

UC Berkeley

UC Berkeley Previously Published Works

Title

The Unconscious ☆

Permalink

<https://escholarship.org/uc/item/6vf126r4>

ISBN

9780128093245

Author

Kihlstrom, John F

Publication Date

2018

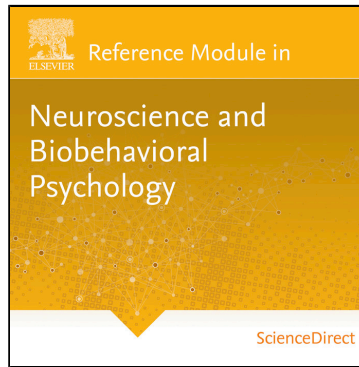
DOI

10.1016/b978-0-12-809324-5.21724-0

Peer reviewed

Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This article was originally published in the online Reference Module in Neuroscience and Biobehavioral Psychology, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<https://www.elsevier.com/about/our-business/policies/copyright/permissions>

From Kihlstrom J.F. The Unconscious. In Reference Module in Neuroscience and Biobehavioral Psychology, Elsevier, 2018. ISBN 9780128093245

ISBN: 9780128093245

© 2018 Elsevier Inc. All rights reserved.

Elsevier

The Unconscious[☆]

John F Kihlstrom, University of California, Berkeley, CA, United States

© 2018 Elsevier Inc. All rights reserved.

Consciousness in Psychology	1
The Initial Discovery of the Unconscious	1
The Unconscious in Pre-psychological Philosophy and Physiology	1
The Psychoanalytic Unconscious	2
The Impact of Behaviorism and the Revival of Consciousness	2
The Rediscovery of the Unconscious	2
Automatic and Controlled Processes	3
The Rise of Neuropsychology	3
The Subliminal	4
The Role of Hypnosis	4
Implicit Cognition	5
The Automaticity Juggernaut	5
Implicit Memory	5
Implicit Perception	5
Implicit Learning	6
Implicit Thought	7
Implicit Motivation and Emotion	7
The Unconscious and the Brain	8
Whither the Unconscious?	9
References	9

Consciousness in Psychology

Consciousness has two aspects: *monitoring* ourselves and the world around us, so that percepts, memories, thoughts, and other mental states are represented in phenomenal awareness; and *controlling* ourselves and the environment, so that we are able to voluntarily initiate and terminate behavioral and cognitive activities. In psychology, *the unconscious* refers to the idea that cognitive, emotional, and motivational states and processes can influence ongoing experience, thought, and action outside of phenomenal awareness and voluntary control.

In the last quarter of the 19th century, as the new science of psychology began to emerge from its roots in philosophy and physiology, consciousness was at the center of the enterprise. Beginning with Wilhelm Wundt and E.B. Titchener, the whole structuralist school of psychology attempted to analyze conscious mental states in terms of their constituent sensations, images, and feelings. Its preferred method of introspection assumed that people had accurate introspective awareness of their own mental states. E.G. Boring summarized the achievements of structuralism with a monograph entitled *The Physical Dimensions of Consciousness* (1927).

Even William James, opposed as he was to the doctrines of structuralism, embraced a version of introspection as his preferred research method (he had a collection of brass instruments, but he hated using them). James began his *Principles of Psychology* with the assertion that "Psychology is the science of mental life" (p. 1). By this he meant *conscious* mental life – as he made abundantly clear in the *Briefer Course*, where he adopted G.T. Ladd's definition of psychology as "the description and explanation of states of consciousness as such".

The Initial Discovery of the Unconscious

At the same time, both James and the structuralists understood that there was more to mental life than was accessible to introspection. The notion that unconscious processes are important elements of mental life is commonly ascribed to Sigmund Freud, the founder of psychoanalysis, but in fact it was an old idea before Freud was even born (Ellenberger, 1970; Klein, 1977; Whyte, 1960).

The Unconscious in Pre-psychological Philosophy and Physiology

At the beginning of the 18th century, long before psychology split from philosophy and physiology, the German philosopher Leibniz had asserted, on logical grounds, that our conscious thoughts must be influenced by sensory stimuli of which we are not aware. At the close of that century, Immanuel Kant's *Anthropology from a Pragmatic Point of View* (1798), the philosopher's

last work and arguably the first comprehensive textbook of psychology, devoted a major section to a discussion "Of the ideas which we have without being conscious of them".

In the 19th century, Johann Friedrich Herbart, drawing on the views of Leibniz, defined the *limen*, or sensory threshold, as a mental battleground where various perceptions, themselves mostly unconscious, competed for representation in consciousness. In the *Treatise on Physiological Optics*, Hermann von Helmholtz argued that our conscious perceptions are determined by unconscious inferences.

The pre-Freudian analysis of unconscious mental life reached its apex with Eduard von Hartmann and his *Philosophy of the Unconscious* (1868), an extremely popular work whose three volumes, running to more than a thousand pages, went through a total of 12 editions. For Hartmann, the universe was ruled by the unconscious, a highly intelligent dynamic force composed of three layers: the absolute unconscious, accounting for the mechanics of the physical universe; the physiological unconscious, underlying the origin, evolution, development, and mechanisms of life; and the relative unconscious, which Hartmann considered to be the origin of conscious mental life. Hartmann's "relative unconscious" is what we would nowadays call *the psychological unconscious* – a term referring to those mental states and processes which influence our experience, thought, and action outside phenomenal awareness and independent of voluntary control.

We owe to Hartmann the Romantic notion, still with us in some quarters today, that the unconscious possesses capacities and powers, which are superior to those available to consciousness. As Hartmann put it: "the Unconscious can really outdo all the performances of conscious reason". In the end, however, Hartmann's ideas proved to be too speculative for the first generation of scientific psychologists. Hermann von Ebbinghaus, discussing Hartmann's book, famously concluded that "what is true is not new, and what is new is not true". William James also offered a warning which would reverberate throughout the 20th century exploration of the unconscious: "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".

The Psychoanalytic Unconscious

Nevertheless, all of this activity, from Leibniz and Kant to Helmholtz and Hartmann, laid the foundation for what Henri Ellenberger, the great historian of psychiatry, called *The Discovery of the Unconscious*. This discovery was consolidated with what Ellenberger called *a new dynamic psychiatry* – the psychiatry of Sigmund Freud and his sometime compatriots, C.G. Jung, and Alfred Adler.

Based on their clinical observations, for example, Freud and his collaborator Josef Breuer concluded that the symptoms of hysteria were caused by unconscious memories of traumatic events. As they put it, "hysterics suffer mainly from reminiscences" – with the proviso that these memories were unconscious, emerging only after skilled intervention by a therapist. In *The Interpretation of Dreams* (1900), Freud proposed a topographical division of the mind into conscious, preconscious, and unconscious compartments. Later, he proposed that our conscious experiences, thoughts, and actions are determined by the interaction between unconscious sexual and aggressive impulses on the one hand, and defense mechanisms such as repression arrayed against them, in order to ward off anxiety caused by the conflict between these primitive motives and sociocultural demands and strictures.

The Impact of Behaviorism and the Revival of Consciousness

Unfortunately, just when the concept of the psychological unconscious was getting up steam, the behaviorist revolution hit. Interest in consciousness disappeared virtually overnight, and interest in the psychological unconscious went with it. For John B. Watson and his comrades in arms, the only way to make psychology truly scientific was to abandon the mental. It was bad enough to try to explain behavior in terms of mental states that couldn't be publicly observed; it was even worse to try to explain behavior in terms of mental states that couldn't be *privately* observed! Of course, psychological interest in consciousness did not die completely with the triumph of behaviorism. So, too, some psychologists maintained an interest in the psychological unconscious. Still, a full-scale revival of academic interest in consciousness had to wait until behaviorism was overthrown by the cognitive revolution.

Somewhat ironically, the early cognitive psychologists hardly ever used the term *consciousness* itself. This is a reflection of two tendencies described by the philosopher Owen Flanagan: on the one hand, there is a *positivistic reserve* reflecting our persisting reluctance to use mentalistic language; and on the other hand, there is a *piecemeal approach* which assumes that big problems like consciousness can be solved by working up from the bottom. But consciousness was there anyway, in the guise of such topics as selective attention, primary or short-term memory, and mental imagery.

The Rediscovery of the Unconscious

A revival of interest in *unconscious* mental life followed shortly thereafter. The seeds for this revival had been planted at the very beginnings of the cognitive revolution, when the linguist Noam Chomsky argued that human language was mediated by "deep" grammatical structures which are inaccessible to conscious introspection, and can be known only by inference. Along the same lines, the philosopher Jerry Fodor argued that many mental functions, such as visual perception, were mediated by dedicated structures which were impenetrable by conscious awareness and voluntary control. Cognitive approaches to perception, as exemplified by the

work of Irvin Rock on visual illusions, entailed a version of Helmholtz's notion of unconscious inference. Finally, the classic multistore model of memory invoked a concept of *preattentive*, or preconscious, information processing.

We are now at a point, however, where interest in the psychological unconscious runs wide and deep within psychology (Bowers and Meichenbaum, 1984; Kelly, 2001; Kihlstrom, 2012; Stein, 1997). This happy state of affairs is the endproduct of at least four quite independent strands of investigation, which together converge on our modern conception of the psychological unconscious: automaticity, cognitive neuropsychology, subliminal perception, and hypnosis.

Automatic and Controlled Processes

One research tradition contributing to the modern interest in the psychological unconscious is the distinction commonly drawn between "automatic" and "strategic" cognitive processes. Skilled reading provides one example of automaticity: we recognize certain patterns of marks on the printed page as letters, and certain patterns of letters as words, and decode the meanings of words in light of the words around them, but we rarely have any conscious awareness of the rules by which we do so. It just happens, as an automatic consequence of having learned to read.

The power of these processes is illustrated by the *color-word effect* discovered by J.R. Stroop. In the basic "Stroop" experiment, subjects are presented with a list of color names printed in different colors, and are asked to name the color in which each word is printed. This task is easy if the ink-color matches the color name (e.g., the word yellow printed in yellow ink); but if the word and its color do not match (e.g., yellow printed in green ink), it is very hard. Despite the subjects' conscious intention to name ink colors, and to ignore the words themselves, they cannot help reading the color names, and this interferes with naming of the colors. It just happens automatically.

According to traditional formulations, automatic processes are inevitably engaged by the appearance of specific environmental stimuli, regardless of the person's conscious intentions. Once invoked, they proceed inevitably to their conclusion, and (in theory at least) their execution consumes no attentional resources. Because they consume no attentional resources, automatic processes leave no traces in conscious memory. Some automatic processes are innate, or nearly so, while others are automatized only after extensive practice with a task. More recently, some of these properties have been called into question by revisionist, memory-based views of automaticity. It's no longer clear that ostensibly automatic processes are *really* executed involuntarily and *really* consume no attentional resources. But even these revisionist views agree that some mental processes, represented in procedural memory, are unconscious in the strict sense of the term: they are inaccessible to phenomenal awareness under any circumstances, and can be known only by inference from task performance.

The Rise of Neuropsychology

Both traditional and revisionist approaches to automaticity assume, at least tacitly, that the mental contents upon which these processes operate are accessible to conscious awareness. However, it is now clear that our experiences, thoughts, and actions can be influenced by mental contents – percepts, memories, thoughts, feelings, and desires, of which we are unaware. Compelling evidence for this proposition began to accumulate about 30 years ago, as cognitive psychology turned into cognitive neuropsychology, and researchers began to see evidence of the psychological unconscious in the behavior of brain-damaged patients.

Pride of place in this history goes to studies of the *amnesic syndrome* resulting from bilateral damage to the hippocampus and related structures in the medial temporal lobe, or, alternatively, to the diencephalon and mammillary bodies. On clinical observation, such patients show a dense anterograde amnesia: after only a few moments of distraction, they cannot consciously remember events that have occurred just recently. But as the study of these patients shifted from clinical description to controlled laboratory investigation, it became apparent that the events apparently covered by the amnesia nonetheless influenced the patients' ongoing experience, thought, and action.

For example, Elizabeth Warrington and Lawrence Weiskrantz showed that amnesic patients, who could not remember having studied a list of words, were nonetheless biased to complete ambiguous word stems or fragments with items from the previously studied list. Past experience influenced their subsequent task performance, even though they had no conscious recollection of the experience itself. Based on "priming" effects such as these, Daniel Schacter (1987) and others drew a distinction between two expressions of memory, *explicit* and *implicit*. Explicit memory refers to one's conscious recollection of the past, as manifested on tasks like recall and recognition. Implicit memory, by contrast, refers to any change in experience, thought, or action that is attributable to a past event, in the absence of conscious recollection of that event.

Priming effects, in which prior exposure to a word like *assassin* makes it easier to complete a fragment like *a__a__i__* than one like *t__p__r__*, are good examples of implicit memory, because the priming effect obviously depends on memory, but the task does not, logically, require conscious recollection of any past event. All the subject has to do is to generate an acceptable word that fits in the spaces provided. The sparing of implicit memory in amnesia shows that some representation of a prior event has been encoded and stored in memory, and influences ongoing experience, thought and action, even though that event cannot be consciously remembered. Implicit memories are unconscious memories.

Neuropsychological research has also revealed unconscious influences in the perceptual domain. Perhaps the most dramatic example is the phenomenon of *blindsight* documented in some patients with damage to the striate cortex of the occipital lobe (Weiskrantz, 1986, 1997). Such patients experience a scotoma – a portion of the visual field where they have no visual experience.

When a stimulus is presented their scotoma, they see nothing at all. Yet when encouraged to make guesses about the properties of the stimulus, their conjectures about presence, location, form, movement, velocity, orientation, and size prove to be more accurate than would be expected by chance alone.

Something similar occurs in cases of visual neglect arising from lesions in the temporoparietal region of one hemisphere (usually the right) that do not affect primary sensory or motor cortices. These patients appear to neglect the corresponding portion of the contralateral sensory field (usually the left). Thus, a patient asked to bisect a set of horizontal lines may ignore the ones on the left side of the page; and for the remainder, the pencil strokes tend to be located about one-quarter of the way in from the right. It is as if the left half of the page, and the left half of each line, is not seen at all. But in at least some cases, it can be shown that these patients respond to information available only the neglected field. For example, pictures presented in the neglected portion of the visual field prime lexical decisions concerning semantically related words presented in the intact portion. Preserved visual functioning in blindsight and in neglect is unconscious perception.

The Subliminal

Perception without awareness can also be observed in neurologically intact subjects, in the form of "subliminal" perception. Before World War II, the question of subliminal perception was raised mostly under the influence of psychoanalysis; after the war, the notion was revived as part of Jerome Bruner's "New Look" in perception – only to be shot down by the withering critiques of Israel Goldiamond and C.W. Eriksen. That might have been the last we heard of the phenomenon, except that in the early 1980s Anthony Marcel presented solid evidence of subliminal semantic priming effects on lexical decision: presentation of a word like *doctor* primed lexical decisions of a semantically related word like *nurse*, even though an intervening mask prevented subjects from consciously perceiving the prime itself.

Other investigators soon confirmed and extended Marcel's findings, but his studies raised a firestorm anyway, with a number of critics essentially repeating the criticisms that Goldiamond and Eriksen had made of the earlier work. A major reason for this response may have been the association of subliminal perception with psychoanalytic theory. But another reason was the simple fact that the cognitive theories of the time tended to describe cognition in terms of a series of ever more complicated processes, and had no room in them for the possibility that the meanings of words could be analyzed unless conscious attention was paid to them. The criticisms had the character of the apocryphal entomologist who found a bug he couldn't classify, so he stepped on it.

In any event, things are vastly different now (Bornstein and Pittman, 1992). Signal detection theory has replaced the notion of an absolute *limen*, or threshold, with the statistical concept of zero sensitivity (where $d' = 0$). A wealth of evidence now supports the validity of subliminal perception, defined broadly as the influence of stimuli that are too degraded by their conditions of presentation to be accessible to conscious perception. The debate now is not so much over whether subliminal perception (so defined) occurs, as over the extent of subliminal processing.

The Role of Hypnosis

The fourth line of research contributing to the rediscovery of the unconscious was hypnosis, a social interaction in which the subject acts on suggestions for experiences involving alterations in perception, memory, and the voluntary control of action. As William James recognized more than a century ago, many of these phenomena involve a division in consciousness such that memories, percepts, and the like influence experience, thought, and action outside of phenomenal awareness.

Consider, for example, posthypnotic amnesia – the phenomenon which gave hypnosis its very name. After receiving appropriate suggestions, many highly hypnotizable subjects come out of hypnosis unable to remember the events and experiences, which transpired while they were hypnotized. For example, subjects who have studied a list of animal names while hypnotized will be unable to remember them afterward. However, these unremembered items will also give rise to priming effects: if asked to generate names of animals, subjects will be more likely to give the names they studied, compared to instances that they did not encounter in hypnosis. Moreover, after the amnesia suggestion has been canceled by a prearranged *reversibility cue*, the subject will regain perfect conscious memory for the studied list. The reversibility of amnesia indicates that, in contrast to the organic amnesias associated with hippocampal damage, posthypnotic amnesia reflects a deficit in retrieval rather than encoding. But the preserved priming effects show that the retrieval deficit affects explicit, but spares implicit, expressions of memory.

Hypnotic suggestion can also affect perceptual functions. When given hypnotic suggestions for blindness, many hypnotizable subjects have the subjectively compelling experience that they no longer can see. However, Richard Bryant and Kevin McConkey presented subjects with cards on which were printed a homophone (i.e., words which have the same sound but two different spellings, such as *pain* and *pane*) together with a disambiguating word (e.g., *body* or *window*). The hypnotically blind subjects did not see the cards, but on a subsequent test they spelled the homophones in accordance with the disambiguating associates with which the words had been paired. This is another kind of priming effect, and it indicates clearly that the words in question were perceptually processed outside of awareness.

Yet a third example of unconscious processing in hypnosis is provided by posthypnotic suggestion – the phenomenon that, so our mythology tells us, gave Freud his first insight into the psychological unconscious. In some sense, posthypnotic suggestion is a special case of implicit memory, because subjects act on a suggestion given during hypnosis, despite the fact that they cannot remember the suggestion itself. Posthypnotic behavior is typically experienced as an involuntary, quasi-automatic response to the cue, but it has had no opportunity to be automatized by extensive practice, and detailed experimental analysis shows that it

does not fit the conventional criteria of automaticity. Despite consuming considerable attentional resources, posthypnotic suggestions are executed automatically – thus extending the boundaries of the psychological unconscious beyond the automatic.

Implicit Cognition

The rediscovery of the unconscious may have gotten a late start, but it is now well along.

The Automaticity Juggernaut

The concept of automaticity has proved to be particularly popular across a wide variety of fields within psychology (Baragh, 1997; Hassin et al., 2005). For example, a set of articles in the July 1999 *American Psychologist*, published under the title of “Behavior – It’s Involuntary” illustrated the increasingly powerful role played by the concept of automaticity in personality, social, and clinical psychology. The general argument is that some attitudes, impressions, and other social judgments, as well as aggression, compliance, prejudice, and other social behaviors, are typically mediated by automatic processes, which operate outside phenomenal awareness and voluntary control. To some extent, what might be called an *automaticity juggernaut* seems to represent a reaction to a cognitive view of social interaction, which seems, to some theorists at least, to inappropriately emphasize conscious, rational, cognitive processes, at the expense of the unconscious, irrational, emotive, and conative.

In many respects, the popularity of automaticity seems to represent a reversion to earlier, pre-cognitive, situationist views within social psychology. After all, the concept of automaticity is at least tacitly modeled on innate stimulus-response connections such as reflexes, taxes, and instincts, as well as those acquired through classical and instrumental conditioning. The automaticity juggernaut is not exactly a reversion to Skinnerian behaviorism, because it entails internal mental representations and processes intervening between stimulus and response, but it is close: if the cognitive processes underlying social cognition and social behavior are indeed largely automatic, then not *too much* thought has gone into them.

Implicit Memory

Like automaticity, the concept of implicit memory has received a huge amount of attention in the field. A whole industry has developed around implicit memory, involving amnesic and demented neurological patients; dissociative disorders such as psychogenic amnesia, fugue, and multiple personality; children and the healthy aged; depressed patients receiving electroconvulsive therapy; surgical patients receiving general anesthesia or conscious sedation; hypnotized subjects (but apparently not sleepers); and even college students who have all their wits about them.

Still, there are some important issues that have to be addressed by further research. For example, almost all the evidence on implicit memory has been collected within a single narrow paradigm, repetition priming, leading to theories of implicit memory which emphasize relatively low-level perceptual processes. But semantic priming occurs too, not just in posthypnotic amnesia but in organic amnesia as well, suggesting that these perception-based theories are not adequate to the phenomenon. Similarly, while explicit and implicit memory are dissociable, they also interact, requiring revisions in theories which hold that these two expressions of memory are mediated by separate memory systems in the brain.

Despite these and other persisting questions, the general acceptance of the distinction between explicit, conscious and implicit, unconscious expressions of memory opens the door to extensions of the explicit-implicit distinction to other domains of mental life.

Implicit Perception

For example, by analogy to implicit memory, *implicit perception* may be defined as the influence of a *current* event, or an event in the very recent past (what William James called the *specious present*), on experience, thought, or action, in the absence of conscious perception of that event. Implicit perception subsumes so-called “subliminal” perception, involving the processing of stimuli which are degraded beyond conscious perception by low intensities, brief durations, or masking stimuli. But it goes beyond the subliminal to include neurological syndromes such as blindsight and neglect, where the stimuli are in no sense subliminal: they are perfectly visible to everyone but the brain-damaged subject. So too, in the conversion syndromes of “hysterical” deafness, blindness, and anesthesia. Similarly, in hypnotic blindness, deafness, anesthesia, and analgesia, the subjects would be clearly aware of the stimuli in question were it not for the hypnotist’s suggestion. On the fringes of consciousness, are cases of so-called *preattentive* processing, where the stimulus in question is nominally supraliminal, but escapes focal awareness by virtue of parafoveal presentation, or presentation over the unattended channel in the dichotic listening or shadowing paradigm. Priming in the absence of conscious perception has also been observed in inattention blindness and various forms of attentional blindness, such as repetition blindness, the attentional blink, and change blindness. Thus, the term *implicit perception* captures a broader range of phenomena than is covered by the term *subliminal perception*, because it covers the processing, outside of conscious awareness, of stimulus events which are clearly perceptible in terms of intensity, duration, and other characteristics. It also has the extra advantage of skirting the difficult psychophysical concept of the *limen*.

What all of these phenomena have in common is a dissociation between explicit and implicit perception, analogous to the dissociation between explicit and implicit memory: the subject’s experience, thought, and action is affected by some event in the

current stimulus environment, in the absence of conscious perception. The distinction between implicit perception and implicit memory is not always easy to make, because both phenomena are revealed by post-exposure priming effects – that is, by performance on a nominal test of implicit memory. Arguably, however, the term *implicit memory* should be reserved for cases where the stimulus event was consciously perceived at the time of encoding; where there is no conscious awareness at the time of encoding, then we can consider priming effects as evidence of implicit perception. Thus, on the assumption that adequately anesthetized patients are really unaware, during their procedures, of what our medical colleagues disarmingly call “surgical stimuli”, evidence collected in the recovery room of priming effects attributable to events presented during surgery constitutes evidence of implicit perception, not just implicit memory. By contrast, priming effects observed in conscious sedation, where subjects are fully aware of the study trials, is a pure implicit memory effect.

Adopting the implicit-explicit distinction may help resolve a persisting controversy over the scope of unconscious perception. For example, Anthony Greenwald and his colleagues have asserted that subliminal perception is analytically limited – some semantic processing of a subliminal stimulus is possible, but not too much of it. On the other hand, advocates of *subliminal symbiotic stimulation* – the “Mommy and I Are One” experiments – assume more processing than Greenwald’s arguments would predict; so do those who employ subliminal techniques in advertising and psychotherapy. Most arguments for the analytic power of subliminal processing appear to be based on Romantic or psychoanalytic notions of the unconscious. On the other hand, there are many ways to render percepts implicit, and how it is done may make a big difference to what you can do with them. Philip Merikle and his colleagues have distinguished between the *objective threshold*, where all response to the stimulus drops to chance levels, and the *subjective threshold*, where the subject simply doesn’t experience the stimulus consciously. Greenwald’s experiments take subjects as close to the objective threshold as they can get, and with such degraded presentations it is not surprising that perception, and memory encoding, is analytically limited. When subjects get closer to the subjective threshold, more extensive analyses, still outside awareness, might be possible. In hypnosis, where the stimulus is in no sense degraded, the possibilities for analysis might be unlimited.

Implicit Learning

Continuing the elaboration of the explicit-implicit distinction to other domains, we can define *implicit learning* as the acquisition of new knowledge and patterns of behavior through experience, in the absence of awareness of the knowledge or behavior so acquired. As it happens, the term “implicit learning” antedates implicit memory, having been coined by Arthur Reber in 1967. Reber demonstrated that subjects who studied a set of letter strings generated by a Markov-process artificial grammar could distinguish between new grammatical and ungrammatical strings without being able to articulate the grammar in question (Reber, 1993).

In some respects, the learning of artificial grammars appears similar to the acquisition of syntax in natural language. After all, we are perfectly fluent speakers and interpreters of our native language long before we learn the rules of grammar in elementary school. But implicit learning has also been observed in a wide variety of other paradigms, including classical and instrumental conditioning, control of complex systems, and the learning of categories and sequential relationships. In each of these cases, the claim is that people’s behavior is shaped by prior experience – the classical definition of learning – even though they are unaware of what they have learned.

As with implicit perception, however, the border between implicit learning and implicit memory is a little vague. Of course, this is as it should be: memory provides the cognitive basis for learning in the first place, and whatever is learned has to be stored in memory. But the problem goes deeper than that. When normal subjects learn an artificial grammar, they certainly remember being asked to study the sample strings, and they may even remember the strings themselves, even if they are unaware of what they have evidently learned about the structure of the grammar itself. By contrast, when brain-damaged amnesic patients acquire new patterns of behavior from experience, as exemplified by mirror tracing and pursuit-rotor learning in Patient H.M. they are amnesic for the whole learning experience. In amnesia, the occurrence of implicit learning also gives evidence of preserved implicit memory, but as in the case of implicit perception the term implicit memory is best reserved for effects which occur in the absence of conscious memory for the original experience. By the same token, implicit learning refers to abilities and patterns of response that are acquired through learning experiences, in the absence of conscious awareness of what has been learned.

At this point, a small industry has developed around implicit learning – culminating in the publication in 1998 of an entire handbook devoted to the topic. Still, it has to be said that the claim of implicit learning remains controversial even after more than 30 years of work. There is a continuing debate over whether implicit learning is really *unconscious* in any meaningful sense of the term. It just may be too much to expect subjects to be able to articulate an entire Markov-process artificial grammar, but subjects might be consciously aware of just enough of the rule to permit them to discriminate at above-chance levels between grammatical and ungrammatical strings. Perhaps this debate will be resolved if we pay a little more attention to the format in which the newly acquired knowledge is represented. Perhaps, as implied by the original artificial grammar studies, the subject acquires a whole system of *If-Then* productions and this procedural knowledge, like all procedural knowledge, is inaccessible to conscious introspection.

On the other hand, perhaps the knowledge acquired during implicit learning is not procedural at all, but declarative in nature. For example, subjects might abstract from the learning trials a prototype of a grammatical string; alternatively, they may simply memorize the instances on the study list. In either case, they may make relatively accurate grammaticality judgments by consciously comparing test items to the summary prototype, or to the specific exemplars they’ve memorized. In any event, amnesic patients can learn from their experience without remembering the learning experience itself, and in that sense, at least, implicit learning gives evidence of unconscious influence.

Implicit Thought

If the concept of implicit learning is more controversial than those of implicit memory or implicit perception, the concept of *implicit thought* is even more so. Still, the literature contains some favorable evidence, if only just a little. For example, some years ago Kenneth Bowers and his associates found that subjects can choose which of two problems is soluble without knowing the answer to the soluble one. In a variant on Sarnoff Mednick's Remote Associates Test (RAT), called "Dyads of Triads", subjects were presented with two RAT-like items: the task in the traditional RAT is to generate a fourth word which is associatively linked to the other three. In Bowers's modification, one triad coheres on a common (if remote) associate, and the other does not (at least barring psychotically loose associations). Bowers et al. found that subjects could choose the coherent triad at better than chance levels, even when they could not say what the solution was.

Subsequent research has found that soluble RAT items primed their solution words for lexical decision, even when the subject failed to produce the solution itself. This priming effect, and the effect on choice observed by Bowers, seem to reflect the preconscious activation of a mental representation of the solution to the problem. But this mental representation is not a percept, and it is not a memory. The mental representation in question is an idea or an image – or, more broadly, a *thought*.

Accordingly, *implicit thought* may be defined as the influence of some cognitive representation, itself neither a percept nor an episodic memory, on experience, thought or action, in the absence of conscious awareness of that representation.

Implicit thought may well underlie some of the most interesting facets of creative thought. In this view, intuition reflects a priming-based "feeling of knowing" similar to what we commonly see in studies of memory; incubation reflects the gradual accumulation of strength of this primed idea; and insight reflects the emergence of the preconscious idea into the full daylight of consciousness (Reder, 1996).

Implicit Motivation and Emotion

Along with automaticity, implicit memory, implicit perception, implicit learning, and implicit thought comprise the cognitive unconscious (Underwood, 1996). But cognition is not all there is to mental life, and so we are led to ask whether there is an *affective* unconscious, and a *conative* unconscious, as well. Of course, emotional and motivational states may arise automatically, and in that sense result from unconscious processes. But can motives and emotions themselves be unconscious, in the same way that implicit memories can?

In fact, the late David McClelland and his associates have articulated a concept of implicit motives – interestingly, without overt reference to the concept of implicit memory. Explicit motivation might be defined as the conscious representation of a conative state, or the desire to engage in some particular activity, as represented by a craving for food, yearning for love, and the like. By contrast, implicit motivation would refer to any change in experience, thought, or action which is attributable to one's motivational state, in the absence of conscious awareness of that state. We are admittedly verging near Freudian territory here, but the motives in question are not seething sexual and aggressive impulses arising from the id; they are motives for achievement, power, affiliation, and intimacy (Schultheiss and Brunstein, 2010).

Moreover, McClelland and his colleagues keep us on a firm empirical base. For them, explicit motives are self-attributed: the person is aware of the motive, can reflect on it, and can report its presence in interviews or on personality questionnaires. Implicit motives, by contrast, are inferred from the person's performance on such tasks as the Thematic Apperception Test. On the basis of an extensive program of empirical research, McClelland and his associates concluded that explicit and implicit motives influence different classes of behavior, and respond to different types of social influence. More work needs to be done, but the evidence so far indicates that explicit and implicit motives are indeed dissociable, in much the same way that explicit and implicit memories are.

Turning to the affective domain, it is possible that dissociations occur between explicit and implicit expressions of emotion, just as they occur between explicit and implicit expressions of memory. Again paralleling the vocabulary of the cognitive unconscious, we define *explicit* emotion as the person's conscious awareness of an emotion, feeling, or mood state; *implicit emotion*, then, refers to changes in experience, thought, or action which are attributable to one's emotional state, in the absence of conscious awareness of that state.

The first thing that should be noted is that conscious emotional responses can serve as expressions of implicit memory and perception (Feldman Barrett et al., 2005). Most of the relevant studies make use of the mere exposure effect documented by Robert Zajonc, in which exposure to an object increases one's preference for that object on a subsequent choice task. According to one theory, the mere exposure effect is a variant on the priming effect familiar from studies of implicit memory and implicit perception. On the memory side, it seems that brain-damaged, amnesic patients show the mere exposure effect, even though they cannot recognize the objects to which they had previously been exposed. With respect to perception, normal subjects show mere exposure effects even when the exposures in question were subliminal, and thus not consciously perceived. In both cases, the subjects preferred previously exposed objects to new ones, even though they had no conscious perception, or conscious recollection, of the visual stimuli themselves.

With respect to the proposition that people can be unaware of emotional states, which nonetheless influence their ongoing experience, thought, and action, the empirical evidence is regrettably somewhat sparse. However, there remain good theoretical reasons for thinking that it might be true. For example, Peter Lang's multiple-systems theory of emotion postulates that every emotional response consists of three components: verbal-cognitive, corresponding to a subjective feeling state such as fear; overt motor, corresponding to a behavioral response such as escape or avoidance; and covert physiological, corresponding to a change in some autonomic index such as skin conductance or heart rate. These three components or systems usually covary, but under some circumstances they can move in different directions – a state which Stanley Rachman and his colleagues have labeled *desynchrony*.

Of special interest in the current context is a particular form of desynchrony in which explicit emotion, as represented by the conscious, subjective feeling state is absent, but behavioral and somatic signs of emotion persist. At present, most evidence for this desynchrony comes in the form of clinical anecdote, rather than controlled experiment, but it is comforting that just such a desynchrony is predicted by a neuropsychological model of fear recently offered by Joseph LeDoux. LeDoux proposes that environmental stimuli are first processed by sensory centers in the thalamus, which then pass information about emotional events to the amygdala, which in turn generates appropriate behavioral, autonomic, and endocrine responses. Information about these responses is also passed to cortical centers supporting working memory, where it is integrated with information provided by thalamic centers about the fear stimulus, thus generating the full-blown subjective experience of *being afraid of something*. However, if a disconnection between thalamus and cortex prevents the fear-eliciting stimulus from being represented in working memory, the person will experience fear without being aware of the fear stimulus. In this case, emotion will serve as an implicit expression of perception or memory, as described earlier. Alternatively, if there is a disconnection between the amygdala and the cerebral cortex, the person will behave in a fearful manner without feeling fear or anxiety. In this case, there will be a dissociation between explicit and implicit emotion.

Neuroscientific theory aside, a potentially interesting approach to implicit emotion has been offered by Anthony Greenwald and Mahzarin Banaji in their application of the explicit-implicit distinction to the social-psychological concept of *attitude*. Attitudes are affective dispositions to like or dislike certain things, and like motives they are usually measured explicitly by self-report scales. However, Greenwald and Banaji have suggested that people may possess positive and negative *implicit attitudes* about themselves and other people, which can affect ongoing social behavior outside of conscious awareness. Like McClelland's implicit motives, implicit attitudes are assessed in terms of task performance rather than self-report. When implicit attitudes diverge greatly from their explicitly expressed counterparts, we have inklings of a dissociation between explicit and implicit emotion. Unfortunately, the experimental literature on implicit attitudes rarely offers a direct contrast with between explicit attitudes, so we do not yet know whether such dissociations actually occur, in what kinds of people, and under what circumstances. Moreover, controversy persists over the IAT itself – whether it actually measures attitudes, and, if so, whether these attitudes are actually unconscious. Still this experimental approach to implicit emotion is very promising.

The Unconscious and the Brain

In the 19th century, Hartmann distinguished between the relative (psychological) unconscious and the physiological and physical unconscious, and we can make a similar distinction between unconscious mental life and unconscious physiological and biochemical processes. Most of our bodily functions proceed unconsciously, without any direct introspective awareness of them or direct voluntary control over them, and this is as true of the nervous system as it is of the other systems in the body. At the molecular and cellular levels of analysis, we have no awareness of the depolarization of cell membranes, the transmission of action potentials down the axon, or the processes of synaptic transmission. We are not aware of the transmission of neural impulses up and down the spinal cord, or the patterns of neural activity which constitute the cortical representation of experience, thought, and action afferent.

We are not aware of the saccades by which we refresh our retinal images, or of the opponent processes which allow us to see various combinations of red, yellow, green, and blue. Spinal reflexes are spared in neurological patients who have been rendered paraplegic or quadriplegic by virtue of a complete break in the spinal cord. But the patients have no direct awareness of these responses, no ability to voluntarily inhibit them, and no ability to initiate similar actions, at will, in the absence of effective stimuli. In the absence of biofeedback technology, we have no awareness of our blood pressure or of whether our brains are generating “alpha”, “beta”, “delta”, or “theta” activity. These physiological and electrochemical processes in the body are as unconscious as photosynthesis and cell division, the tides and plate tectonics. But these things are not what “the unconscious” is all about. To refer to them as unconscious, as Hartmann did, is to make what the philosopher Gilbert Ryle called a *category mistake*: the error of ascribing to one domain a feature attributable only to another. Consciousness belongs to the domain of the mental, and so does the psychological unconscious. It makes no sense to apply the distinction to anything other than mental life.

The revival of interest in consciousness among psychologists and other cognitive scientists has led to a revival of interest in its physiological substrates. A number of neural correlates of consciousness (NCCs) have been proposed, such as activity in the reticular formation or the left hemisphere. However, the reticular formation appears to be responsible for maintaining cortical arousal, which is not the same as consciousness; and careful studies of split-brain patients make it clear that the right hemisphere has a consciousness of its own, even if lacks the ability to communicate its experiences verbally. More recent proposals have similar problems. For example, Francis Crick and Christof Koch have proposed that the NCC consists of synchronized firings (e.g., at 40 Hz) of different various cortical neurons representing features of an object; however, it also may be that such firings serve merely to bind various features of an object together to form a unified mental representation, and that other neural processes determine whether that representation will be consciously accessible. Giulio Tononi and Gerald Edelman have proposed that consciousness results from the firings of neurons distributed widely over the thalamocortical system; but this may be true of any mental representation, conscious or not.

The 19th-century phrenologists did not assign the “faculty” of consciousness to any portion of the brain, perhaps because they viewed consciousness as an intrinsic property of mental life. However, documented dissociations between explicit and implicit cognition, emotion, and motivation offer the possibility of finding portions of the brain that are involved in conscious but not unconscious, or unconscious but not conscious, mental life. For example, Schacter has hypothesized the existence of a *conscious*

awareness system (CAS), a brain module which supports conscious awareness in various domains such as perception and memory. Connection between a module supporting memory and the CAS would render memories accessible to conscious awareness; a disconnection (to use a term coined by Norman Geschwind) between these modules would impair conscious recollection, but spare implicit expressions of memory.

Schacter's proposal is congruent with a general principle of the modularity of the mind and brain, but some considerations suggest that there might be multiple CASs, each supporting conscious awareness in different mental domains. Thus, while it appears that the hippocampus plays a special role in explicit memory, it plays essentially no role in conscious vision. The striate cortex (Area V1) mediates conscious visual experience, but plays no special role in conscious recollection. As noted earlier, LeDoux has proposed that conscious emotional experiences (or, at least, the experience of fear) are mediated by connections between the amygdala and cortical structures supporting working memory, while unconscious behavioral and physiological expressions of emotion are mediated by connections between the amygdala and subcortical structures. Perhaps newly emerging brain-imaging techniques applied to patients or subjects who display a dissociation between explicit and implicit memory, or the like, will be able to reveal cortical structures which are differentially involved in conscious and unconscious mental life.

On the other hand, it may be that brain-imaging approaches will not reveal the neurophysiological differences between conscious and unconscious mental life after all. For example, it has been proposed that all conscious mental states refer to the *self* as the agent or patient of some action, or the stimulus or experience of some state; unconscious mental states lack this kind of self-reference (Kihlstrom, 1997). If so, then conscious awareness is not a matter of the activity of one or more brain modules, but rather of a connection of some sort between the mental representation of the action or state in question, and a mental representation of the self. Such knowledge representations, we now believe, are generated by ensembles of neurons distributed widely across the cortex, rather than by specific cortical loci (e.g., "grandmother" cells). If this notion is correct, then brain imaging techniques such as fMRI, which are designed to identify specific neural loci involved in one kind of mental activity or another, but cannot discriminate between neural representations of specific mental contents, cannot reveal the NCC.

Whither the Unconscious?

The initial discovery of the unconscious, which was consolidated at the turn of the 20th century, has been revived, and the process of rediscovery is well along as we turn to the 21st. There is incontrovertible evidence for automatic mental processes and for implicit memories. Implicit perception is, perhaps, less convincingly established at this point, and implicit learning remains controversial as well. Still the evidence favoring both concepts cannot be dismissed out of hand. Research on implicit thought is admittedly immature, but the evidence in hand is quite provocative. Based on the evidence for the cognitive unconscious, implicit motivation and implicit emotion cannot be dismissed out of hand, but we still require convincing evidence that they can be dissociated from their explicit counterparts. Still, it is clear that the paradigms developed in the study of implicit memory provide a vehicle for exploring all aspects of the psychological unconscious. In response to Immanuel Kant, we can say that priming and other methodologies do in fact enable us to infer that we have ideas, even though we are not conscious of them. And in response to William James, we can say that these same methodologies, rigorously applied, will prevent us from believing whatever we like about the unconscious mind.

In that respect, it must be emphasized that the scope of the psychological unconscious, broad as it is, does not appear to be so broad as to encompass the unconscious of psychoanalytic theory. There is no evidence, in any of the evidence summarized in this article, favoring Freud's view that the unconscious is the repository of primitive, infantile, irrational, sexual and aggressive impulses, repressed in a defensive maneuver to avoid conflict and anxiety (Bornstein and Masling, 1998; Levy, 1996; Shevrin et al., 1999). Nor is there any evidence to support the more extreme clinical lore concerning unconscious representations of trauma, or the excesses of the recovered memory movement in psychotherapy (Conway, 1997). In this case, as James warned, the unconscious does indeed seem to be a tumbling-ground for whimsies.

References

- Bargh, J.A., 1997. The automaticity of everyday life. In: Wyer, R.S. (Ed.), *Advances in Social Cognition*, vol. 10. Erlbaum, Mahwah, N.J., pp. 1–61
- Bornstein, R.F., Pittman, T.S., 1992. *Perception without Awareness: Cognitive, Clinical, and Social Perspectives*. Guilford, New York.
- Bornstein, R.F., Masling, J.M. (Eds.), 1998. *Empirical Perspectives on the Psychoanalytic Unconscious*. American Psychological Association, Washington, D.C.
- Bowers, K.S., Meichenbaum, D. (Eds.), 1984. *The Unconscious Reconsidered*. Wiley Interscience, New York.
- Conway, M.A. (Ed.), 1997. *Recovered Memories and False Memories*. Oxford University Press, Oxford.
- Ellenberger, H.L., 1970. *The Discovery of the Unconscious: The History and Evolution of Dynamic Psychiatry*. Basic, New York.
- Feldman-Barrett, L., Niedenthal, P.M., Winkielman, P. (Eds.), 2005. *Emotion and Consciousness*. Guilford, New York.
- Hassin, R.R., Uleman, J.S., Bargh, J.A., 2005. *The New Unconscious*. Oxford University Press, New York.
- Kelly, W.L., 2001. *The Psychology of the Unconscious: Mesmer, Janet, Freud, Jung, and Current Issues*. Prometheus, Buffalo, N.Y.
- Kihlstrom, J.F., 1997. Consciousness and me-ness. In: Cohen, J.D., Schooler, J.W. (Eds.), *Scientific Approaches to Consciousness*. Erlbaum, Mahwah, N.J., pp. 451–468
- Kihlstrom, J.F., 2012. Unconscious processes. In: Reisberg, D. (Ed.), *Oxford Handbook of Cognitive Psychology*. Oxford University Press, Oxford, pp. 176–186.
- Klein, D.B., 1977. *The Unconscious: Invention or Discovery? a Historico-critical Inquiry*. Goodyear, Santa Monica, CA.

* *Change History*: July 2017. John F. Kihlstrom, with major changes to the body of the text and major updates to the reference list.

- Levy, D., 1996. *Freud Among the Philosophers: The Psychoanalytic Unconscious and its Philosophical Critics*. Yale University Press, New Haven, CT.
- Reber, A.S., 1993. *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford University Press, New York.
- Reber, L.S., 1996. *Implicit Memory and Metacognition*. Erlbaum, Hillsdale, N.J.
- Schacter, D.L., 1987. Implicit memory: history and current status. *J. Exp. Psychol. Learn. Mem. Cognition* 13, 501–518. <https://doi.org/10.1037/0278-7393.13.3.501>.
- Schultheiss, O.C., Brunstein, J.C., 2010. *Implicit Motives*. Oxford University Press, Oxford.
- Shevrin, H., Bond, J.A., Brakel, L.A.W., Hertel, R., Williams, W.J., 1999. *Conscious and Unconscious Processes: Psychodynamic, Cognitive, and Neurophysiological Convergences*. Guilford, New York.
- Stein, D.J., 1997. *Cognitive Science and the Unconscious*. American Psychiatric Press, Washington, D.C.
- Underwood, G., 1996. *Implicit Cognition*. Oxford University Press, Oxford.
- Weiskrantz, L., 1986. *Blindsight: A Case Study and its Implications*. Oxford University Press, Oxford.
- Weiskrantz, L., 1997. *Consciousness Lost and Found: A Neuropsychological Exploration*. Oxford University Press, New York.
- Whyte, L.L., 1960. *The Unconscious before Freud*. Basic, New York.