The Mythomanias: The Nature of Deception and Self-Deception

Edited by
Michael S. Myslobodsky
Tel-Aviv University

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1997 Mahwah, New Jersey

3

Self-Knowledge and Self-Deception: Further Consideration

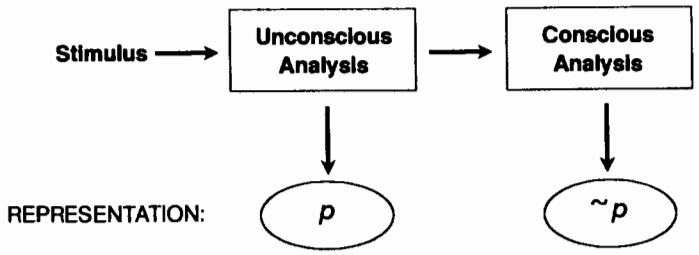
Anthony G. Greenwald

The term *self-deception* describes the puzzling situation in which a person appears both to know and not know one and the same thing. Consider as an example a cancer patient who maintains the expectation of recovery even while surrounded by the signs of an incurable malignancy. Presumably this patient knows unconsciously that the disease is incurable, but manages to prevent that knowledge from becoming conscious. Interestingly, one of the reasons for concluding that the patient unconsciously knows of the incurable malignancy is the very success of the defense. How could that defense be maintained so effectively without using knowledge of the unwelcome fact to anticipate the forms in which it might try to intrude into consciousness?

THE PARADOX OF SELF-DECEPTION

The sense in which this example is puzzling, or paradoxical, is shown in Fig. 3.1. Some encountered situation, or stimulus, is assumed to receive both unconscious and conscious analyses. The unconscious analysis, which is assumed to occur first, identifies a threatening, or anxiety-evoking, aspect of the stimulus. In Fig. 3.1, the anxiety-evoking stimulus is represented as some proposition, *p*—such as, "I have a terminal malignancy." Conscious analysis, however, fails to apprehend this proposition.

There are three puzzling aspects of this situation. First, how can the person manage unconsciously to reach the conclusion that proposition *p* is true while not also reaching that conclusion consciously? Second, what good does it do for the person not to know consciously that *p* is true? Should it not produce anxiety just to know unconsciously that *p* is



Three Questions

1. **How does the person both know and not know *p*?**
2. **What good does it do not to know *p* consciously?**
3. **Why is the faster, more accurate, system unconscious?**

FIG. 3.1. The paradox (or puzzle) of self-deception. In this and other figures in this chapter, theorized stages of cognitive analysis are represented by rectangles, representation outputs are represented by ovals, and observable events (stimuli and responses) are identified leading to or emerging from these entities. In this figure, a proposition and its negation are represented as *p* and ~*p*, respectively.

true? Third, and most puzzling of all, why does the unconscious system give both a faster and a more thorough analysis than the conscious system: Would it not be sensible to have one's most acute cognitive abilities available to consciousness?

Interest in Self-Deception

Self-deception has attracted the interest of scholars of several different disciplines, and for several different reasons. For clinical psychologists and psychiatrists, self-deception is seen as a means of protection from painful knowledge (Murphy, 1975; Sackeim & Gur, 1978; Schafer, 1976). At the same time, it seems a strangely cumbersome method of defense. That is, it appears to create more problems for the psyche than it can possibly solve. How, therefore, can it protect? From this clinical perspective, understanding self-deception has implications for the conduct of psychotherapy.

For cognitive psychologists and philosophers (e.g., Fingarette, 1969), self-deception is seen as a paradoxical condition of knowledge. How does a knowledge system accommodate an apparent internal contradiction? From this epistemological perspective, achieving an understanding of self-deception will shed light on the organization of human knowledge.

For ethologists, self-deception is seen as a strategy that could provide an advantage in animal social interaction (e.g., Lockard, 1980; Trivers, 1985). By unconsciously deceiving itself, an animal might become a more effective deceiver of others. From this perspective, the investigation of self-deception might justify placing the psychological concept of the unconscious under the explanatory umbrella of sociobiology, making it a topic within the emerging subdiscipline of evolutionary psychology.

The intellectual perspectives of clinical psychology, psychiatry, cognitive psychology, epistemology, and sociobiology collectively yield a set of questions that might be answered by a successful analysis of self-deception. First, and most fundamentally, how is self-deception to be conceived of in terms of knowledge organization: If it requires unconscious cognition, how does that unconscious cognition relate to conscious cognition? Second, what is the function of self-deception: What psychic gain results from the combination of knowing something unconsciously while not knowing it consciously? Third, how common is self-deception: Is it an ordinary phenomenon of everyday life, or is it an exotic, even pathological phenomenon? And fourth, what empirical criteria can be used to identify self-deception: How can it be studied in the laboratory?

The theoretical analysis used in this chapter addresses these four questions and, in doing so, describes two theoretical alternatives to the paradox-laden psychoanalytic account of defenses against cognitive threat. One alternative borrows from cognitive psychology the well-established paradigm of an ordered series of *stages*, or *levels*, of information processing. The second theoretical account uses the newer paradigmatic approach of *parallel distributed processing*, or *neural network modeling*. Both of these theoretical approaches allow nonparadoxical interpretation of effective cognitive defenses.[1]

Previous Analyses

The most thorough intellectual analysis of self-deception was provided by philosopher Herbert Fingarette in the 1969 book *Self-Deception*. Fingarette sought to develop a paradox-free account of self-deception. Ultimately, Fingarette's attempt to avoid paradox must be judged unsuccessful. Nevertheless, Fingarette's review and analysis advanced the topic considerably, and provided a stimulating entry point for researchers. Notable among subsequent researchers were Gur and Sackeim (1979; see also Sackeim & Gur, 1978), who provided a careful statement of the

[1]The earlier version of this chapter (Greenwald, 1988b) described how the first of these two approaches (information-processing stages) could produce a nonparadoxical account of apparent self-deceptions. The second approach (network modeling) is newly included in the present chapter.

self-deception paradox and offered a laboratory procedure for its investigation.

Fingarette's (1969) Analysis. Fingarette started by criticizing previous attempts by philosophers to analyze self-deception, and concluded that previous attempts to resolve the paradox of self-deception either (a) had not addressed themselves to the proper phenomena of self-deception, or (b) rather than resolving the paradox, had merely portrayed it in a "variant form." Fingarette's own analysis went part of the way toward a solution, but unfortunately did not escape reintroducing the paradox. This reintroduction occurred in the form of an unnamed mechanism that analyzes the true (threatening) import of circumstances and, on the basis of the knowledge so obtained, purposefully prevents the emergence into consciousness of both the threatening information and the defense against it.

Fingarette's unnamed mechanism was capable of inference and intention in a way that required sophisticated symbolic representation. Yet Fingarette assumed that this mechanism operated outside of the ordinary machinery of inference and symbolic representation—that is, outside of conscious cognition. The paradoxical aspects of Fingarette's unnamed mechanism seem indistinguishable from the paradoxical aspects of Freud's censor (the agency of repression). For Freud, the censor operated from a base within the conscious ego, and although it appeared to have ego's reasoning powers, nevertheless was assumed to operate without ego's consciousness (Freud, 1923/1961). The three questions in Fig. 3.1, which define the paradox of self-deception, apply as much to Fingarette's analysis as to Freud's.

Gur and Sackeim's (1979) Analysis. In seeking to demonstrate the paradoxical character of self-deception, Gur and Sackeim (1979) adapted a voice-recognition task that had been developed about a half-century earlier by Wolff (1932; see also Huntley, 1940). In this task, after making recordings of samples of their own voice, subjects were asked to judge whether each of a series of played-back samples was or was not their own voice. The critical evidence comes from examining the relationship between occurrences of skin conductance response (SCRs) and overt verbal identification responses to the voice stimuli. The SCR is assumed to indicate unconscious own-voice recognition, whereas verbal identification indicates conscious recognition. Self-deception is judged to occur when the SCR occurs on an own-voice trial, yet the subject fails to identify the voice as self.

Why are such trials paradoxical? It is not simply that the SCR and verbal response appear to disagree. That disagreement could be explained nonparadoxically (and not very interestingly) by assuming, for example, that the skin conductance system is more prone to error (perhaps by influence from stray events), or that it is susceptible to sources of error that differ from those that disrupt verbal identification. The response disagreement becomes interestingly paradoxical, however, when one concludes that the SCR reflects an unconscious own-voice identification that plays a role in the purposeful blocking of conscious identification. It is therefore relevant that Gur and Sackeim demonstrated a correlation between individual differences in voice identification accuracy and scores on a Self-Deception Questionnaire measure, suggesting a motivated blocking of conscious voice recognition that is initiated by a knowing observer operating outside of conscious cognition.

Resolving the Paradox by Changing Assumptions

The Assumption of Personal Unity. Paradoxes stimulate theoretical advance by making it apparent that there is a troublemaker lurking somewhere among one's theoretical assumptions. One candidate troublemaker, in the case of the self-deception paradox, is the assumption of personal unity that implicitly underlies much psychological theory (cf. Greenwald, 1982). This is the assumption that each person's knowledge is organized into a single, unified system. It is the assumption that implicitly justifies use of the word *individual* (i.e., an indivisible entity) to refer to the person. Alternatively, it might be assumed that there are dissociations within personal knowledge systems (Hilgard, 1977). For example, in the case of the voice-recognition task, one might assume that the right hemisphere (or some other modular brain subsystem; see Gazzaniga, 1985) controls the SCR independently of the left hemisphere's control of verbal-identification responses. With such an abandonment of the assumption of unity within the knowledge system, discrepancies between SCR and verbal identification of own-voice stimuli are no longer paradoxical—no more than it would be paradoxical for two different people to disagree in identifying the same voice.

Abandoning the assumption of personal unity seems a drastic step. At the same time that one gains the ability to explain findings of discrepancy between response systems, one gives up at least some of the ability to explain *relationships* between response systems—relationships of the sort that are heavily appealed to in psychological theory, for example, in the influential mediationist behaviorisms of Spence (1956), Mowrer (1960), or Osgood (1953), in the information processing theories of the cognitive revolution (e.g., Smith, 1968; Sternberg, 1969), and in cognitive interpretations of emotion such as those of Schachter and Singer (1962) or Lazarus (1984). In the last decade, however, the assumption of personal unity has received a substantial indirect attack, in the develop-

ment of the concept of parallel distributed processing, or neural network modeling (e.g., Rumelhart & McClelland, 1986).

The Assumption of a Coordinate Unconscious. There is another possible troublemaking assumption that many will find more easily sacrificed than the assumption of personal unity. This expendable troublemaker is the assumption of a *coordinate unconscious*—the assumption that unconscious and conscious cognition are coordinate, or equivalent in power, and therefore capable of the same types of mental operations. When the assumption of a coordinate unconscious is abandoned, it becomes possible to set unconscious cognition into a multilevel conception of mental representations—a conception that readily provides nonparadoxical explanations of phenomena to which the *self-deception* label has been attached.

An attractive alternative to the coordinate unconscious is a conception of unconscious cognition that is decidedly weaker in analytic power than conscious cognition—a *subordinate unconscious*. The subordinate unconscious assumption is described later after introducing a multilevel interpretation of human representational abilities. In the multilevel view, with its subordinate unconscious, unconscious cognition continues to play an important role in cognitive defense, but that role is served by mechanisms that are much weaker in analytic power than is the coordinate unconscious of psychoanalytically inspired theories of cognitive defense.

THE NONPARADOX OF KNOWLEDGE AVOIDANCE

This section explains in some detail how cognitive defenses, including ones that appear to involve paradoxical self-deception, can be explained theoretically without paradox when the assumption of a coordinate unconscious is replaced by the assumption of a subordinate unconscious.

Levels of Representation

The cognitive psychological concept of an ordered set of information-processing stages (e.g., Smith, 1968) provides the basis for a multilevel analysis of mental representations. Figure 3.2 shows a minimal levels-of-representation analysis, with just two stages or levels. The first stage produces a relatively crude representation of an experienced event. This initial representation can control some action directly while providing input for a second, higher, level of analysis. The second level, in turn, produces its own representation, which can control a different response to the event.

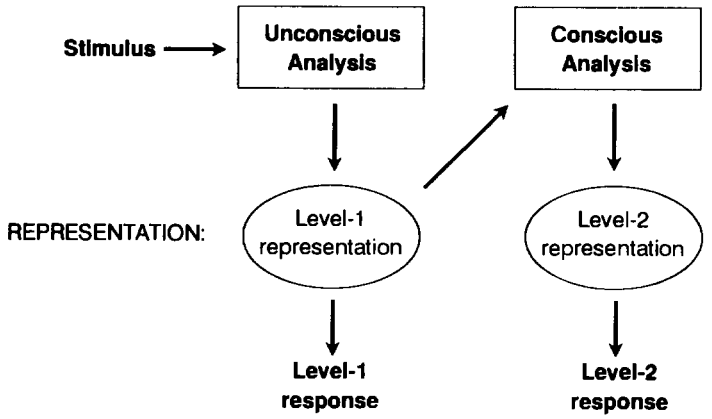


FIG. 3.2. Two-stage levels of representation scheme.

This familiar device of assuming that cognitive analyses occur in series or stages, illustrated minimally with two stages in Fig. 3.2, provides the basis for replacing the paradoxical concept of cognitive defense by self-deception with the nonparadoxical concept of cognitive defense by *knowledge avoidance*. To see how this levels-of-representation analysis avoids paradox, consider an analogy: a two-level model of behavioral (not cognitive) avoidance for the mundane problem of dealing with the contents of one's mailbox.

Junk-Mail Model of Knowledge Avoidance

The annoyance of dealing with unsolicited mass mailings—of material such as advertisements and requests for funds from various organizations—is partly captured by their common designation as "junk" mail. Fortunately, there are easily perceived cues that warn recipients of the likely uninterestingness of an envelope's contents. The postage may be lower than the rate for personal letters, the address likely printed by machine, the recipients's name given in unusual fashion (e.g., to "occupant"), and the envelope made from low-quality paper. Certainly many people have the habit of discarding, without opening, envelopes that provide such warnings. This is a useful avoidance response—one saves the time required to open and read the undesired contents of such mail.

The two stages of the junk-mail model (see Fig. 3.3) are (a) examining the exterior of the envelope, and (b) reading the contents. It is clear that the second stage's processing can be avoided by using results from the first stage's analysis. In other words, one need not know specifically what is inside the envelope to judge that it should be discarded.

In order to connect junk mail to self-deception, let us return to the

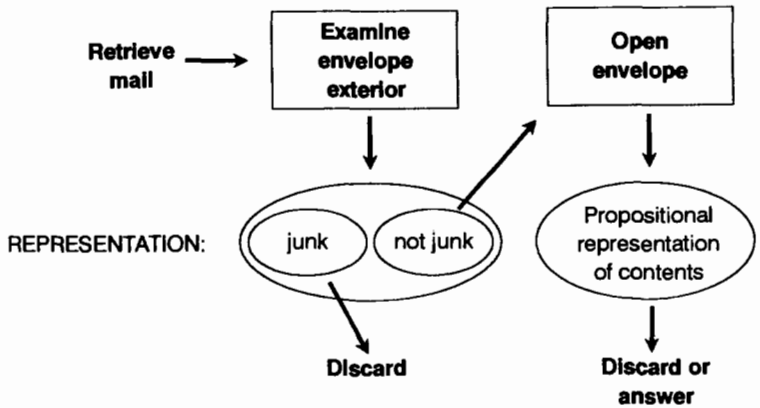


FIG. 3.3. Junk-mail version of the two-stage model.

example of the terminally ill cancer patient. Perhaps the patient picks up cues indicating that *some* unwelcome knowledge may be available (like seeing the outside of the envelope), and then avoids learning precisely what the unwelcome knowledge is (like discarding the letter). There is no more paradox in the cancer patient's avoiding sure knowledge of terminal illness than there is in the junk-mail recipient's avoiding sure knowledge of the contents of an unopened envelope. This analysis, which proposes that an avoidance response can be based on partial analysis of a stimulus, is a close relative of ones offered previously by Allport (1955), Eriksen and Browne (1956), and Kempler and Wiener (1963), in their reviews of research on perceptual defense.

Nonparadoxical Account of the Voice-Recognition Experiment

Figure 3.4 analyzes Gur and Sackeim's (1979) voice-recognition procedure in terms of levels of representation. In the two-level model of Fig. 3.4, the SCR is controlled by the first level, which analyzes the acoustic features of a voice sample. The SCR may be elicited by voice-spectrum features that resemble one's own voice. This sensory-feature-based SCR is *not* equivalent to voice identification any more than examining the outside of an envelope is equivalent to reading its contents. Voice identification occurs only at the second stage of analysis, perhaps based on additional, more complex (paralinguistic) cues, such as accent, speech rate, and inflection. As was the case for the two levels of the junk-mail model, the two levels of the voice-recognition model involve different types of analysis. The second stage requires more complex analysis than

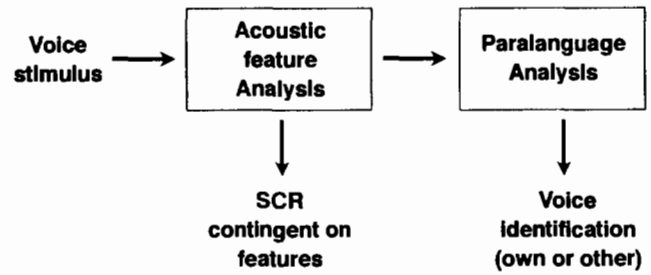


FIG. 3.4. Two-stage model of the voice-recognition experiment.

the first, and it produces a more abstract representation. Because the first-stage SCR and the second-stage self-identification can be based on different stimulus information, there is no paradox when a first-stage SCR is accompanied by nonidentification of own voice at the second stage.

Levels of Representation Elaborated

Figure 3.5 expands the two-level model of Fig. 3.2 into a four-level structure that is rich enough to account for a broad variety of human cognitive capabilities (based on Greenwald & Leavitt, 1984; elaborated further in Greenwald, 1988a). At the lowest level is sensory-feature analysis—a process that is assumed to operate automatically and without leaving memory traces—that is, unconsciously. In the model in Fig. 3.5, unconscious cognition is identified with this first (lowest) level, which does not produce representations more abstract than sensory features. The placement of a dividing line between unconscious and conscious cognition within the series of levels of analysis makes this model one of a subordinate, rather than a coordinate, unconscious. The second level identifies objects and accesses word meanings. The third level encodes verbal information into propositional representations (i.e., sentence meanings). The fourth and highest level uses stored conceptual knowledge to generate inferences from the third level's propositional representations.[2,3]

[2]In the more detailed development of this analysis (Greenwald, 1988a), the second level is split into two functions—object identification and categorization—that can be treated as separate levels.

[3]No attempt has been made in this chapter to relate the hypothesis of a series of cognitive stages of analysis, as in Fig. 3.4, to theorization concerning neural apparatus that could support such function. However, modern theorization concerning the orienting reflex (e.g., in the tradition of Sokolov, 1963) provides a conception of central nervous system organization that is quite compatible with the cognitive distinction between preattentional (unconscious) and attentional (conscious) levels.

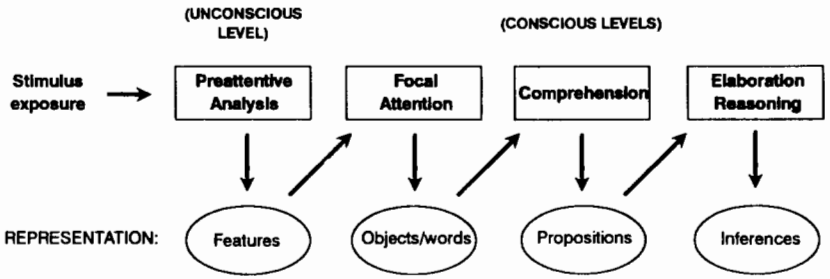


FIG. 3.5. Four-level model of cognition.

"Self-Deception" as Avoidance of Inference

In the example of unawareness of terminal cancer, critical processing occurred at Figure 3.5's third level—the level at which events are analyzed in terms of propositions such as "The doctor said they removed all of the tumor." After that processing occurs, it is still necessary to use the fourth level—the level of reasoning from conceptual knowledge—to draw inferences such as "The doctor didn't tell me to expect complete recovery. That means my chances aren't so good." By not going beyond the third level—by not drawing inferences—one *avoids* an unwelcome conclusion. That is, cognition does not proceed to the level of identifying the threat specifically—the unwelcome news remains unknown. This avoidance of a painful conclusion differs fundamentally from the paradoxical self-deception of Fig. 3.1. In the self-deception analysis, the painful conclusion is simultaneously known (unconsciously) and not known (consciously). By contrast, in the avoidance analysis, the painful conclusion is not known at either a conscious or unconscious level.

In addition to providing an account of apparent self-deception associated with terminal illness, the cognitive-defense-by-avoidance-of-inference analysis applies to daily acts of avoidance that everyone must engage in routinely. Every time we hear news of personal risks—such as diseases associated with foods we eat (such as caffeine or eggs), or accidents associated with behavior in which we engage (flying or driving), or of possible local hazards (such as crime, earthquake, flood, or toxic waste spill), or even of global threats (such as terrorism, ozone depletion, or catastrophic nuclear reactor failure)—we can infer that our well being is threatened. However, most of us spend little time contemplating such threats. The analysis of this avoidance is quite similar to that of the terminally ill cancer patient. Our avoiding the unwelcome conclusion that we will fall victim to one of numerous threats may be credited to habits of interrupting trains of thought that lead to unpleasant conclusions (cf. Dollard & Miller, 1950). Of course, not all people avoid drawing frightening conclusions about their personal vulnerability. But, then, neither do all terminally ill patients avoid drawing the conclusion that they are dying.

Avoidance of Inference by Drawing Alternative Inference

Consider a possible example of cognitive defense that received much publicity in 1991, when Clarence Thomas was nominated to the position of Associate Justice of the United States Supreme Court. During Thomas's U.S. Senate confirmation hearing, law professor Anita Hill accused Thomas of sexual harrassment when Thomas had been her supervisor in a previous job. Thomas vigorously denied the accusation. The accusation and denial were such that it appeared that one of the two must be lying. A possibility that was not considered in the news and commentary accounts of the time is that both could have been telling what they perceived to be the truth! This would be possible if Hill and Thomas drew different inferences or conclusions from their participation in the same situation. What Hill concluded to be sexual harrassment might have been interpreted quite differently by Thomas, perhaps as an unsuccessful attempt to establish a friendly relationship with a colleague.

It has long been considered normal for different participants in a social interaction to draw different inferences (or make different attributions) about the interaction. Two well-researched sources of systematic variation in these attributions are associated with the differing perspectives of *actor* and *observer:* (a) Actors tend to interpret their actions as being responsive to events occurring in the situation, whereas observers are more likely to interpret the same behavior as indicating some characteristic personality trait of the actor (Jones & Nisbett, 1972); and (b) actors are likely to interpret and remember their own actions in a self-serving or self-enhancing way (Greenwald, 1980). The differences between Clarence Thomas's and Anita Hill's interpretations of their interaction could be a case of Thomas's drawing an inference that was self-servingly different from Hill's following interactions that they viewed and interpreted from different perspectives. As in the case of the terminal cancer patient, this can be a cognitive defense that is explainable in nonparadoxical fashion by the sequential-stage levels-of-representation model.

THE ORDINARINESS OF KNOWLEDGE AVOIDANCE

The illustrations of avoiding knowledge of terminal illness and avoiding anxiety about various risks and environmental hazards have been interpreted as cognitive defenses that occur between the third and fourth levels of the model in Fig. 3.5. Experienced events are analyzed to the level of propositions (such as "Amazon rain forests are being cut down"), but anxiety-producing inferences are avoided. Failure to draw such inferences may be the phenomenon that has most frequently been identified, in previous analyses, as involving (paradoxical) self-deception. In contrast with the present analysis, those previous analyses have assumed that the inference must be made at an unconscious level at the same time that it is avoided consciously. The model in Fig. 3.5, however, provides no mechanism for achieving inferences unconsciously, and does not require the occurrence of such inferences as a condition of successful avoidance.

The model in Fig. 3.5 allows knowledge avoidance to occur not only at the transition from its third to fourth level, but also in its lower level transitions. The following consideration of these possibilities suggests that cognitive defense by knowledge avoidance is a pervasively ordinary phenomenon.

Avoidance of Comprehension

Avoidance of third-level processing would occur if the words in a message were perceived individually (second-level processing), but the perceiver avoided comprehending their sentence-level meaning. Such avoidance of comprehension occurs commonly in dealing with the content of mass media. Television and radio programs are frequently interrupted by short commercial announcements in which one is uninterested; newspapers and magazines contain advertisements and uninteresting articles interspersed among their more interesting contents. In dealing with mass media, the perceiver may be consciously aware of the individual words of a message while nevertheless avoiding comprehension of their sentence-level meaning. Hearing or seeing a brand name may suffice to classify the surrounding message as uninteresting, which in turn leads to diverting attention elsewhere, thereby avoiding the effort of comprehending that message. Similarly, the title of a magazine article or the headline of a newspaper story may contain a name or topic word that is sufficient to forestall further analysis. Avoided comprehension after perceiving individual words may be what is happening when one reacts to another's "unattended" remark by asking for it to be repeated, but then readily retrieves the individual words in sequence and does the higher

level work of comprehension before the remark before is actually restated.

Avoidance of Attention

Treisman and Gelade (1980) described the cognitive act of attention as involving the integration of sensory features into perceived objects. Avoiding attention can therefore occur when first-level analysis of sensory features is not followed by further perceptual analysis. The well-known "cocktail-party effect"—being able to focus on a single one of several simultaneously heard voices—is an example of avoiding attention. The listener successfully avoids attending to the words of extraneous conversations while nevertheless analyzing their sensory features, such as voice pitch and spatial location (Broadbent, 1958; Moray, 1970). A second example is an experience, familiar to most skilled automobile drivers, that occurs when, immediately after completing some portion of a familiar route, one is unable to recall stimuli that must have been processed recently, such as whether the last traffic light was red or green. In this case, it is not so much that perceiving the object is undesired as that it is unnecessary. For experienced drivers, driving is so well learned that it can be performed automatically, with habitual actions occurring in response to important stimulus features (i.e., after analysis only at the lowest level of Fig. 3.5), leaving those features unintegrated into perceptually attended objects.

Avoidance of Exposure

Perhaps the most common type of knowledge avoidance is one that cannot be located between stages of the levels-of-representation model because it involves complete nonexposure to stimuli that might lead to useless or otherwise unwelcome cognitive analyses. For example, consider the consequence of a heavy smoker not engaging in physical exercise. The smoker thus avoids encounters with stimuli (excessive fatigue, difficulty breathing, etc.) that could indicate adverse physical effects of smoking. In a similar fashion, by soliciting no student evaluations, a professor can avoid negative feedback that would injure self-esteem. And, to take an almost trivial example, many recreational tennis players effectively avoid discovering that they routinely commit the error of foot-faulting (i.e., stepping into the playing area before hitting a serve) because, consistent with good tennis form, they simply do not look at their feet while serving. Because there is no exposure to events that could lead to unwelcome knowledge, the avoiding-exposure strategy can be very effective. The avoider has no basis for suspecting that any-

thing is being avoided. In terms of the four-level model, such behavioral avoidance of exposure to unwelcome stimuli preempts the first level of analysis. It is located off the left side of the model. The junk-mail model of Fig. 3.3 is itself an example of behavioral avoidance of exposure.

TWO NEW THEORETICAL DEVELOPMENTS

Neural Network Modeling

Since publication of the earlier version of this chapter, there has been active development of a new paradigmatic approach in psychology, alternatively labeled *parallel distributed processing*, *connectionism*, or *neural network modeling* (Rumelhart & McClelland, 1986). Figure 3.6 gives a schematic representation of this approach, indicating how it accommodates the distinction between conscious and unconscious cognition.

The extensive parallelism of the network model of Fig. 3.6 is its fundamental difference from the sequential-stage information-processing model of Fig. 3.5. The network model in Fig. 3.6 incorporates representations of two forms of conscious cognition.[4] One of these—conscious cognition as network operation that boosts activation to resonantly stable high levels in subnetworks—corresponds to an interpretation of *conscious cognition as a focus of attention* on some thought or percept. The network's second representation of conscious cognition is its possibility of having verbal outputs that, by virtue of their connections to inner nodes ("hidden units") of the network, are able to report (in some sense) on internal network status. These verbal outputs correspond to an interpretation of *conscious cognition as a capacity for introspective report* (or "self-consciousness"). Importantly, the structural principles of the neural network provide no assurance that these verbal reports will provide valid descriptions of network status.

In the information-processing form of theory (Fig. 3.5), cognitive defenses were explained nonparadoxically by supposing a hierarchical division between conscious and unconscious cognition, and assuming that unconscious cognition was associated with hierarchically lower and less complex levels of analysis. The neural network account avoids paradox by the more radical device of abandoning the assumption of personal unity. To illustrate: In a neural network account of the voice-recognition experiment (for which an information-processing account was portrayed in Fig. 3.4), the (nonverbal) SCR and the (verbal) voice-identification response would be treated as outputs from analyses that occur in parallel. Because these two analyses might take independent paths through the network, they need be under no constraint to yield correlated outputs.

[4]The distinction between these two forms of unconscious cognition is developed more fully in Greenwald (1992).

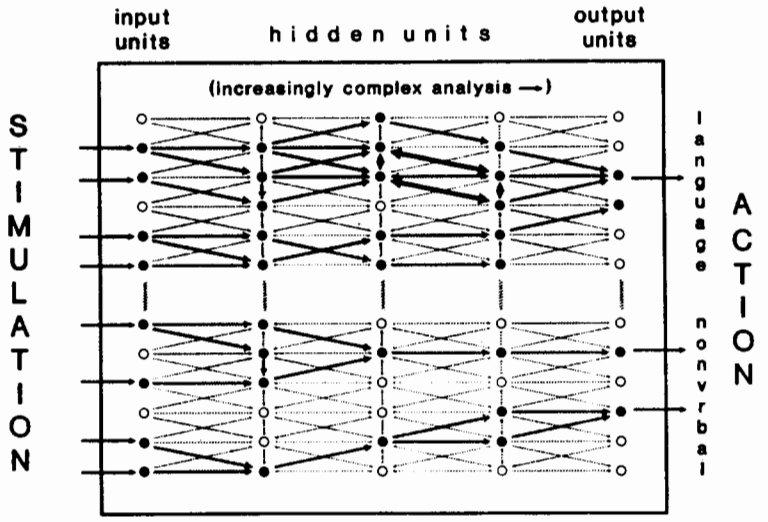


FIG. 3.6. Representation of conscious and unconscious cognition in the format of neural network (connectionist or parallel distributed processing) models. (This figure is duplicated with permission from Fig. 2 of Greenwald, 1992.)

Implicit Cognition

In the early 1980s, cognitive psychologists began to investigate a large class of phenomena inspired by Edouard Claparède's (1911/1951) observation of a surprising manifestation of memory in a Korsakoff-syndrome patient. Characteristically for the illness, Claparède's patient lacked ordinary recollection, and was unable to recognize Claparède from one visit to the next. During one visit, Claparède deliberately pricked the patient's finger with a hatpin when they were shaking hands. On the next visit, the patient hesitated to shake hands with Claparède—whom, as usual, the patient did not recognize as a familiar acquaintance.

Jacoby and colleagues (Jacoby & Dallas, 1981; Jacoby & Witherspoon, 1982) have reported experimental tests that established the repeatability of observations (like Claparède's) of "remembering without awareness," not only with Korsakoff-syndrome patients, but also with normal undergraduate students. In Jacoby's research, unrecallable events have

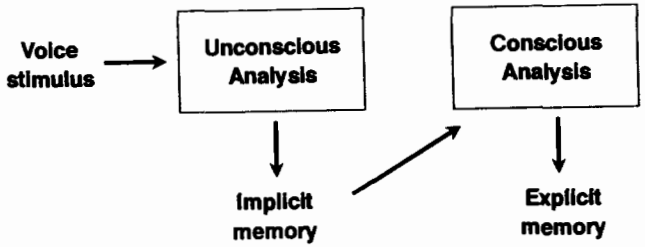


FIG. 3.7. Two-stage analysis of implicit cognition.

been shown to potently influence judgments made in response to stimuli that re-present some portion of the earlier event. A familiar example is that subjects will complete a word stem (e.g., *can___*) with a word that was presented earlier in the experiment (e.g., *candle*), even when they are unable to recall that *candle* was presented earlier.[5] Schacter (1987) reviewed the rapidly growing literature on such memory phenomena, and Greenwald and Banaji (1995) established that parallel forms of *implicit cognition* also occur pervasively in expressions of social attitudes, stereotypes, and self-esteem.

The defining characteristic of implicit cognition is that some judgment draws on information contained in past experience while the judge nevertheless remains ignorant of the influence of that past experience (see Greenwald & Banaji, 1995). This generic definition of *implicit cognition* encompasses the situation in which a skin conductance response to a playback of the subject's voice indicates that the subject is responding under the influence of past experience with the sound of own voice, even while not recognizing this influence. Not surprisingly, then, implicit cognition can be analyzed with a stage model of the same type used for the voice-recognition experiment (see Fig. 3.7).

EMPIRICAL ANALYSIS OF THE RELATIONSHIP BETWEEN CONSCIOUS AND UNCONSCIOUS COGNITION

Of the several procedures that have been used in attempted laboratory models of psychoanalytically conceived defense mechanisms, none has escaped criticism that the resulting evidence is inconclusive. For every sympathetic review of findings on topics such as perceptual defense (Dixon, 1981; Erdelyi, 1974) or repression (Erdelyi & Goldberg, 1979; Shevrin & Dickman, 1980), one can point to forceful opposing reviews (e.g., Eriksen, 1958; Holmes, 1974; Loftus & Loftus, 1980). In drawing conclusions from a review of the research literature, Erdelyi (1985) observed that a consistent shortcoming of laboratory models of cognitive defense was their failure to demonstrate that "the perceiver can intentionally and selectively reject perceptual inputs [of emotional stimuli]" (p. 256). Sackeim and Gur (1978) presented the following list of four criteria for an empirical demonstration of (paradoxical) self-deception:

1. The individual holds two contradictory beliefs (*p* and not-*p*).
2. These beliefs are held simultaneously.
3. The individual is not aware of holding one of the beliefs (*p* or not-*p*).
4. The act that determines which belief is and which belief is not subject to awareness is a motivated act. (p. 150)

The levels-of-representation analysis of knowledge avoidance (Fig. 3.3, 3.4, and 3.5) and the neural network model (Fig. 3.6) provide perspectives from which the difficulty of achieving research demonstrations that meet these four criteria is easily understood. In the levels-of-representation approach, because cognitive avoidance does not involve knowledge of what is being avoided, Sackeim and Gur's first, second, and fourth criteria are inappropriate. That is, the expectation of data patterns that match Sackeim and Gur's four criteria is contingent on the assumption of a coordinate unconscious—an assumption that seems quite unnecessary for the explanation of successful cognitive defense.

In the neural network approach, it is possible for conscious and unconscious cognition to be as independent of one another as if they were taking place in the left and right cerebral hemispheres of a split-brain patient. This possibility of independent, parallel cognitive paths bears on the reference to an individual (implying an indivisible entity) in Sackeim and Gur's first and third empirical criteria. In the neural network, there is no necessary assumption of personal unity, meaning that two mutually contradictory beliefs could be represented in the network without there being any system locus that has simultaneous access to them. This makes it possible for the neural network to account nonparadoxically for even exotic dissociations that could not be explained by information-processing models such as that in Fig. 3.5.[6]

[5]The reader who completed this stem with *cancer* may have experienced the same kind of automatic influence routinely experienced by subjects in Jacoby's experiments.

[6]Greenwald (1992) described a variant information-processing approach that assumed the possibility of independent processing paths, each composed of sequential stages. (Information-processing models standardly assume only a single series of stages or levels.) When modified to permit multiple parallel paths, the information-processing approach may be paradigmatically indistinguishable from the neural network approach.

Recent empirical and theoretical developments have not only made cognitive psychologists much more comfortable with the idea of unconscious cognition than they were a decade ago, but have also shaped a view in which unconscious cognition is seen as operating alongside, and sometimes even independently of, conscious cognition. The theoretical development of neural network modeling has made the idea of parallel conscious and unconscious cognition decidedly nonparadoxical. In recent years, researchers have increasingly employed assumptions of independence (or dissociation) between conscious and unconscious cognition in their interpretations of experimental findings (e.g., Greenwald, Klinger, & Schuh, 1995; Jacoby, Lindsay, & Toth, 1992). These dissociations can be given nonparadoxical interpretation by either (or both) a levels-of-representation or a network approach.

In summary, Sackeim and Gur's (1978) list of criteria for an empirical demonstration of self-deception implies two assumptions that appear unneeded and excessive in light of cognitive psychological research of the past few decades. These two questionable assumptions are (a) *unconscious knowledge of threat*—that successful cognitive defense requires a prior, complete, unconscious representation of the knowledge that is being defended against; and (b) *single-agency coordination*—that successful cognitive defense represents the coordinated achievement of a single agency. In combination, these two assumptions comprise the view that was described earlier in this chapter as the "coordinate unconscious" conception. The knowledge-of-threat assumption is easily sacrificed in the context of a levels-of-representation view, in which low-level, partial analysis of a threatening stimulus allows avoidance or modification of a later, more complex analysis. The single-agency assumption survives in levels-of-representation models, but is unneeded in neural network models, which move away from a conception of personal unity and toward understanding the person as a distributed processor with multiple, concurrent cognitive processes.

CONCLUSION: A VIEW OF UNCONSCIOUS COGNITION

This chapter has described a view of unconscious cognition that has evolved, in the last few decades, from a previously dominant view that rested on the psychoanalytic conception of coordinate conscious and unconscious cognition. The psychoanalytic view obliged the assumption that cognitive defense could involve paradoxical simultaneous knowledge and ignorance of threatening or anxiety-arousing situations. The assumed validity of this paradoxical psychoanalytic interpretation justified referring to some cognitive defenses as self-deceptions.

Contemporary cognitive psychology provides two theoretical interpretations of the relationship between conscious and unconscious cognition that provide nonparadoxical accounts of a wide variety of cognitive defenses, including those that have been labeled *self-deceptions*. The longer established of these interpretations is the concept of a hierarchical series of stages of information processing (or levels of representation). In this sequential-stage view, a low-level analysis can guide both the avoidance of threat and the avoidance of higher level processing that is needed to identify the exact nature of the threat. The more recently developed approach of neural network modeling, by accommodating independent paths of cognition initiated by the same stimulus, is theoretically powerful enough to provide nonparadoxical explanations of even exotic cognitive dissociations, such as multiple personality. One attraction of the sequential-stage information-processing view is its ability to provide a nonparadoxical account of cognitive defense, while accommodating the lay conception of unity of the normal personality. In the context of the neural network approach, personal unity may be treated merely as an illusion of the normal personality.

These new interpretations of unconscious cognition are important not only because they demystify phenomena previously considered to be paradoxical self-deceptions, but because they portray unconscious cognition as relatively weak in its cognitive analytic power. The conception of unconscious cognition as cognitively weak appears especially in the sequential information-processing interpretation, which associated unconscious cognition with early (and relatively crude) stages of processing. This conception of unconscious cognition as relatively weak in analytic power was implied by the junk-mail model, which was first described in the 1988 version of this chapter. The case for regarding unconscious cognition as weak in analytic power was developed in much more detail by Greenwald (1992).

ACKNOWLEDGMENTS

This chapter is a major revision and update of the author's contribution to the volume, *Self-Deception: An Adaptive Mechanism?*, edited by Joan S. Lockard and Delroy L. Paulhus (Englewood Cliffs, NJ: Prentice-Hall, 1988 [Greenwald, 1988a]). Preparation of the original chapter was aided by a grant from National Science Foundation (BNS-8217006). Preparation of this revision was aided by grants from National Science Foundation (SBR-9422242) and National Institute of Mental Health (MH-41328). The

author thanks Joan S. Lockard, Delroy L. Paulhus, and Prentice-Hall for permission to use portions of the earlier chapter.

REFERENCES

Allport, F. H. (1955). *Theories of perception and the concept of structure.* New York: Wiley.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon.

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

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

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy.* New York: McGraw-Hill.

Erdelyi, M. H. (1974). A new look at the new look: Perceptual defense and vigilance. *Psychological Review, 81,* 1–25.

Erdelyi, M. H. (1985). *Psychoanalysis: Freud's cognitive psychology.* New York: Freeman.

Erdelyi, M. H., & Goldberg, B. (1979). Let's not sweep repression under the rug: Toward a cognitive psychology of repression. In J. F. Kihlstrom & F. Evans (Eds.), *Functional disorders of memory* (pp. 355–402). Hillsdale, NJ: Lawrence Erlbaum Associates.

Eriksen, C. W. (1958). Unconscious processes. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958* (pp. 169–227). Lincoln: University of Nebraska Press.

Eriksen, C. W., & Browne, C. T. (1956). An experimental and theoretical analysis of perceptual defense. *Journal of Abnormal and Social Psychology, 52,* 224–230.

Fingarette, H. (1969). *Self-deception.* London: Routledge & Kegan Paul.

Freud, S. (1961). The ego and the id. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 3–66). London: Hogarth. (Original work published in 1923)

Gazzaniga, M. S. (1985). *The social brain: Discovering the networks of the mind.* New York: Basic Books.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35,* 603–618.

Greenwald, A. G. (1982). Is anyone in charge? Personalysis versus the principle of personal unity. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1, pp. 151–181). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenwald, A. G. (1988a). *Levels of representation.* Unpublished manuscript, University of Washington, Department of Psychology, Seattle.

Greenwald, A. G. (1988b). Self-knowledge and self-deception. In J. S. Lockard & D. L. Paulhus (Eds.), *Self-deception: An adaptive mechanism?* (pp. 113–131). Englewood Cliffs, NJ: Prentice-Hall.

Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *American Psychologist, 47,* 766–779.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General, 124,* 22–42.

Greenwald, A. G., & Leavitt, C. (1984). Audience involvement in advertising: Four levels. *Journal of Consumer Research, 11,* 581–592.

Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology, 37,* 147–169.

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

Holmes, D. S. (1974). Investigations of repression: Differential recall of material experimentally or naturally associated with ego threat. *Psychological Bulletin, 81,* 632–653.

Huntley, C. W. (1940). Judgments of self based upon records of expressive behavior. *Journal of Abnormal and Social Psychology, 35,* 398–427.

Kempler, B., & Wiener, M. (1963). Personality and perception in the recognition threshold paradigm. *Psychological Review, 70,* 349–356.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306–340.

Jacoby, L. L., Lindsay, D. S., & Toth, J. P. (1992). Unconscious influences revealed: Attention, awareness, and control. *American Psychologist, 47,* 802–809.

Jacoby, L. L., & Witherspoon, D. (1982). Remembering without awareness. *Canadian Journal of Psychology, 36,* 300–324.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist, 39,* 124–129.

Lockard, J. S. (1980). Speculations on the adaptive significance of self-deception. In J. S. Lockard (Ed.), *The evolution of social behavior* (pp. 257–275). New York: Elsevier.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist, 35,* 409–420.

Moray, N. (1970). *Attention: Selective processes in vision and hearing.* New York: Academic Press.

Mowrer, O. H. (1960). *Learning theory and behavior.* New York: Wiley.

Murphy, G. (1975). *Outgrowing self-deception.* New York: Basic Books.

Osgood, C. E. (1953). *Method and theory in experimental psychology.* New York: Oxford University Press.

Rumelhart, D. E., & McClelland, J. L. (Eds.) (1986). *Parallel distributed processing* (2 vols.). Cambridge, MA: MIT Press.

Sackeim, H. A., & Gur, R. C. (1978). Self-deception, self-confrontation, and consciousness. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (Vol. 2, pp. 139–197). New York: Plenum.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 65,* 379–399.

Schacter, D. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schafer, R. (1976). *A new language for psychoanalysis.* New Haven, CT: Yale University Press.

Shevrin, H., & Dickman, S. (1980). The psychological unconscious: A necessary assumption for all psychological theory? *American Psychologist, 35,* 421–434.

Smith, E. E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin, 69,* 77–110.

Sokolov, E. N. (1963). *Perception and the conditioned reflex.* New York: Pergamon.

Spence, K. W. (1956). *Behavior theory and conditioning.* New Haven, CT: Yale University Press.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57,* 421–457.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Trivers, R. (1985). *Social evolution.* Menlo Park, CA: Benjamin/Cummings.

Wolff, W. (1932). Selbstbeurteilung and fremdbeurteilung im wissentlichen und unwissentlichen versuch. *Psychologische Forschung, 16,* 251–329.