1036

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. V. H. Winston & Sons.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A., & Pollak, S. (2019). Emotional expressions reconsidered: Challenges to inferring emotion in human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10 .1177/1529100619832930

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. https://doi.org/10.1111/j .1745-6916.2007.00051.x

Binet, A., & Fere, C. (1888). *Animal magnetism*. Appleton-Century.

Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, *22*(1), 72–110. https:// doi.org/10.1016/0010-0285(90)90004-N

Brown, C. L., Van Doren, N., Ford, B. Q., Mauss, I. B., Sze, J. W., & Levenson, R. W. (2020). Coherence between subjective experience and physiology in emotion: Individual differences and implications for well-being. *Emotion*, *20*(5), 818–829. https://doi.org/10.1037/emo0000579

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*(3), 237–273. https://doi.org/10.1016/0010-0277(83)90006-9

Bryant, R. A., & Fearns, S. (2007). Taking the feeling out of emotional memories—A study of hypnotic emotional numbing. *International Journal of Clinical and Experimental Hypnosis*, *55*(4), 426–434. https://doi.org/10.1080/00207140701506508

Bryant, R. A., & Kapur, A. (2006). Hypnotically induced emotional numbing: The roles of hypnosis and hypnotizability. *International Journal of Clinical and Experimental Hypnosis*, *54*(3), 281–291. https://doi.org/10.1080/00207140600689462

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. Guilford Press.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, *36*, 387–395. https://doi.org/10.3758/BF03202793

Claparede, E. (1995). Recognition and selfhood. *Consciousness and Cognition*, *4*(4), 371–378. https://doi.org/10.1006/ccog.1995.1044 (Original work published 1911)

Claparede, E. (1951). Recognition and me-ness. In D. Rapaport (Ed.), *Organization and pathology of thought: Selected sources* (pp. 58–75). Columbia University Press. (Original work published 1911)

Clore, G. L. (1994). Why emotions are never unconscious. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 285–290). Oxford University Press.

Clore, G. L., Storbeck, J., Robinson, M. D., & Centerbar, D. B. (2005). Seven sins in the study of unconscious affect. In L. Feldman Barrett, P. M. Niedenthal, & P. Winkielman (Eds.), *Emotion and consciousness* (pp. 384–408). Guilford Press.

Crews, F. (1995). *The memory wars: Freud's legacy in dispute*. New York Review of Books.

Crews, F. C. (2017). *Freud: The making of an illusion*. Metropolitan.

Dijksterhuis, A., & Strick, M. (2016). A case for thinking without consciousness. *Perspectives on Psychological Science*, *11*(1), 117–132. https://doi.org/10.1177/1745691615615317

Dixon, N. F. (1971). *Subliminal perception: The nature of a controversy*. McGraw-Hill.

Dixon, N. F. (1981). *Preconscious processing*. Wiley.

Dorfman, J., Shames, V. A., & Kihlstrom, J. F. (1996). Intuition, incubation, and insight: Implicit cognition in problem solving. In G. Underwood (Ed.), *Implicit cognition* (pp. 257–296). Oxford University Press.

Eich, E., Macaulay, D., Lowenstein, R. J., & Dihle, P. H. (1997). Memory, amnesia, and dissociative identity disorder. *Psychological Science*, *8*(6), 417–422. https://doi.org/10.1111/j.1467-9280.1997.tb00454.x

Fallon, A. E., & Rozin, P. (1985). Sex differences in perception of desirable body shape. *Journal of Abnormal Psychology*, *94*(1), 102–105. https://doi.org/10.1037/0021-843X.94.1.102

Feldman Barrett, L., Niedenthal, P. M., & Winkielman, P. (Eds.). (2005). *Emotion and consciousness*. Guilford Press.

Fillmore, C. (1968). The case for case. In E. Bach & R. Harms (Eds.), *Universals in linguistic theory*. Holt, Rinehart and Winston.

Fisk, G. D., & Haase, S. J. (2007). Exclusion failure does not demonstrate unconscious perception. *The American Journal of Psychology*, *120*(2), 173–204.

Flanagan, O. (1992). *Consciousness reconsidered*. MIT Press.

Fleck, J. I., Beeman, M., & Kounios, J. (2013). Insight. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology* (pp. 779–794). Oxford University Press.

Freud, S. (1953). *Fragment of an analysis of a case of hysteria* (Vol. 7). Hogarth Press and the Institute of Psycho-Analysis. (Original work published 1905)

Geschwind, N. (1965a). Disconnexion syndromes in animals and man: Part I. *Brain: A Journal of Neurology*, *88*(2), 237–294. https://doi.org/10.1093/brain/88.2.237

Geschwind, N. (1965b). Disconnexion syndromes in animals and man: Part II. *Brain: A Journal of Neurology*, *88*(3), 585–644. https://doi.org/10.1093/brain/88.3.585

Greenwald, A. G., & Banaji, M. R. (2017). The implicit revolution: Reconceiving the relation between conscious and unconscious. *American Psychologist*, *72*(9), 861–871. https://doi.org/10.1037/amp0000238

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, *273*, 1699–1702. https://doi.org/10.1126/science.273.5282.1699

Greenwald, A. G., & Lai, C. K. (2020). Implicit social cognition. *Annual Review of Psychology*, *71*, 419–445. https://doi.org/10.1146/annurev-psych-010419-050837

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Harris, B. (1988). Key words: A history of debriefing in social psychology. In J. G. Morawski (Ed.), *The*

rise of experimentation in American psychology (pp. 188–212). Yale University Press.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, 208–216. https://doi.org/10.1037/h0022263

Hartmann, E. V. (1931). *Philosophy of the unconscious: Speculative results according to the inductive method of physical science*. Routledge and Kegan Paul. (Original work published 1868)

Hassin, R. R., Uleman, J. S., & Bargh, J. A. (Eds.). (2005). *The new unconscious*. Oxford University Press.

Helmholtz, H. V. (1968). The facts of perception. In R. M. Warren & R. P. Warren (Eds.), *Helmhotz on perception: Its physiology and development* (pp. 205–231). Wiley. (Original work published 1878)

Hilgard, E. R. (1969). Pain as a puzzle for psychology and physiology. *American Psychologist*, *24*(2), 103–113. https://doi.org/10.1037/h0027146

Hilgard, E. R. (1973). A neodissociation interpretation of pain reduction in hypnosis. *Psychological Review*, *80*(5), 396–411. https://doi.org/10.1037/h0020073

Hilgard, E. R. (1977). *Divided consciousness: Multiple controls in human thought and action*. Wiley-Interscience.

Hilgard, E. R., & Hilgard, J. R. (1975). *Hypnosis in the relief of pain*. Kaufman.

Hodgson, R. E., & Rachman, S. (1974). Desynchrony in measures of fear. *Behaviour Research and Therapy*, *12*(4), 319–326. https://doi.org/10.1016/0005-7967(74)90006-0

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306–340. https://doi.org/10.1037/0096-3445.110.3.306

James, W. (1980a). *Principles of psychology*. Harvard University Press. (Original work published 1890)

James, W. (1980b). *Psychology: Briefer course*. Harvard University Press. (Original work published 1892)

Janet, P. (1889). *L'automatisme psychologique* [Psychological automatisms]. Alcan.

Janet, P. (1907). *The major symptoms of hysteria*. Macmillan.

Kahneman, D. (2011). *Thinking fast and slow*. Farrar, Straus & Giroux.

Kihlstrom, J. F. (1984). Conscious, subconscious, unconscious: A cognitive perspective. In K. S. Bowers & D. Meichenbaum (Eds.), *The unconscious reconsidered* (pp. 149–211). Wiley.

Kihlstrom, J. F. (1987). The cognitive unconscious. *Science*, *237*(4821), 1445–1452. https://doi.org/10.1126/science.3629249

Kihlstrom, J. F. (1992). Dissociation and dissociations: A comment on consciousness and cognition. *Consciousness and Cognition*, *1*(1), 47–53. https://doi.org/10.1016/1053-8100(92)90044-B

Kihlstrom, J. F. (1993a). The psychological unconscious and the self. *Experimental and theoretical studies of consciousness* (pp. 147–167). Wiley.

Kihlstrom, J. F. (1993b). What does the self look like? In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 5, pp. 79–90). Erlbaum.

Kihlstrom, J. F. (1995a). Memory and consciousness: An appreciation of Claparede and *Recognition et Moiite*. *Consciousness and Cognition*, *4*(4), 379–386. https://doi.org/10.1006/ccog.1995.1045

Kihlstrom, J. F. (1995b). The rediscovery of the unconscious. In H. Morowitz & J. L. Singer (Eds.), *The mind, the brain, and complex adaptive systems* (pp. 123–143). Addison-Wesley Publishing.

Kihlstrom, J. F. (1995c). The trauma-memory argument. *Consciousness and Cognition*, *4*(1), 63–67. https://doi.org/10.1006/ccog.1995.1004

Kihlstrom, J. F. (1996a). Perception without awareness of what is perceived, learning without awareness of what is learned. In M. Velmans (Ed.), *The science of consciousness: Psychological, neuropsychological and clinical reviews* (pp. 23–46). Routledge.

Kihlstrom, J. F. (1996b). The trauma-memory argument and recovered memory therapy. In K. Pezdek & W. P. Banks (Eds.), *The recovered memory/false memory debate* (pp. 297–311). Academic Press.

Kihlstrom, J. F. (1997a). Consciousness and me-ness. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 451–468). Erlbaum.

Kihlstrom, J. F. (1997b). Suffering from reminiscences: Exhumed memory, implicit memory, and the return of the repressed. In M. A. Conway (Ed.), *Recovered memories and false memories* (pp. 100–117). Oxford University Press.

Kihlstrom, J. F. (1998). Exhumed memory. In S. J. Lynn & K. M. McConkey (Eds.), *Truth in memory* (pp. 3–31). Guilford Press.

Kihlstrom, J. F. (2002). Demand characteristics in the laboratory and the clinic: Conversations and collaborations with subjects and patients. *Prevention & Treatment*, *5*(1), Article 36c. https://doi.org/10.1037/1522-3736.5.1.536c

Kihlstrom, J. F. (2004). Implicit methods in social psychology. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *The Sage handbook of methods in social psychology* (pp. 195–212). Sage Publications.

Kihlstrom, J. F. (2005). Dissociative disorders. *Annual Review of Clinical Psychology*, *1*, 227–253. https://doi.org/10.1146/annurev.clinpsy.1.102803.143925

Kihlstrom, J. F. (2009). "So that we might have roses in December": The functions of autobiographical memory. *Applied Cognitive Psychology*, *23*, 1179–1192. https://doi.org/10.1002/acp.1618

Kihlstrom, J. F. (2012a). Searching for self in mind and brain. *Social Cognition*, *30*(4), 367–379. https://doi.org/10.1521/soco.2012.30.4.367

Kihlstrom, J. F. (2012b). Unconscious processes. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology* (pp. 176–186). Oxford University Press.

Kihlstrom, J. F. (2019). The motivational unconscious. *Personality & Social Psychology Compass*, *13*(5), Article e12466. https://doi.org/10.1111/spc3.12466

Kihlstrom, J. F. (2020a). Posthypnotic amnesia: Using hypnosis to induce forgetting. In D. Groome & M. Eysenck (Eds.), *Forgetting: Explaining memory failure*. SAGE Publications.

Kihlstrom, J. F. (2020b). Varieties of recollective experience. *Neuropsychologia, 137*, Article 107295. https://doi.org/10.1016/j.neuropsychologia.2019.10 7295

Kihlstrom, J. F. (2021). Ecological validity and "ecological validity". *Perspectives on Psychological Science*. Advance online publication. https://doi.org/10.1177/1745691620966791

Kihlstrom, J. F., Barnhardt, T. M., & Tataryn, D. J. (1992). Implicit perception. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness: Cognitive, clinical, and social perspectives* (pp. 17–54). Guilford Press.

Kihlstrom, J. F., Cantor, N., Albright, J. S., Chew, B. R., Klein, S. B., & Niedenthal, P. M. (1988). Information processing and the study of the self. In L. Berkowitz (Ed.), Advances in experimental social psychology: Vol. 21. *Social psychological studies of the self: Perspectives and programs* (pp. 145–178). Academic Press.

Kihlstrom, J. F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 1–47). Academic Press.

Kihlstrom, J. F., & Cork, R. C. (2017). Anesthesia. In S. Schneider & M. Velmans (Eds.), *Blackwell companion to consciousness* (2nd ed., pp. 682–694). Blackwell.

Kihlstrom, J. F., & Cunningham, R. L. (1991). Mapping interpersonal space. *Person schemas and maladaptive interpersonal patterns* (pp. 311–336). The University of Chicago Press.

Kihlstrom, J. F., Marchese-Foster, L. A., & Klein, S. B. (1997). Situating the self in interpersonal space. In U. Neisser & D. A. Jopling (Eds.), *The conceptual self in context: Culture, experience, self-understanding* (pp. 154–175). Cambridge University Press.

Kihlstrom, J. F., Mulvaney, S., Tobias, B. A., & Tobis, I. P. (2000). The emotional unconscious. In E. Eich, J. F. Kihlstrom, G. H. Bower, J. P. Forgas, & P. M. Niedenthal (Eds.), *Cognition and emotion* (pp. 30–86). Oxford University Press.

Kihlstrom, J. F., Schacter, D. L., Cork, R. C., Hurt, C. A., & Behr, S. E. (1990). Implicit and explicit memory following surgical anesthesia. *Psychological*

Science*, *1*(5), 303–306. https://doi.org/10.1111/j.1467-9280.1990.tb00222.x

Kihlstrom, J. F., Shames, V. A., & Dorfman, J. (1996). Intimations of memory and thought. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 1–23). Erlbaum.

Klein, S. B. (2012). The self and science: Is it time for a new approach to the study of human experience? *Current Directions in Psychological Science*, *21*(4), 253–257. https://doi.org/10.1177/0963721412447623

Kunzendorf, R. G. (1988). Self-consciousness as the monitoring of cognitive states: A theoretical perspective. *Imagination, Cognition and Personality*, *7*(1), 3–22. https://doi.org/10.2190/TW3U-0886-Q5TA-XHM4

Kunzendorf, R. G. (2015). *On the evolution of conscious sensation, conscious imagination, and coonsciousness of self*. Baywood.

Kunzendorf, R. G. (2022). Source monitoring as an explanation for the illusion of "self as subject." *Psychology of Consciousness*, *9*(1), 64–77. https://doi.org/10.1037/cns0000219

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*(2), 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Lang, P. J. (2014). Emotion's response patterns: The brain and the autonomic nervous syste. *Emotion Review*, *6*(2), 93–99. https://doi.org/10.1177/1754073913512004

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. M. Schlein (Ed.), *Research in psychotherapy* (Vol. 3, pp. 90–103). American Psychological Association.

LeDoux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology*, *46*, 209–235. https://doi.org/10.1146/annurev.ps.46.020195.001233

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

LeDoux, J. E., & Brown, R. (2017). Emotions as higher-order states of consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(10), E2016–E2025. https://doi.org/10.1073/pnas.1619316114

LeDoux, J. E. (1994). Emotional processing, but not emotions, can occur unconsciously. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 291–292), Oxford University Press.

Macmillan, M. (1997). *Freud evaluated: The completed arc*. MIT Press. (Original work published 1991)

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and

word recognition. *Cognitive Psychology*, *15*(2), 197–237. https://doi.org/10.1016/0010-0285(83)90009-9

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954–969. https://doi.org/10.1037/0003-066X.41.9.954

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209–237. https://doi.org/10.1080/02699930802204677

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*(4), 690–702. https://doi.org/10.1037/0033-295X.96.4.690

Mead, G. H. (1934). *Mind, self, and society*. University of Chicago Press.

Mita, T. H., Dermer, M., & Knight, J. (1977). Reversed facial images and the mere-exposure hypothesis. *Journal of Personality and Social Psychology*, *13*(8), 89–111. https://doi.org/10.1037/0022-3514.35.8.597

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology*, *67*, 263–287. https://doi.org/10.1146/annurev-psych-122414-033550

Moors, A. (2013). Automaticity. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology* (pp. 163–175). Oxford University Press.

Murphy, G. L. (2002). *The big book of concepts*. MIT Press.

Murray, H. A. (1943). *Thematic Apperception Test*. Harvard University Press.

Nelson, T. O. (1978). Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(5), 453–468. https://doi.org/10.1037/0278-7393.4.5.453

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.), *The Psychology of learning and motivation* (Vol. 26, pp. 125–173). Academic Press.

Nisbett, R. E., & Wilson, D. S. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–253. https://doi.org/10.1037/0033-295X.84.3.231

Ohman, A. (1999). Distinguishing unconscious from conscious emotional processes: Methodological considerations and theoretical implications. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 321–352). Wiley.

Ohman, A., Flykt, A., & Lundqvist, D. (2002). Unconscious emotion: Evolutionary perspectives, psychophysiological data, and neuropsychological mechanisms. In R. D. Lane & L. Nadel (Eds.), *The cognitive neuroscience of emotion* (pp. 296–327). Oxford University Press.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776–783. https://doi.org/10.1037/h0043424

Orne, M. T. (1973). Communication by the total experimental situation: Why it is important, how it is evaluated, and its significance for the ecological validity of findings. In P. Pliner, L. Krames, & T. Alloway (Eds.), *Communication and affect* (pp. 157–191). Academic Press.

Peters, M. A. K., Kentridge, R. W., Phillips, I., & Block, N. (2017). Does unconscious perception really exist? Continuing the ASSC20 debate. *Neuroscience of Consciousness*, *2017*(1). https://doi.org/10.1093/nc/nix015

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Wiley.

Prince, M. (1908). The unconscious. *Journal of Abnormal Psychology*, *3*(4), 261–297. https://doi.org/10.1037/h0073698

Prince, M. (1978). *The dissociation of a personality*. Oxford University Press. (Original work published 1905)

Rachman, S., & Hodgson, R. E. (1974). Synchrony and desynchrony in measures of fear. *Behaviour Research and Therapy*, *12*(4), 311–318. https://doi.org/10.1016/0005-7967(74)90005-9

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X

Reber, A. S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition*, *1*(2), 93–133. https://doi.org/10.1016/1053-8100(92)90051-B

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.) *Advances in experimental social psychology* (Vol. 10, pp. 173–220). Academic Press.

Sadler-Smith, E. (2015). Wallas' four-stage model of the creative process: More than meets the eye? *Creativity Research Journal*, *27*(4), 342–352. https://doi.org/10.1080/10400419.2015.1087277

Sandberg, K., Del Pin, S. H., Bibby, B. M., & Overgaard, M. (2014). Evidence of weak conscious experiences in the exclusion task. *Frontiers in Psychology*, *5*, 1–8. https://doi.org/10.3389/fpsyg.2014.01080

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 501–518. https://doi.org/10.1037/0278-7393.13.3.501

Schilder, P. (1938). *Image and appearance of the human body*. Kegan, Paul, Trench, & Tribner.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schultheiss, O. C., & Brunstein, J. C. (Eds.). (2010). *Implicit motives*. Oxford University Press.

Schultheiss, O. C., Yankova, D., Diritkov, B., & Schad, D. J. (2009). Are implicit and explicit motive measures statistically independent? A fair and balanced test using the Picture Story Exercise and a cue- and response-matched questionnaire measure. *Journal of Personality Assessment*, *91*(1), 72–81. https://doi.org/10.1080/002238908024

Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.

Sherman, J. W., Gawronski, B., & Trope, Y. (Eds.). (2014). *Dual-process theories of the social mind*. Guilford Press.

Shiffrin, R. M. (1997). Attention, automatism, and consciousness. *Scientific approaches to consciousness* (pp. 49–64). Erlbaum.

Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, *135*(1), 94–120. https://doi.org/10.1037/a0014212

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. MIT Press.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108–131. https://doi.org/10.1207/S15327957PSPR0402_01

Taylor, E. (1983). *William James on exceptional mental states: Reconstruction of the 1896 Lowell lectures*. Scribner's.

Topolinski, S. (2018). The sense of coherence: How intuition guides reasoning and thinking. In L. J. Ball & V. A. Thompson (Eds.), *Routledge international handbook of thinking and reasoning* (pp. 559–574). Taylor & Francis.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12. https://doi.org/10.1037/h0080017

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. Academic Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Wallace, D. F. (2005). *This is water* [Commencement Address at Kenyon College]. https://fs.blog/2012/04/david-foster-wallace-this-is-water/

Wallas, G. (1926). *The art of thought*. Harcourt Brace.

Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature*, *217*, 972–974. https://doi.org/10.1038/217972a0

White, P. A. (2020). The perceived present: What is it, and what is it there for? *Psychonomic Buletin & Review*, *27*(4), 583–601. https://doi.org/10.3758/s13423-020-01726-7

Winkielman, P., Berridge, K., & Sher, S. (2011). Emotion, consciousness, and social behavior. In J. Decety & J. T. Cacioppo (Eds.), *Oxford handbook of social neuroscience* (pp. 195–211). Oxford University Press.

Winkielman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychological Science*, *13*(3), 120–123. https://doi.org/10.1111/j.0963-7214.2004.00288.x

Winograd, T. (1972). *Understanding natural language*. Academic Press.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, *40*, 663–680. https://doi.org/10.3758/s13421-012-0205-5