To appear in *International Encyclopedia of the Social and Behavioral Sciences*

Abstract

Characterizing the relationships between conscious and unconscious processes is one of the most important and long-standing goals of cognitive psychology. Renewed interest in the nature of consciousness — long considered not to be scientifically explorable —, as well as the increasingly widespread availability of functional brain imaging techniques, now offer the possibility of detailed exploration of the neural, behavioral, and computational correlates of conscious and unconscious cognition. This entry reviews some of the relevant experimental work, highlights the methodological challenges involved in establishing the extent to which cognition can occur unconsciously, and situates ongoing debates in the theoretical context provided by current thinking about consciousness.

Conscious and unconscious processes in cognition

Axel Cleeremans

Université Libre de Bruxelles, Brussels, Belgium

Everyday experience suggests that we often seem to know more than we can tell. Riding a bicycle, playing tennis or driving a car, for instance, all involve mastering complex sets of motor skills, yet we are at a loss when it comes to explaining exactly how we perform such physical feats. These dissociations between our ability to report on cognitive processes and the behaviors that involve these processes are not limited to action but extend to higher-level cognition as well. Most native speakers of a language are unable to articulate the grammatical rules they nevertheless follow when uttering expressions of the language. Likewise, expertise in domains such as medical diagnosis or chess, as well as social or aesthetic judgments, all involve intuitive knowledge that one seems to have little introspective access to.

We also often seem to tell more than we can know. In a classic article, social psychologists Nisbett and Wilson (1977) reported on many experimental demonstrations of the fact that accounts of our own behavior frequently reflect reconstructive and interpretative processes rather than genuine introspection. In one such study, Nisbett and Wilson asked patrons of a department store to choose the best quality item among four identical pairs of nylon stockings in the context of what subjects believed was a consumer survey. There was a strong position effect: Items located on the right, which were inspected last, were chosen much more often than items in other positions. Yet, when asked to motivate their choice, not a single subject mentioned position as a relevant factor, and many denied that it might have had an effect when asked directly about it. More generally, while it is generally agreed that cognitive processes are not in and of themselves open to any sort of introspection, Nisbett & Wilson (1977) further claimed that we can sometimes be "(a) unaware of the existence of a stimulus that importantly influenced a response, (b) unaware of the existence of the response, and (c) unaware that the stimulus has affected the response". (p. 231).

Just as for knowledge of our own cognitions, our knowledge of the world turns out to be much less complete that it might appear, thus further deconstructing Cartesian conceptions of cognition in which conscious awareness reigns supreme. In change blindness and inattentional blindness experiments, for instance, people are often surprisingly unable to detect large changes in visual stimuli that they are allowed to observe at their leisure. One such experiment (e.g., Simons & Chabris, 1999) involved watching a movie depicting two teams of three basketball players and counting the number of times that a ball was passed between the players of each team. Under these conditions, that is, with attention fully engaged in tracking one of the two balls, people often failed to notice that, halfway through the movie, an actor dressed as a gorilla appeared on the right of the screen, walked slowly towards the middle of the scene where he repeatedly banged his chest, subsequently resuming his stroll to disappear on the left!

Demonstrations of dissociations between conscious awareness and stimulation and/or behavior have now been reported in many domains of psychology. Frith, Perry and Lumer (1999) have suggested to organize paradigms through which to study the “neural correlates of consciousness” in nine groups resulting from crossing two dimensions: (1) three classes of psychological processes involving respectively knowledge of the past, present, and future — memory, perception, and action —, and (2) three types of cases where subjective experience is incongruent with the objective situation — cases where subjective experience fails to reflect changes in either (1) the

stimulation or in (2) behavior, and (3) cases where subjective experience changes whereas stimulation and behavior remain constant.

Among such paradigms, the study of dissociations between conscious and unconscious perceptual processes has by far the longest history, extending back to the beginnings of scientific psychology in the 19th century. Subliminal perception (see Dixon, 1971) is said to occur when stimuli presented below the threshold of awareness influence information processing, emotion, or action. In a classic, but highly controversial series of studies, Anthony Marcel suggested that semantic priming, that is, the facilitation of decisions about semantic aspects of a target stimulus when it follows a related stimulus (the prime), can occur in conditions when the prime is presented so rapidly that subjects fail to be able to decide whether it was present or not (e.g. Marcel, 1983). While many subsequent studies have reported similar findings with different stimuli such as spoken words or faces, many have also failed to replicate such findings and the domain remains highly controversial today (Holender, 1986). Nevertheless, it appears that considerable information can be perceived and processed in the absence of conscious awareness when it is assessed through subjective measures: People can believe that they are guessing yet perform better than chance would predict in a variety of tasks involving semantic and affective decisions.

Some forms of learning and memory can also be unconscious. Memory for previous events can be expressed explicitly, as a conscious recollection, or implicitly, as automatic, unconscious influences on behavior. Numerous studies have demonstrated dissociations between implicit and explicit memory. Amnesic patients, for instance, who exhibit severe or total loss in their ability to explicitly recall previous experiences, nevertheless retain the ability to learn novel procedural skills (Milner, 1962) or to exhibit sensitivity to past experiences of which they are not conscious (e.g., Warrington & Weiskrantz, 1968). Some studies have explored the extent to which patients undergoing general anesthesia are able to learn novel associations, and one study has even demonstrated learning in foetuses (van Heteren et al., 2000).

Similar dissociations can be observed with normal subjects (see Cleeremans, Destrebecqz, & Boyer, 1998). Arthur Reber, in a classic series of studies conducted in 1967, first suggested that learning might be "implicit", to the extent that people appear to be able to learn new information without intending to do so and in such a way that the resulting knowledge is difficult to express. In Reber's seminal study of artificial grammar learning, subjects were asked to memorize a set of meaningless letter strings generated by a simple set of rules embodied in a finite-state grammar. After this

memorization phase, they were told that the strings followed the rules of a grammar, and were asked to classify novel strings as grammatical or not. In this experiment and in many subsequent replications, subjects were able to perform this classification task better than chance would predict, despite remaining unable to describe the rules of the grammar in verbal reports. This dissociation between classification performance and verbal report is the finding that prompted Reber to describe learning as implicit, for subjects appeared sensitive to and could apply knowledge (the rules of the grammar) that they remained unable to describe and had had not intention to learn.

In a series of experiments that attracted renewed interest in implicit learning, Berry and Broadbent (1984) showed that success in learning how to control a simulated system (e.g., in one set of experiments, a system described to subjects as a "sugar factory") so as to make it reach certain goal states was independent from ability to answer questions about the principles governing subject's inputs and the system's output: Practice selectively influenced ability to control the system, whereas verbal explanations about how the system works selectively influenced ability to answer questions.

Today, a third paradigm — sequence learning — has become dominant in the investigation of implicit learning. Nissen and Bullemer (1987) first demonstrated that subjects asked to respond as fast and as accurately as possible to a series of visual events progressively learned about the sequential structure of the stimulus sequence in spite of showing little evidence of being aware that the material contained structure. Numerous subsequent studies of this effect have indicated that subjects can learn about complex sequential relationships despite remaining unable to fully elicit this knowledge in corresponding direct, explicit tasks (e.g., Cleeremans & McClelland, 1991).

One of the most convincing demonstrations of dissociations between conscious knowledge and behavior was obtained in a eye-blink conditioning situation (Perruchet, 1985). In this experiment, people were exposed to a series of identical tones, 50% of which could be followed after a short interval by an air puff directed to the left cornea. Immediately after each tone was presented (and before the puff occurred in reinforced trials), people were asked to indicate the extent to which they expected the tone to be followed by an air puff on a 0-7 points scale. A trial-by-trial analysis of the results indicated that eye blink responses were increasingly more likely to occur after presentation of a tone if the corresponding trial had been preceded by a series of reinforced trials (i.e., trials during which the tone had indeed been followed by an air puff). In stark contrast, however, people’s subjective expectancy of the

occurrence of an air puff tended to decrease with the number of reinforced trials that preceded the trial under consideration. In other words, people's eye blink responses were completely dissociated from their conscious expectations about when each tone would be followed by an air puff.

Finally, the neuropsychological literature also contains many striking cases of dissociation between subjective experience and behavior. In addition to the dissociations exhibited by patients with amnesia, other syndromes have also attracted attention over the last decades. Blindsight patients, for instance, suffer from damage to the primary visual cortex, and report being unable to see in some portions of their visual field. Larry Weiskrantz (1986) showed that such patients are nevertheless often able to act appropriately when stimulated in their blind field, for instance by initiating correct reaching or grasping movements towards a target. Likewise, the study of patients with form agnosia or neglect has offered valuable insight into the neural correlates of consciousness. For instance, Milner and Goodale (e.g., 1998) have described a patient, D.F., who, while unable to offer verbal or manual reports on the orientation of a slot, nevertheless retains the ability to actually insert a card in it, accurately preparing her movement well before the slot is reached. These observations led Milner and Goodale to suggest that distinct "what" and "where" processing pathways exist in the human visual system, and that only one of these pathways (the ventral, "what" system) is associated with conscious visual experience.

2 Consciousness and theories of the psychological unconscious

These different findings all suggest that unconscious influences on behavior are pervasive. This raises the question of how to best characterize the relationships between conscious and unconscious processes, and in particular whether one should consider that mental representations can be unconscious. The idea that mental life includes both conscious and unconscious events has been expressed most clearly by Freud (e.g., 1949), whose psychoanalytical theory has profoundly influenced both scientific thought and public conceptions of the mind. Even though one can find earlier characterizations of cognition in which computations that are not accompanied by awareness contribute to conscious decisions (e.g. Leibniz, 1704/1981; de Biran, 1804), most thinkers up until Freud, including René Descartes and, arguably, William James, considered that mental life consists exclusively of conscious events. This position is still endorsed by some authors today, even though there is wide consensus about the idea that cognitive processing is largely unconscious.

Freud's characterization of the "dynamic unconscious", however, makes very specific assumptions — specifically that there exist unconscious mental representations and that these representations can reflect semantic and affective dimensions of processing. Further, the unconscious, as Freud depicted it, is as dynamic and causally efficacious as the conscious — it is a mental bubbling cauldron of sorts, replete with repressed instincts, thoughts and feelings vying to get access to awareness. Few thinkers would endorse Freud's characterization of the unconscious in full today — often choosing instead to deny its existence altogether or considering that only shallow aspects of processing can take place unconsciously. In particular, while popular belief has often tended to ascribe powerful abilities to the unconscious, empirical exploration of the level of analysis at which processing occurs in the absence of consciousness has failed to offer convincing demonstrations that unconscious processing can be as flexible or as deep as conscious processing. Rather, the evidence suggests that unconscious processing can bias conscious processing in a way that reflects strong or habitual responses.

An important issue in this context is to determine the extent to which conscious and unconscious processing depend on separable neural systems, and what their relationship should be. Contemporary theories of consciousness make widely different assumptions about its underlying mechanisms. Atkinson, Thomas & Cleeremans (2000) proposed to organize such theories along two dimensions defined by (1) whether the theory assumes that consciousness depends on the involvement of certain processes or rather on the properties of certain mental representations, and by (2) whether the theory assumes that consciousness depends on the involvement of specific modules or whether is can occur anywhere in the brain. Computational models of information processing play a significant role in fostering the development of novel theories of consciousness. In neural network — or connectionist — models, for instance, task-relevant knowledge is embedded in the same structures that support processing itself. In contrast to traditional information processing frameworks, knowledge is thus implicit in such models to the extent that it cannot be separated from the mechanisms that subserve processing.

If consciousness does not depend on neural systems specifically dedicated to subserve subjective experience but should instead be viewed as an emerging property of the processing conducted by many different regions of the brain, one might expect to find indications that consciousness fails to be a unified, all-or-none phenomenon. Marcel (1993) described a relevant series of experiments probing the unity of consciousness in normal subjects. In these experiments, people were asked to detect changes in the luminance of very faint stimuli in three different ways: by blinking, by pressing on a

button, or by a verbal response ("yes"). When these responses had to be produced simultaneously, they often tended to be dissociated, with subjects answering positively, for instance, through a blinking response, but negatively through the other modalities. Another experiment indicated that the probability of correct guessing varied significantly across modalities, thus suggesting that they each afforded varying degrees of availability to conscious awareness.

3 Methodological challenges

Because there is no accepted operational definition of what it means for an agent to be conscious of something, complex measurement challenges arise in the study of the relationships between conscious and unconscious cognition.

First, consciousness is not a single process or phenomenon, but rather encompasses many dimensions of experience. A first important challenge thus arises in delineating which aspects of consciousness count when assessing whether a subject is aware or not of a particular piece of information: awareness of the presence or absence of a stimulus, conscious memory for a specific previous processing episode, awareness of one's intention to use some information, or awareness that one's behavior is influenced by some previous processing episode. Different aspects of conscious processing are engaged by different paradigms. In subliminal perception studies, for instance, one is concerned with determining whether stimuli that have not been consciously encoded can influence subsequent responses. In contrast, implicit memory research has been more focused on retrieval processes, that is, on the unintentional, automatic effects that previously consciously perceived stimuli can exert on subsequent decisions. In studies of implicit learning, it is the relationships between ensembles of consciously processed stimuli that remain purportedly unconscious. These subtle differences in which specific aspects of the situation are available to awareness illustrate the need to carefully distinguish between awareness during encoding and awareness during retrieval of information. Further, both encoding and retrieval can concern either individual stimuli or relationships between sets of stimuli, and both can either be intentional or not.

A second important challenge is to devise an appropriate measure of awareness. Most experimental paradigms dedicated to exploring the relationships between conscious and unconscious processing have relied on a simple dissociation logic aimed at comparing the sensitivity of two different measures to some relevant information: A measure C of subjects' awareness of the information, and a measure P of behavioral sensitivity to the same information in the context of some task. Unconscious

processing, according to the simple dissociation logic, is then demonstrated whenever P exhibits sensitivity to some information in the absence of correlated sensitivity in C. There are several potential pitfalls with the simple dissociation logic, however. First, the measures C and P cannot typically be obtained concurrently. This "retrospective assessment" problem entails that finding that C fails to be sensitive to the relevant information need not necessarily imply that information was processed unconsciously during encoding, but that, for instance, it might have been forgotten before retrieval. A second issue is to ensure that the information revealed through C is indeed relevant to perform the task. As Shanks & St. John (1994) have suggested, many studies of implicit learning have failed to respect this "information" criterion. For instance, successful classification in an artificial grammar learning task need not necessarily be based on knowledge of the rules of the grammar, but can instead involve knowledge of the similarity relationships between training and test items. Subjects asked about the rules of the grammar would then understandably fail to offer relevant explicit knowledge. A third issue is to ensure that C and P are both equally sensitive to the relevant information. At first sight, verbal reports and other subjective measures such as confidence ratings would appear to offer the most direct way through which to assess the contents of subjective experience, but such measures are often difficult to operationalize in a sufficiently controlled manner. For instance, people might simply refrain from reporting on knowledge held with low confidence, or might offer reports that are essentially reconstructive in nature, as Nisbett and Wilson's experiments indicate. For this reason, many authors have advocated using so-called objective measures of awareness. Objective measures of awareness include forced-choice tests such a recognition, presence-absence decisions, or identification.

Even if these different conditions are fulfilled, however, it might be elusive to hope to be able to obtain measures of awareness that are simultaneously exclusive and exhaustive with respect to knowledge held consciously. In other words, finding null sensitivity in C, as required by the dissociation paradigms for unconscious processing to be demonstrated, might simply be impossible because no such absolute measure exists. A significant implication of this conclusion is that, at least with normal participants, it makes little sense to assume that conditions exist where awareness can simply be "turned off". Much of the ongoing debate about the existence of subliminal perception can be attributed to a failure to recognize the limitations of the dissociation logic.

It might therefore instead be more plausible to assume that any task is always sensitive to both conscious and unconscious influences. In other words, no task is process-pure. Two methodological approaches that specifically attempt to overcome

the conceptual limitations of the dissociation logic have been developed. The first was introduced by Reingold & Merikle (1988), who suggested that the search for absolute measures of awareness should simply be abandoned in favor of approaches that seek to compare the sensitivity of direct measures and indirect measures of some discrimination. Direct measures involve tasks in which the instructions make explicit reference to the relevant discrimination, and include objective measures such a recognition or recall. In constrast, indirect measures, such as stem completion in memory tasks, make no reference to the relevant discrimination. By assumption, direct measures should exhibit greater or equal sensitivity than indirect measures to consciously held task-relevant information, for subjects should be expected to be more successful in using conscious information when instructed to do so than when not. Hence, demonstrating that an indirect task is more sensitive to some information than a comparable direct task can only be interpreted as indicating unconscious influences on performance.

The second approach — Larry Jacoby's "Process Dissociation Procedure" (Jacoby, 1991) constitutes one of the most significant advances in the study of differences between implicit and explicit memory. It is based on the argument that, just as direct measures can be contaminated by unconscious influences, indirect measures can likewise be contaminated by conscious influences: Particular tasks can simply not be identified with particular underlying processes. The process dissociation procedure thus aims to tease apart the relative contributions of conscious and unconscious influences on performance. To do so, two conditions are compared in which conscious and unconscious influences either both contribute to performance improvement, or act against each other. For instance, subjects might be asked to memorize a list of words and then, after some delay, to perform a stem completion task in which word stems are to be completed either so as to form one of the words memorized earlier (the inclusion condition) or so as to form a different word (the exclusion condition). If the stems nevertheless tend to be completed by memorized words under exclusion instructions, then one can only conclude that memory for these words was implicit, since if subjects had been able to consciously recollect them, they would have avoided using them to complete the stems. Numerous experiments have now been designed using the process dissociation procedure. They collectively offer convincing evidence that performance can be influenced by unconscious information in the absence of conscious, subjective awareness.

4 Treating consciousness as a variable: Qualitative differences

Rather than pursuing the elusive goal of demonstrating unconscious cognition, a more fruitful approach to exploring the relationships between conscious and unconscious cognition thus consists of treating consciousness as a variable, that is, of exploring the functional, neural, and computational differences that exist between tasks performed with awareness and without. This leads to the design of experiments in which qualitative differences between cognition with and without awareness can be established in carefully controlled conditions. In this respect, the "contrastive approach" advocated by Baars (1988) and Frith, Perry & Lumer (1999) offers great promise in helping solve the methodological challenges described above. Numerous qualitative differences between conscious and unconscious processes have been reported. For instance, based on behavioral methods, Arthur Reber and others have suggested that implicit cognition involves processes that are more resistant to insult or injury, that tend to be less sensitive to individual differences or affective state, and less sensitive to attentional demands. Recent advances in brain imaging methods make it possible to go beyond purely behavioral methods and to actually integrate different online measures of awareness and of brain activity when attempting to contrast conscious and unconscious cognition. One of the most interesting paradigms through which to combine so-called first person (subjective) and third-person (objective) methodologies is binocular rivalry — a phenomenon that takes place when different images are shown to each eye. Under these conditions, one's perception of the stimulus spontaneously alternates between the complete perception of one stimulus and the complete perception of the other. Using functional brain imaging, one can then correlate people's subjective reports of what it is they actually perceive at each point in the experiment with the cerebral regions that are most activated, and in so doing establish which of these regions correlate most with the unchanging stimulus and which correlate most with perception as assessed by subjective reports.

Numerous studies combining various online measures of brain activity have likewise attempted to identify which regions of the brain are most involved in processes such as conscious recollection, intentional retrieval, or non-conscious influences in different memory tasks (see Shacter, Buckner & Koutsaal, for a review). While it is too early to draw definite conclusions from such studies, it appears, for instance, that anterior prefrontal regions of the brain appear to be most clearly associated with processes of effortful, conscious retrieval. This and similar findings should help us further develop theories of consciousness that have a firm rooting in our knowledge of the brain.

5 Conclusions

The study of differences between conscious and unconscious processing is a major endeavor for the cognitive neurosciences. Conscious and unconscious processing differ on several dimensions, including depth and specificity of processing. Action can be initiated without consciousness, and some aspects of perception or memory can also take place without consciousness. Such unconscious processing always tend to reflect habitual or strong responses. From this perspective, unconscious processing is best characterized as the indirect effects of conscious processing. The extent to which conscious and unconscious processes involve distinct brain regions remains an open issue, but it is clear that some regions of the brain tend to be more associated with conscious processing than others. The distinction between conscious and unconscious processing thus appears to involve both a graded and dynamic continuum (in terms of the underlying mechanisms) and a dichotomy (in terms of subjective experience). Functional brain imaging methods, when used together with sufficiently sensitive behavioral methods, offer the promise of combining first-person and third-person perspectives to elucidate the relationships between conscious and unconscious processes in cognition.

References

Atkinson, A., Thomas, M., & Cleeremans, A. (2000). Consciousness: Mapping the theoretical landscape. *Trends in Cognitive Sciences, 4(10),* 372-382.

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.

Berry, D.C. & Broadbent, D.E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, *36*, 209-231.

Cleeremans, A. & McClelland, J.L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology : General*, *120*, 235-253.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406-416.

de Biran, M. (1804/1929). *The influence of habit on the faculty of thinking.* Baltimore, Williams & Wilkins.

Dixon, N.F. (1971). *Subliminal Perception*. McGraw-Hill.

Freud, S. (1949). *An Outline of Psychoanalysis*. (J. Strachey, Tr.). London: Hogarth Press.

Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: An experimental framework. *Trends in Cognitive Sciences*, *3*, 105-114.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9*, 1-66.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513-541.

Leibniz, G.W. (1704/1981). *New essays on human understanding* (P. Remnant & J. Bennet, Trans. & Eds.). Cambridge, England: Cambridge University Press.

Marcel, A.J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, *15*, 197-237.

Marcel, A. J. (1993). Slippage in the unity of consciousness. In *Experimental and Theoretical Studies in Consciousness (*C*iba Foundation Symposium 174)*. John Wiley and Sons.

Milner, A.D., & Goodale, M.A. (1998). *The visual brain in action*. Oxford University Press.

Milner, B. (1962). Les troubles de la mémoire accompagnant des lésions hippocampiques biltatérales. In *Physiologie de l'hippocampe*. Paris: Centre National de la Recherche Scientifique.

Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

Nissen, M.J. & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1-32.

Perruchet, P. (1985). A pitfall for the expectancy theory of human eyelid conditioning, *Pavlovian Journal of Biological Science*, *20*, 163-170.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855-863.

Reingold, E.M., & Merikle, P.M. (1988). Using direct and indirect measures to study perception without awareness, *Perception and Psychophysics*, *44*, 563-575.

Shanks, D.R., & St. John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367-447.

Shacter, D.L., Buckner, R.L., & Koutsaal, W. (1998). Memory, consciousness and neuroimaging. *Philosophical Transactions of the Royal Society of London* B, *353*, 1861-1878.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception, 28*, 1059-94.

van Heteren, C. F., Boekkooi, P. F., Jongsma, H.W., & Nijhuis, J.G. (2000). Fetal learning and memory. *Lancet*, *356*, 1169-1170.

Warrington, E.K., & Weiskrantz, L. (1968). New method of testing long term retention with special reference to amnesic patients, *Nature*, *217*, 972-974.

Weiskrantz, L. (1986). Blindsight: A case study and implications. Oxford, England: Oxford University Press.