

Characteristics of dissociable human learning systems

David R. Shanks

Department of Psychology, University College London, London WC1E 6BT, England

Electronic mail: david.shanks@psychol.ucl.ac.uk

Mark F. St. John

Department of Cognitive Science, University of California at San Diego, La Jolla, CA 92093

Electronic mail: mstjohn@cogsci.ucsd.edu

Abstract: A number of ways of taxonomizing human learning have been proposed. We examine the evidence for one such proposal, namely, that there exist independent explicit and implicit learning systems. This combines two further distinctions, (1) between learning that takes place with versus without concurrent awareness, and (2) between learning that involves the encoding of instances (or fragments) versus the induction of abstract rules or hypotheses. Implicit learning is assumed to involve unconscious rule learning. We examine the evidence for implicit learning derived from subliminal learning, conditioning, artificial grammar learning, instrumental learning, and reaction times in sequence learning. We conclude that unconscious learning has not been satisfactorily established in any of these areas. The assumption that learning in some of these tasks (e.g., artificial grammar learning) is predominantly based on rule abstraction is questionable. When subjects cannot report the “implicitly learned” rules that govern stimulus selection, this is often because their knowledge consists of instances or fragments of the training stimuli rather than rules. In contrast to the distinction between conscious and unconscious learning, the distinction between instance and rule learning is a sound and meaningful way of taxonomizing human learning. We discuss various computational models of these two forms of learning.

Keywords: artificial grammar; categorization; connectionism; consciousness; explicit/implicit processes; instances; learning; memory; rules

1. Introduction

A classic issue faced by researchers attempting to understand the basic laws of learning is whether there is more than one basic learning mechanism. Can all the phenomena of learning be accommodated by a unitary mechanism, or do we need to posit the existence of independent and dissociable human learning systems? In this target article we consider some of the experimental evidence – much of it very recent – that has addressed this issue.

We will consider two dimensions on which it has been suggested that functionally distinct learning systems differ. The first dimension concerns the role of awareness during learning. Many authors (e.g., Hayes & Broadbent 1988; Lewicki et al. 1987; Reber 1989a) have argued that in addition to having a learning system whose functioning is accompanied by concurrent awareness of what is being learned, humans have a quite separate system that operates independently of awareness. The second dimension, which turns out to be closely related to the first, concerns the content of learning. Distinct learning systems encode very different sorts of information; one system induces rules (e.g., Lea & Simon 1979; Nosofsky et al. 1989), whereas a second system memorizes instances (e.g., Brooks 1978; Medin & Schaffer 1978).

We believe it is important to evaluate the current evidence for and against the multiple-systems view for at least two reasons. First, each of the separate systems that

has been hypothesized has tended to encourage researchers to develop a set of explanatory constructs that are unique to that system and that allow its characteristic phenomena to be explained. A drawback, however, is that experimental results are often interpreted exclusively in terms of these restricted concepts, with no consideration of whether they might also be understood (and possibly better understood) in terms of more general principles.

The second and perhaps more pressing reason for evaluating the evidence for dissociable learning systems is that there has been considerable interest, over the last few years, in whether there exist dissociable *memory* systems (for reviews, see Richardson-Klavehn & Bjork 1988; Schacter 1987; 1989; Squire 1992). The mounting positive evidence comes from a variety of sources. For example, amnesic patients have been shown to be dramatically impaired on certain direct tests of memory, such as free recall, but less impaired or even unimpaired on indirect tests of memory, such as motor skills (see Squire 1992). Although dissociations between performance on direct and indirect tests do not force us to conclude that there are dissociable memory systems (e.g., Jacoby & Kelley 1991; Roediger 1990), some researchers have argued at length that the experimental results, together with current understanding of brain functioning, strongly imply the existence of separable underlying systems (e.g., Schacter 1989; Squire 1992).

Few would argue that learning and memory can be

studied independently. On the contrary, the possible characteristics of dissociable learning systems should be considered in research on the issue of dissociable memory systems and vice versa. Indeed, if there really are dissociable memory systems, it seems very likely that there are also dissociable learning systems that supply them with information. Yet, as several authors have noted (e.g., Berry & Dienes 1991; Reber 1989a), research on learning and on memory has tended to proceed independently. We hope to help memory researchers in their attempts to understand information storage and retrieval by examining carefully the question of whether distinct learning systems exist and by analyzing the properties of the learning mechanisms that acquire information.

1.1. Proposed distinctions between types of learning

Distinctions between different types of learning have been common in psychology for many years. One such distinction is between declarative and procedural learning, that is, between the acquisition of factual knowledge and the acquisition of skills, respectively (e.g., Cohen & Squire 1980; Morris 1984; Winograd 1975). Other distinctions include the acquisition of "habits" versus "memories" (Mishkin et al. 1984) and "taxon" versus "locale" learning (O'Keefe & Nadel 1978). Of course, if independent memory systems require independent learning mechanisms, then many more distinctions might be needed. For instance, we might require separate learning systems to feed semantic and episodic memory stores (Neely 1989; Tulving 1983; see also multiple book review: *BBS* 7(2) 1984).

Of these distinctions, the one between declarative and procedural learning has probably attracted the most attention, with a variety of empirical phenomena being interpreted in that framework. For example, Cohen and Squire (1980) suggested that amnesics have normal or near-normal procedural learning but impaired declarative learning, a theoretical notion that has been widely taken up by other researchers in the amnesia field. This distinction has in recent years been largely eclipsed, however, by the alternative distinction between "explicit" and "implicit" learning. (Note that some authors have replaced the original declarative/procedural distinction with the terms "declarative" and "nondeclarative" [e.g., Shimamura & Squire 1989; Squire 1992].) The main reason for the shift in terminology and emphasis toward the terms "explicit" and "implicit" is dissatisfaction with the original terminology, the term "procedural" apparently being too narrow to encompass the relevant learning effects. For example, the learning that is preserved in amnesia is not always of a procedural nature: it includes a variety of priming effects involving, for instance, the ability to complete word stems (Graf et al. 1984) and an increase in the likelihood of judging a nonfamous name famous as a result of prior exposure (e.g., Squire & McKee 1992).

The term "implicit learning" was first coined by Reber (1967), who is responsible for much of the recent interest in the issue of distinct learning systems (see Reber 1989a for a review). Different authors have used a variety of definitions to capture the fine detail of the explicit/implicit learning distinction (see Mathews et al. 1989, for examples), but the key factor is the idea that implicit

learning occurs without concurrent awareness of what is being learned and represents a separate system from the one that operates in more typical learning situations, where learning does proceed with concurrent awareness (i.e., explicitly). At the same time, it is clear that many authors have been concerned with the possibility that different learning tasks might give rise to different kinds of knowledge (e.g., Mathews et al. 1989; Reber 1989a; Vokey & Brooks 1992), one kind abstract or rule based and the other based on separate fragments or instances. For Reber, implicit learning is not only unconscious but also involves the acquisition of abstract information.

The paradigm case is language learning, where people are assumed to be able implicitly to learn abstract grammatical rules. Few nonlinguists are aware of or are able to articulate the grammatical rules supposed to underlie their linguistic performance, so it makes sense to imagine that those rules are acquired, if at all, without ever being directly represented in consciousness. The rules are abstract in the sense that they apply equally to any linguistic tokens, including novel ones, that come from the appropriate syntactic categories.

Because the aware/unaware and rules/instances dimensions are logically distinct, we believe that they must be treated independently, in this target article we accordingly review evidence for these two dimensions separately. In what follows, we reserve the term "unconscious learning" for learning without awareness, regardless of what sort of knowledge is being acquired. At the same time, we use the terms "rule learning" and "instance learning" to refer to the acquisition of abstract and fragmentary knowledge, respectively, regardless of whether such learning is conscious.

Most of the article is devoted to whether unconscious learning is indeed supported by empirical evidence. In section 2 we survey a wide range of learning paradigms, from subliminal learning phenomena to Pavlovian conditioning to artificial grammar learning and serial reaction-time tasks. The stimuli and specific processes involved in performing and learning each of these tasks differ widely and may share some basic characteristics or may exhibit some basic differences. Across these diverse paradigms we find little actual support for unconscious rule induction (i.e., for implicit learning), or for the unconscious learning of any other type of information. However, in section 3 we do find evidence for a dissociation between a rule-induction system and an instance-memorization system; we review evidence for this dissociation obtained in explicit, or conscious, learning tasks. Within each system, the range of different processes and information is still large, but they nevertheless seem to form two distinct types: slow, effortful hypothesis testing on the one hand, and fast, efficient memorization of instances and fragments of instances on the other.

We concentrate throughout on data from normal subjects. It is clear, however, that amnesic patients have learning difficulties, and these difficulties have been widely interpreted within the explicit/implicit framework (e.g., Squire 1992). For our present purposes, the data from such subjects are tangential, because the question of awareness during learning has not been directly considered in amnesics (but see Knopman 1991). In section 4 we comment briefly on the interpretation of learning data from this population of subjects.

2. Can learning occur without awareness?

Proponents of the explicit/implicit distinction have argued that there are clear demonstrations of subjects' ability to encode new information without being aware of that information, and hence that awareness is the key dimension on which separable learning systems differ. The question of whether learning can occur without awareness goes back many decades (e.g., Adams 1957; Dulany 1961; Eriksen 1960; Krasner 1958; Thorndike & Rock 1934). In addition to the recent work of Reber, which we consider below, in the last five or six years there have been a large number of *sequence learning* reaction time studies that have adopted an interesting and novel technique for assessing the relationship between awareness and learning. A substantial part of our review concerns results obtained using this task. We also consider evidence from a variety of conditioning procedures. We begin with some comments on experimental methodology.

2.1. The logic of dissociations

Almost all studies of unconscious learning have adopted a very constrained version of the logic of dissociation. Separate indices of learning and awareness are used in the attempt to find circumstances in which exposure to a set of stimuli leads to detectable learning unaccompanied by any reliable degree of awareness. On the face of it, such an approach could lead to unequivocal evidence of unconscious learning, but researchers using similar logic to try to establish the existence of unconscious perception have noted several problems (e.g., Reingold & Merikle 1988). What counts as a suitable test of awareness? Can we discount the possibility that our index of awareness is contaminated by unconscious information? Can we be sure it is sufficiently sensitive to detect exhaustively all conscious information? As we shall see, these are deep problems, and researchers have adopted a variety of strategies to try to circumvent them.

Firmer evidence for unconscious learning may emerge from experiments based on alternatives to this particular dissociation paradigm. To test unconscious perception, for example, Reingold and Merikle (1988) have proposed a new and interesting procedure, whereby one looks for greater sensitivity to some variable in an indirect test in which instructions make no reference to the variable as compared with an otherwise identical direct test in which the instructions do refer to the variable. Alternatively, one could try to demonstrate the independence of two learning systems by trying to establish *qualitative* differences between them (e.g., Merikle & Reingold 1992), such that, for example, one system is affected in one way by a variable, the other in the opposite way.

We know of only one study that has even come close to establishing such qualitative differences; this case will accordingly be considered in some detail. Hayes and Broadbent (1988) began by postulating two independent systems: an unconscious system that would slowly accumulate information about predictive events in the environment and a conscious system that would test hypotheses. They further assumed that the conscious system would be highly dependent on a limited-capacity working memory system, and the unconscious system would be

independent of it. [See also Broadbent: "The Maltese Cross" *BBS* 7(1) 1984.]

A rather straightforward prediction emerges from this plausible model of the cognitive system. Because the conscious learning mechanism relies on working memory, there should be situations where learning is profoundly affected by loading the working memory system with a secondary task, such as generating random numbers. At the same time, because the unconscious system does not depend on working memory, other (implicit) learning tasks should be unaffected by such a secondary task. Indeed, Hayes and Broadbent went so far as to say that unconscious learning might be facilitated by a secondary task if it prevented the conscious system from exerting an interfering influence on the unconscious system. The importance of the Hayes and Broadbent study is that, in accordance with their model, they appeared to have found two learning tasks that differed in only a minor way, one of which was inhibited and the other facilitated by a secondary task.

In their experiments Hayes and Broadbent contrasted performance in two versions of the computer "person" task. On each trial, the subject entered an attitude (e.g., polite) into the computer, which then responded with its attitude (e.g., unfriendly). The subject's task was to try to get the computer to be friendly. If we designate the 12 possible attitudes – going from very unfriendly to loving – with the numbers 1 . . . 12, then the computer's attitude on each trial was a simple numerical function of the subject's input. In one (No-Lag) condition, the computer's attitude (O_t) on each trial was a function of the subject's attitude (I_t) on the same trial:

$$O_t = I_t - 2 + r \quad (1)$$

where r is a random number ($-1, 0, \text{ or } 1$) and the attitudes have the 12 numerical values mentioned above. In the other (Lag) condition, I_t was replaced by I_{t-1} , so that the computer's attitude was determined by the subject's attitude on the preceding trial:

$$O_t = I_{t-1} - 2 + r \quad (2)$$

Performance was measured in terms of the number of trials in which the subject's input was one that could (given the random element) have produced a friendly response from the computer person. Although learning occurred in both groups, Hayes and Broadbent found that subjects could give highly accurate verbal reports about the No-Lag task, indicating that their learning had been accompanied by awareness, whereas the verbal reports of subjects in the Lag version were very poor. This result encourages the view that learning in the No-Lag task can be readily achieved by the explicit system, but that the Lag task requires the implicit system. Thus we might predict that a concurrent secondary task would have an effect on learning in the No-Lag condition but not in the Lag condition.

To test this, Hayes and Broadbent (1988) gave subjects a block of learning trials using either Equation 1 (No-Lag group) or Equation 2 (Lag group). After 30 trials in the No-Lag condition and 50 trials in the Lag condition, performance was approximately equated, and at this point Hayes and Broadbent changed the rules by replacing the -2 in the equations with $+2$. They then presented a further 30 (No-Lag group) or 50 (Lag group) relearning

trials. Under single-task conditions (Experiment 1), performance in the Lag condition was affected more detrimentally than in the No-Lag condition by this rule change. In contrast, when subjects were required to perform a concurrent secondary task (generating random letters or digits; Experiments 2 and 3), a change in the rule interfered more with performance in the No-Lag than in the Lag task, exactly the opposite of the result obtained when there was no secondary task. The results conform to Hayes and Broadbent's theory – and hence to their conception of separate implicit and explicit learning systems – if we simply assume that the secondary task occupied the conscious working memory system and therefore interfered with the explicit system, whereas removal of the working memory system allowed the implicit system to operate without any interfering influence from the explicit system.

Unfortunately, Green and Shanks (1993) were unable to replicate Hayes and Broadbent's results. In the single-task groups, Green and Shanks found that the introduction of the equation change had similar effects on performance in the No-Lag and Lag groups, thus failing to replicate Hayes and Broadbent's (Experiment 1) finding that performance was more detrimentally affected in the Lag condition. Under dual-task conditions the situation was the same: performance was approximately equally affected in the two groups. There was not the slightest hint that performance in the Lag group was less affected by the equation change, and hence Hayes and Broadbent's (Experiment 2 and 3) dual-task results were likewise not replicated. Green and Shanks suggest that Hayes and Broadbent may have obtained the results they did owing to the inappropriate inclusion of subjects who had learned very little prior to the equation change.

Hayes and Broadbent's dissociation posed a genuine problem for theories of learning relying on a single learning mechanism. Because the secondary task appeared to have opposing effects on the two primary tasks, Hayes and Broadbent's data seemed to support the claim that there exist dissociable learning systems. Obviously, the fact that their results could not be replicated undermines those conclusions.

With the exception of Hayes and Broadbent's study, implicit learning experiments have universally adopted the dissociation logic of attempting to demonstrate learning in the absence of any detectable degree of awareness. As we shall see, various methodological problems with the dissociation procedure make it doubtful whether unconscious learning has yet been established. It is worth bearing in mind, however, that future experiments using alternative methods may license stronger inferences concerning the dissociability of learning systems. We now begin our discussion of the empirical evidence.

2.2. Unconscious learning with subliminal stimuli

Most studies of unconscious learning have asked whether people can learn about *relationships* between stimuli without being aware of those relationships, but before discussing the results of such studies we will briefly consider evidence from experiments asking a more direct question, Can people learn about stimuli when they are unaware of the existence of these stimuli, that is, when the stimuli are subliminal? A situation in which uncon-

scious learning would, on the face of it, be fairly straightforward to establish is one in which a subject is entirely unaware that the critical stimulus in the learning phase is present at all, yet still shows evidence of leaning something about that stimulus.

There have, of course, been a large number of experiments in which subjects are presented with brief or low-intensity stimuli intended to be below the threshold of awareness and in which an attempt is made to measure effects of such stimulation on subsequent behavior. We ignore much of this literature, for two reasons: first, in some cases such effects may be only tenuously related to learning. For example, many subliminal activation experiments ask whether the way a stimulus is interpreted may be biased by a supposedly subliminal stimulus presented a few hundred milliseconds previously (e.g., Marcel 1983). It is doubtful, however, that such biasing effects would occur over longer intervals: instead, they are typically interpreted as examples of some sort of short-lived facilitation. Needless to say, it is difficult to draw a sharp line between perception and learning, but if unconscious learning is to have any real significance, it must be demonstrable over reasonable intervals of time (at the very least seconds or minutes rather than milliseconds). Second, many subliminal activation experiments that do appear to show longer-lasting effects (e.g., Eich 1984) have already been the subject of extensive criticism in this journal (see Holender 1986, and accompanying commentaries). We have no wish to repeat arguments made previously except to point out that in such experiments it is extremely difficult to be confident that the stimuli are indeed below the threshold of conscious perception.

We accordingly focus in this section on studies that avoid these problems. Andrade (in press), Bornstein (1992), Ghoneim and Block (1992), Greenwald (1992), and Schacter (1987) review a number of relevant studies examining learning with subliminal stimuli. Although there have been some positive results, a corresponding number of negative findings leads us to suggest that unconscious learning with subliminal stimuli has not yet been conclusively demonstrated.

Subliminal stimuli may be presented to awake subjects as auditory messages at extremely low intensity or in some scrambled form, or as images presented for very brief durations or embedded in other figures; alternatively, they may be presented to subjects during sleep or anesthesia. There is a widespread popular belief in the ability of such subliminal messages to condition attitudes or preferences or otherwise to influence behavior. Indeed, this belief is so powerful that the families of two young men who died from self-inflicted gunshot wounds sought more than \$6 million in damages from the rock group Judas Priest on the grounds that subliminal messages on one of the group's records had caused the men to commit suicide (see Loftus & Klinger 1992). Recent investigations, however, suggest that the concern is misplaced. Controlled experiments attempting to see whether subliminal messages can influence behavior or whether people can use self-help audiotapes as learning aids have yielded exclusively negative results (British Psychological Society 1992; Greenwald et al. 1991; Vokey & Read 1985). It seems unlikely that unconscious learning can occur in such situations.

Several investigations of spared cognitive functions under general anesthesia have obtained evidence of small but reliable amounts of learning, but these are matched by a comparable number of negative results (see Andrade, in press; Ghoneim & Block, 1992, for reviews). If the anesthetic has been adequately administered and renders the patient entirely unconscious, then spared learning must in turn be unconscious. A typical positive result was reported by Jelacic et al. (1992). They gave anesthetized patients repeated auditory presentations of two words (e.g., yellow, green) from a semantic category. Later, when the anesthetic had worn off, subjects were asked in a priming test to generate members of those categories. Subjects were significantly more likely to produce the preexposed words than were control subjects who had not been read the words during anesthesia. Thus some information does seem to have been encoded while the subjects were unconscious.

Another positive result was reported by Kihlstrom et al. (1990). They gave anesthetized patients lists of strongly associated cue-target word pairs, with each list being presented about 67 times during the operation. Later, when the anesthetic had worn off, subjects were given a cued recall and a recognition memory test; in a third test, they were read the cue words and had to say the first word that came to mind. Although the recall and recognition tests yielded no evidence of retention, on the generation test subjects were more likely to produce target items to preexposed cue words than to nonpreexposed cue words, whether the test was relatively soon after the exposure phase (median 87 min) or much later (median 14 days). Thus, again, some degree of unconscious registration seems to have occurred.

In contrast to this are the many negative results that have been published. Some of these are particularly revealing because they come from experiments using procedures very similar to those of studies that have found positive results. For example, Cork et al. (1992) failed to replicate the Kihlstrom et al. (1990) results using a different anesthetic but otherwise identical procedures. Furthermore, despite the likelihood that sleep renders a person less unconscious than general anesthesia, in a well-controlled experiment Wood et al. (1992) were unable to obtain evidence of learning during sleep, again with procedures similar to those used in the Kihlstrom et al. (1990) study. Similarly, Ghoneim et al. (1992) found no evidence of Pavlovian conditioning in anesthetized patients; they used experimental procedures that did reveal conditioning in nonanesthetized subjects.

This pattern of results might simply indicate that learning under anesthesia is a genuine phenomenon, but that relatively subtle methodological factors determine whether a given study will or will not obtain evidence of it. However, Andrade (in press) discusses a large number of studies, including over 20 published reports of failures, and is unable to find any clear factors that determine whether learning will or will not occur. For example, it does not seem to be especially related to the type of stimuli used. More significantly, it remains an open possibility that many positive results have been due to inadequately administered anesthetic that left some or all of the patients at least partially conscious. It is worth noting that in the Cork et al. (1992) study three subjects were excluded from the analysis because they had explicit mem-

ory of the study items! As Cork et al. say, "the extent to which implicit expressions of memory are affected by general anesthesia remains uncertain" (p. 897).

2.2.1. Conclusions. Experiments in which subjects are presented with stimuli that they are likely to be unaware of at the time of exposure yield some evidence of unconscious learning, but this is offset by a substantial body of negative evidence. At present, it would be premature to conclude from the available studies that unconscious learning is feasible.

2.3. Criteria for establishing unconscious learning with supraliminal stimuli

In the rest of this section we focus on situations where the stimuli are above the threshold for detection and identification. In such situations, subjects may be unaware of the relationships between stimuli even though they are aware of the stimuli themselves. Learning of inter-stimulus *relationships* may therefore be unconscious.

We argue that just about all unconscious learning experiments with supraliminal stimuli can be conceptually reduced to the arrangement shown in Figure 1. The figure illustrates an associative learning episode in which subjects have the opportunity to learn that two events, A and B, stand in a predictive relationship. Event A might be a tone conditioned stimulus (CS) and event B a shock unconditioned stimulus (US); the measure of learning might be a galvanic skin response (GSR) at time t_2 when the CS is presented again. Or event A might be a feature or set of features, event B might be a category, and the measure of learning might be the probability of making the category response at t_2 . We are interested in whether subjects can learn the predictive relationship in the absence of concurrent awareness of that relationship. We assume for the sake of simplicity that there is just one learning trial.

Learning itself presumably takes place during or after presentation of event B; we wish to ascertain the subject's state of awareness during this learning episode. Unfortunately, there are likely to be profound technical difficulties in assessing awareness of a predictive relationship at just the moment learning itself occurs. Apart from anything else, asking subjects at time t_1 whether they are aware of the relationship between stimuli A and B is likely to direct their attention to that relationship. As an illustration, in a study by Baeyens et al. (1990a) that will be discussed in more detail later, the proportion of A-B relationships which the subjects appeared to be aware of on a postconditioning recognition test increased from 18% to 77% when subjects *also* gave concurrent estimates of awareness during the learning stage. Clearly, the concurrent index of awareness directed subjects' attention to the relationship and affected the very entity it was designed to measure.

Hence, we will usually have to settle for assessing awareness after the target learning trial. At this time (t_2 in Fig. 1), suppose we present event A (a tone previously paired with shock) and measure the GSR as well as asking subjects whether they have any particular expectancy of event B. If we obtain a GSR but no evidence of a conscious expectancy of event B, we have obtained the crucial finding that lies at the heart of all attempts to demonstrate

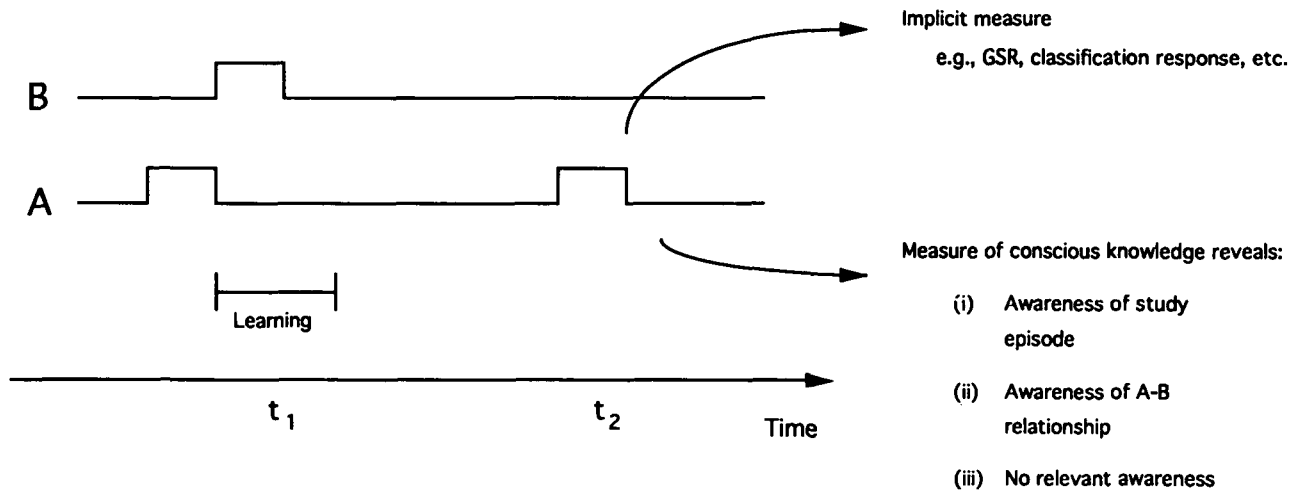


Figure 1. Schematic illustration of events in experiments that investigate the role of awareness in the learning of predictive relationships. Subjects witness a predictive relationship between stimuli A and B, with learning presumed to occur during the interval marked t_1 . At some time later (t_2) stimulus A is presented again. Performance at t_2 is taken as an index of learning at t_1 , whereas a concurrent measure of awareness at t_2 is used to infer the content of the subject's awareness at t_1 .

implicit learning with supraliminal stimuli. For if subjects have no expectancy of event B at t_2 , we have some basis for inferring that they were not aware of the A-B relationship at t_1 .

This might seem to be a very strong inference, but we believe such inferences will have to be accepted if unconscious learning is to be established. It is unavoidably difficult to assess awareness concurrently with learning, so one is forced to rely on some later test. Of course, we also make a backward inference concerning learning itself: if performance at time t_2 is no better than we would expect by chance, we often infer that learning did not occur at t_1 . Conversely, if performance is better at t_2 than we would expect by chance, we conclude that learning did occur.

2.3.1. The relationship between unconscious learning and implicit memory. The basic design shown in Figure 1 allows us to see the intimate relationship between unconscious learning and implicit retrieval: demonstrations of unconscious learning are a proper subset of the larger set of demonstrations of implicit retrieval.

Implicit retrieval is defined as occurring when information from some prior episode can be retrieved and can hence influence current processing, but in the absence of conscious recollection of that prior episode (e.g., Schacter 1987; we use the term "implicit retrieval" rather than the more common term "implicit memory" to emphasize that we are specifically considering what happens during the retrieval process). Thus, implicit retrieval requires the absence of a conscious reexperience of the study episode. Now, lack of awareness of a contingency at t_2 presumably means the absence of any consciously recallable episodic memory traces in which that contingency is embedded, and hence any piece of evidence that allows us to infer unconscious learning must also be an example of implicit retrieval: this is case (iii) shown in Figure 1. The converse does not hold, however; an example of implicit retrieval does not necessarily represent evidence of unconscious learning.

Suppose that a subject emits a GSR when presented at

test with a tone stimulus. There are three possible scenarios, shown in Figure 1:

1. The subject remembers the study episode, in which case the GSR response does not count as an example of implicit retrieval according to Schacter's (1987) definition. Because remembering the episode entails remembering the content of that episode (i.e., the A-B contingency), the learning could not have been implicit either.

2. The subject does not remember the study episode, but is aware – that is, has semantic knowledge – that this tone predicts shock (cf. source amnesia). Although this qualifies as a case of implicit retrieval, we would not infer that learning itself had been unconscious, since at t_2 the subject is aware that A predicts B. (Note that this ignores the possibility that subjects could have been unaware of the A-B relationship at t_1 , but aware of it at t_2 , for example, as a result of observing their own behavior. Observation of a GSR in response to the tone might lead the subject to believe that the tone must therefore predict shock. How one might exclude this possibility is a difficult question.)

3. The subject neither remembers the study episode nor has conscious semantic knowledge of the A-B relationship. This final case again qualifies as implicit retrieval. More important, we now have evidence that is relevant to unconscious learning, as lack of awareness of the relationship at retrieval licenses the inference that learning too took place without awareness.

Thus, in order for us to infer unconscious learning from implicit retrieval, the subject must be unaware of the relevant relationship that occurred in the study episode, in addition to being unaware of the episode itself. In summary, an unconscious learning experiment just is an implicit retrieval experiment, but with the added component of meeting this further condition. For researchers in the field of implicit retrieval, all that is of interest is whether the subject is unaware of the relevant study episode, as in cases (ii) and (iii). But only case (iii) is relevant to the question of unconscious learning; the subject must also be unaware of the relationship that

occurred in that episode. It is for this reason, we argue, that much of the data obtained from amnesics is irrelevant to the question of unconscious learning (see sect. 4).

2.3.2. Dissociation of task performance and verbal reports. Within the dissociation paradigm (Reingold & Merikle 1988), many studies have shown that subjects can acquire information without being able to report it verbally at a later time. Such findings have been taken as support of the claim of unconscious learning. Suppose that subjects are presented with some information at time t_1 and that a subsequent performance test indicates they have encoded this information. We argue that if the aim is to establish what the subjects' state of awareness was at t_1 , examining the content of their verbal reports at t_2 is certainly not the only way to do this and may not be the best one.

To illustrate this, note that the condition mentioned above (that the backwards inference must be valid) can be made more specific by dividing it into two further criteria. The first concerns the match between the information responsible for performance changes and the information revealed by the test of awareness. We call this the Information Criterion. The second criterion concerns the sensitivity of the test for awareness. We call this the Sensitivity Criterion.

Information Criterion: Before concluding that subjects are unaware of the information that they have learned and that is influencing their behavior, it must be possible to establish that the information the experimenter is looking for in the awareness test is indeed the information responsible for performance changes.

This criterion is intended to exclude situations such as the following: suppose the experimenter sets up a task in which performance can be improved if the subjects learn information I . Performance does indeed improve, and subjects are apparently unaware at time t_2 that they have learned I . However, an adequate explanation of the improvement in performance is that subjects are not learning I , but I^* . By the experimenter's criteria, awareness of I^* would be disregarded as irrelevant, and so the experimenter would erroneously conclude that the subjects' performance was under the control of some information or knowledge of which they were unaware. The Information Criterion is closely related to the notion of "correlated hypotheses" introduced by Adams (1957) and Dulaney (1961) and which will be discussed in section 2.6.1.

Our second criterion is far from new (e.g., Brewer 1974; Brody 1989; Dawson & Schell 1985; Ericsson & Simon 1980; 1984; Eriksen 1960; Reingold & Merikle 1988). It is simply that tests of unconscious learning must achieve an adequate level of sensitivity:

Sensitivity Criterion: To show that two dependent variables (in this case, tests of conscious knowledge and task performance) relate to dissociable underlying systems, we must be able to show that our test of awareness is sensitive to all of the relevant conscious knowledge.

Unless this criterion is met, the fact that subjects are able to transmit more information in their task performance than in a test of awareness may simply be due to the greater sensitivity of the performance test to whatever conscious information the subject has encoded. Let us

take as our null hypothesis the claim that there is a single source of conscious knowledge that can manifest itself on both the performance and the awareness test. If performance is above chance, but there is no detectable awareness, an immediate inference is that our test of awareness is simply less sensitive than the performance test to the available resource of conscious information. Or, to put it another way, there is conscious knowledge that is not being detected by the supposed test of awareness but is contributing to task performance.

To rule out this possibility, we must have either (1) some independent reason to believe that the test of awareness is sensitive to all of the potentially relevant conscious information, or (2) some reason to believe that the awareness test is at least as sensitive as the performance test in terms of its ability to detect relevant conscious information. The first of these requires demonstrating that the awareness test is *exhaustive*, something that Reingold and Merikle (1988) have noted is likely to be very difficult to do. In contrast, the second requirement can be met if we try to make the performance and awareness tests as similar as possible in terms of retrieval context, differing only in terms of task instructions. If the instructions in the awareness test encourage the subject to retrieve as much conscious information as possible, and if the retrieval contexts in the two tests are approximately matched, then the Sensitivity Criterion may be met, because it is unlikely that the performance test would elicit the retrieval of more conscious information than the awareness test when the latter has provided subjects with a stronger motivation to do so. If we still obtain a dissociation between performance and awareness under such circumstances, we will have good evidence of unconscious learning.¹

As an illustration of the application of these criteria, consider a widely cited implicit learning study by Lewicki et al. (1987). In the first phase, each trial consisted of the presentation of a target item in one of the four quadrants of a computer screen (which, for purposes of discussion, we can designate as A, B, C, and D); the subjects' task was simply to press a button corresponding to that quadrant as quickly as possible. The basic idea of these experiments can be simply stated: the choice of target location on each trial was nonrandom, and the question was whether the subjects would be able to detect this nonrandomness.

Subjects were presented with sequences of seven trials, with rules constructed so that target locations on the seventh trial could be predicted from its locations on trials 1, 3, 4, and 6. On each of the first six trials, the digit 6 appeared on its own in one of the quadrants of the screen, but on trial 7 (the "complex" trial), it was embedded in a display containing 36 digits. Reaction time on the seventh trial was the measure of interest. Again, the rules specifying target location were deterministic: thus, if the target appeared in locations C, A, D, and B on trials 1, 3, 4, and 6 respectively, then on trial 7 the target would be in location A.

In common with many other such results (which will be reviewed in sect. 2.7 below), Lewicki et al. (1987, Experiment 1) found that reaction times (RTs) on the target trials decreased significantly across 4,608 complex trials. In addition, RTs increased significantly when, toward the end of the experiment, the rules were changed so that on

the complex trials the target now appeared in the quadrant diagonally opposite where it had appeared previously. This latter finding rules out nonspecific factors as the locus of the speedup effect. In a second experiment, Lewicki et al. applied deterministic rules only on two out of three sets of seven trials: on the remaining sets, target location on trial 7 was random. Here, a change in the rules only affected RTs in the sets that were rule determined and not in those that were not.

Lewicki et al. found that none of their subjects came even close to being able to report any of the rules. In fact "none of the subjects were even able to correctly specify which four out of six simple trials were the crucial ones" (1987, p. 529). Thus we appear to have good evidence of a dissociation between performance and reports. It is highly doubtful, however, whether these results meet either the Information or the Sensitivity Criterion. With regard to the former, Lewicki et al. required subjects to try to report "at least one pair of co-occurring elements (i.e., a sequence of four target locations in simple trials and the corresponding location of the target in the subsequent matrix-scanning [complex] trial)" (p. 528). Thus subjects were classified as able to report something about the sequence, and hence as aware, only if they were able to specify a complete sequence of four simple trials and one complex trial. The problem with this classification, however, is that to show a speedup in RT, complete knowledge of the sequences was not necessary.

Analysis of the sequences, for example, shows that even the last simple trial on its own was informative about target location on the seventh trial: if the target was in quadrant A on trial 6, it was twice as likely to be in quadrants A and D on trial 7 as in quadrants B or C. Trial 6 provided a great deal of information on its own about target location on trial 7. Knowledge about trials 4 and 6 provided still more information about target location on trial 7, but if the subjects could report this sort of regularity, it would still not have counted as correct according to Lewicki et al.'s criterion. It is true that knowledge of the sequence across the four relevant simple trials provided absolute certainty about the seventh trial, but our point is that considerable amounts of speedup in RT could be attributable to fragmentary knowledge of "microrules" that Lewicki et al. would not have counted as evidence of awareness, even if the subjects could articulate them.

Turning to the Sensitivity Criterion, we may ask whether the verbal report test is an adequate measure of the subject's awareness in this procedure. We suggest that it is not. First, we cannot be sure that the performance and awareness tests are matched in terms of the conscious information they pick up, because quite different retrieval contexts are provided for the two tests. In the case of RTs, performance is elicited in a context where (1) stimuli are presented on the computer screen, (2) responses are made on the keyboard, (3) a horizontal and a vertical line appear on the screen dividing it into quadrants, (4) a response is made very soon after the preceding response, and so on. All these cues are pertinent, in that they were present during the learning phase (which is just the RT task). In the case of verbal report, none of these cues is present. Instead, the subject is required to retrieve the sequence rules from memory, without the aid of any of the aforementioned cues.

Second, we have little reason to believe that the verbal

report test provides an exhaustive index of conscious information, since there are other tests such as recognition that manifestly detect information left undetected by verbal report tests. For example, Nelson (1978) compared the sensitivity of recognition and verbal recall in the following way. Suppose we have two memory tests, A and B. Subjects learn a list of items and are then given test A. Then, test B is applied only to those study items that test A failed to detect. If test B detects any of these items, it is said to be more sensitive than test A. It is important also to apply the tests in the reverse order – test B, then test A – and to fail to observe an increase in sensitivity. Using such a procedure, Nelson showed that recognition tests can detect items not detected by free recall tests, but the converse was not true. Hence, recognition is a more sensitive test than free recall, and the latter is therefore not exhaustive.

Moreover, note that it is possible that subjects misinterpret free report questions to mean they should only report rules. They might believe that fragmentary information is not supposed to be reported. Many researchers have attempted to avoid this problem by asking more and more specific questions about what stimuli may begin or end a sequence, and so on. Such questions are somewhat better from a sensitivity standpoint because they are more specific (and provide more cues), and may be better from an informational standpoint if they ask about the information that subjects actually learn.

In sum, we suggest that the Information and Sensitivity Criteria are not met in Lewicki et al.'s (1987) experiment. The default hypothesis – that there is only a single resource of conscious information – may be correct, with less of that knowledge being detected by the verbal report test than by the RT task. There is no evidence that the knowledge used to perform the RT task is any different or is in any way acquired independently of the knowledge that the subject's reports are based on. Verbal reports are impoverished compared to task performance simply because less of the available information is retrieved in the test of reportable knowledge. If the subject were given enough retrieval cues, there is every reason to believe that this knowledge could be brought to consciousness and reported; it is simply that a normal test of verbal report does not do this. Last, if sufficient cues could make the information conscious, there is every reason to believe that it was conscious at the time of encoding.

It is important to note that we are not denying the empirical fact that performance and verbal reports can be dissociated. On the contrary, we acknowledge that there have been numerous satisfactory demonstrations of this (for example, in Lewicki et al.'s [1987] experiment), and that this has interesting implications for applied psychology. Subjects' performance indicates that they have learned something, yet they are poor at articulating verbally what they have learned. Instead, we are suggesting that this dissociation is only very weak evidence for the claim that the original learning was unconscious, and that it provides no evidence at all for the functional dissociation of conscious and unconscious learning. Its status is exactly the same as the difference that commonly emerges between tests of recall and recognition. For the same reason, amnesic patients' inability to recall information that an earlier test shows they had learned (e.g., Nissen & Bullemer 1987) is not in its own right evidence

of unconscious learning. Since we are claiming that a dissociation between performance and verbal report is not compelling evidence for unconscious learning, we place special weight (below) on studies that have tried to use more sensitive tests of awareness.

It is also important to recognize that our criteria do not make unconscious learning undemonstrable. As Bowers (1984) has noted, it is pointless to argue about a possible unconscious process if one's criteria for its existence make it a logical impossibility. But the Information Criterion can readily be met in any study that establishes unequivocally what it is that the subject is learning, and the Sensitivity Criterion can be met by tests that adequately reinstate the learning context or that attempt to be exhaustive with respect to conscious information. Indeed, we will see in section 2.7 below that a replication of Lewicki et al.'s experiment by Stadler (1989) met both of these criteria by using an alternative test of awareness. Furthermore, successful demonstrations of unconscious *perception* have been possible in experiments that use tests of awareness that meet these criteria (e.g., Merikle & Reingold 1990). In sum, Lewicki et al.'s (1987) experiments demonstrate the dangers of asking the wrong questions and of ignoring substantial differences between different types of test.

With these considerations in mind, we now turn to other evidence for learning without awareness. In the following sections, we focus on four areas of experimental evidence: conditioning, artificial grammar learning, instrumental learning, and sequential pattern acquisition.

2.4. Awareness and conditioning

2.4.1. Pavlovian conditioning. We begin with a consideration of whether classical or Pavlovian conditioned responses can be acquired in the absence of awareness of the scheduled contingency of reinforcement. Since many researchers regard conditioning as representing a relatively primitive learning system (see Boakes 1989), it is plausible to imagine that learning without awareness can occur in this context. The conclusion from a huge number of studies, however, is quite the opposite: there is no compelling evidence for conditioning in human subjects without awareness of the reinforcement contingency. This conclusion was first reached in a classic review by Brewer (1974), and more recent studies have not changed the situation (see Boakes 1989; Dawson & Schell 1985, for reviews). Such conclusions have not always been heeded, however, because there are still claims in the literature to the effect that conditioning can occur without awareness (e.g., Musen et al. 1990, p. 1074) and is hence an instance of implicit, unconscious learning.

There have been two general approaches to examining the relationship between conditioning and awareness. First, some studies have sought to ascertain whether instructions to the subject concerning the nature of the relationship between a cue and a reinforcer affect conditioning as measured, for instance, by GSRs. The rationale is that if conditioning is a relatively automatic form of learning that can proceed independently of awareness, then changes in the subjects' conscious beliefs ought to have little effect on their behavior. Using this logic, Grings et al. (1973), for example, presented subjects with two conditioned stimuli (CSs), one of which (CS+) was

followed by a shock unconditioned stimulus (US), and one of which (CS-) was not. At the end of the training stage, CS+ elicited a larger conditioned GSR than did CS-. Prior to the second stage, subjects were correctly told that the relationship between stimuli and shocks would now be reversed, with shocks following CS- but not CS+.

As has been observed in many other studies, these instructions had a powerful effect on conditioned responding. Grings et al. found that their subjects responded on the first trial of the second stage to CS- but not to CS+, indicating that their knowledge at least partially controlled their responding. Significantly, the response to CS+, a stimulus that had been paired several times with shock, was no greater than the response to a control stimulus that in the first stage had been presented with uncorrelated USs. Similar results of verbal instruction have been obtained in experiments using phobic stimuli such as pictures of snakes (Davey 1992), where it was once thought that conditioned responding could proceed independently of instructions (e.g., Hugdahl & Ohman 1977).

Although such results are unresponsive of the notion that conditioning can proceed without awareness, they do not address the issue directly because awareness itself is not examined. A recent experiment by Lovibond (1992) exemplifies the approach of eliciting measures of awareness concurrently with conditioned responses. Lovibond presented subjects with two stimuli (slides depicting flowers or mushrooms), one of which (the CS+) was paired with shock while the other (CS-) was nonreinforced. Awareness of the relationship between the stimuli and shock was measured in two ways. First, during the learning phase subjects continually adjusted a pointer to indicate their moment-by-moment expectation of shock (note that asking for a rating of shock expectancy does not specifically direct attention to the A-B relationship); and second, at the end of the experiment they were given a structured interview designed to assess their awareness.

It should be apparent how the design conforms to the basic procedure depicted in Figure 1, except that there are four learning trials. In Lovibond's experiments, each of trials 2-4 in fact represents a new learning trial, an assessment of whether learning occurred on the preceding trial(s), and an assessment of the subject's awareness on the preceding trial(s). The Information Criterion should not raise particular problems here, because there is little doubt that the information the subjects learn (the contingency between the CS and US) corresponds with what the awareness test asks them to report.

In each of the experiments, some subjects gave no indication, on either of the tests of awareness, that they associated A with shock to a greater extent than B. Critically, these subjects also gave no hint of stronger conditioned responding to A than to B. For subjects who were aware of the conditioning contingencies, GSRs were stronger to A than to B. Thus, on the basis of these results we would have to conclude that learning about a CS-shock relationship does not occur in the absence of awareness of that relationship. It is also worth noting that Lovibond's experimental design is well suited to demonstrating that our criteria for implicit learning do not make it a logical impossibility. If his results had been different - something which is simply an empirical matter - the criteria

would have been met and implicit learning could have been firmly established.

Other studies have tried to mask the CS-US relationship and again compare awareness and conditioning. The results have been clear: so long as awareness is measured by an immediate test, usually a recognition test, significant conditioning only occurs in situations where the subject is aware of the contingency (see Boakes 1989; Dawson & Schell 1985). One recent experiment serves to illustrate the typical result. Marinkovic et al. (1989) presented their subjects with a recognition memory task for odors. On each trial, one odor was presented for 8 sec as a "target," followed in succession by three further odors. Subjects' primary task was to say which of the three was the same as the target. One of the three recognition odors was in fact either the CS+ or the CS-. If it was CS+, a shock was presented at its offset; skin conductance was measured as the conditioned response. The question of interest was whether acquisition of GSRs could occur without concurrent awareness of the contingency between the CS+ and the shock. Marinkovic et al. measured awareness with a test in which subjects were required to indicate their expectancy of the shock during each odor on a 7-point scale. Because awareness was measured during the CSs, this again represents a concurrent assessment of awareness, rather than a *post hoc* one.

The outcome was that differential conditioning to CS+ was only observed in subjects classified as aware, indicating that awareness is necessary for conditioning. In addition, Marinkovic et al. obtained some evidence that when conditioned responding did occur, it only started after the onset of awareness. In sum, results from conditioning experiments appear to contradict the notion that this type of learning can proceed without concurrent awareness.

For a variety of reasons, some researchers have questioned whether GSRs condition in the same way other responses, such as the eyeblink or salivary reflexes, do. Thus it is worth noting that correspondences between awareness and conditioning seem to occur with other response systems as well (e.g., for eyelid conditioning, Baer & Fuhrer 1982).

The conclusion from these studies is clear, and confirms Brewer's (1974) earlier analysis: Pavlovian conditioning, which is often cited as a fundamental form of learning, does not seem to occur in the absence of awareness of the reinforcement contingency.

2.4.2. Evaluative conditioning. Evaluative conditioning refers to a form of learning that manifests itself in changes in affective response to a stimulus (Martin & Levey 1978). Specifically, it refers to the transfer of affect from a US to a CS. Some authors (e.g., Baeyens et al. 1990a; Martin & Levey 1987) have suggested that – unlike standard Pavlovian conditioning – this form of learning can proceed in the absence of awareness of the CS-US relationship. We briefly review some of the relevant evidence.

Baeyens et al. (1990a) presented subjects with 10 repetitions of a CS-US pair of slides, in which the CS slide had been previously evaluated by the subject as affectively neutral and the US slide as either liked, neutral, or disliked. Evaluative conditioning was observed in that on a postconditioning test of affect, the CS slides became affectively positive (liked) if they had been paired with a liked US, negative (disliked) if they had been paired with

a disliked US, and they remained neutral if they had been paired with another neutral stimulus.

As a test of awareness, at the end of the learning phase Baeyens et al. showed the subjects each of the CS pictures and asked them to identify which had been the relevant US. If subjects failed to respond correctly they were then asked whether the US had been liked, neutral, or disliked. They were classified as "unaware" of the CS-US relationship if they failed on both of these questions. Evidence that evaluative conditioning occurred without awareness emerged in the observation that conditioning was the same for CS-US pairs, regardless of whether or not the subject was aware of the relationship.

Of course, the test of awareness may have been an insensitive one. Baeyens et al. accordingly tried to use a more sensitive concurrent measure of awareness. One group of subjects was required to indicate during the 4-sec interval between the onset of the CS and US slides whether they expected a liked, neutral, or disliked US stimulus on that trial. Subjects were classified as "unaware" if they failed to respond correctly on the final three pairings of each stimulus combination. Unfortunately, results from this group undermine the notion of unconscious learning. As discussed in section 2.3, subjects could accurately report most of the pairings, and for those few they could not report, there was no significant evaluative conditioning. Further, in another study, Baeyens et al. (1992) found that groups of subjects given increasing numbers of CS-US pairings showed an increase in both the magnitude of evaluative conditioning and the level of awareness as measured by a postconditioning test. In sum, these studies of evaluative conditioning have failed to show that it can occur unconsciously. (See Shanks & Dickinson 1990, for further criticisms of this research.)

Although they are not usually classified as studies of evaluative conditioning, Lewicki's (1986; Lewicki et al., 1989) experiments on the learning of nonsalient contingencies can be readily conceived as such. Lewicki presented subjects with photographs of people accompanied by personality descriptions such as "kind" or "capable." For some subjects all "kind" people had long hair and all "capable" people had short hair, while for other subjects the opposite was the case. Lewicki reported that on test trials in which subjects had to affirm or disconfirm statements classifying new people as either "kind" or "capable," they responded "yes" more often when the description preserved the study-phase correlation than when it broke the correlation. (They also consistently took longer to answer "yes" when the correlation was preserved.)

Lewicki's (1986) subjects were apparently unaware of the relationship between hair length and personality description, because "not one subject mentioned haircut or anything connected with hair" (p. 138) in a test of verbally reportable knowledge. If we take the personality description as being an evaluative response conditioned to the cue of hair length, the results would again appear to suggest unconscious evaluative learning. However, that conclusion requires us to assume, without any supportive evidence, that the Sensitivity Criterion has been met in these studies. In addition, some of Lewicki's results have proven hard to replicate (see de Houwer et al., in press; Dulany & Poldrack 1991); so we must at this stage reserve judgment on whether this form of learning indeed can occur unconsciously.

2.4.3. Conclusions. In experiments examining the relationship between learning and awareness in Pavlovian conditioning, researchers have striven to meet the Sensitivity Criterion by using multiple tests of awareness. The Information Criterion does not raise particular problems, because there is little doubt that the information the subjects learn (the contingency between CS and US) corresponds to what the awareness test asks them to report. Thus these studies provide a reasonably good test of the role of awareness in learning. The results we have surveyed give little reason to believe that unconscious learning can occur in these situations. For evaluative conditioning the evidence is less clear-cut, but we have few reservations in suggesting that unconscious evaluative learning has not yet been adequately established.

2.5. Awareness in artificial grammar learning tasks

Studies of subjects learning artificial grammars present the classic pattern of unconscious learning: subjects clearly learn something about the input domain, but they appear unable either to report the rules of the grammar or to explain their performance. Such studies provide evidence of unconscious learning if learning involves rule induction. In this section we examine the evidence for unconscious learning of artificial grammars and conclude that memorization rather than rule induction is the principal process involved; we conclude that evidence for unconscious learning is weak. Later, in section 3.5, we review several further studies that have examined conscious hypothesis testing in artificial grammar tasks.

In a prototypical experiment, Reber (1967) required subjects to memorize either a series of letter strings generated from a small finite-state grammar or a series of strings generated at random (see Fig. 2). Subjects who learned the rule-governed strings then performed a grammaticality test in which they were asked to accept novel strings that fit the rules and reject novel strings that did not. They categorized 79% of the 44 test strings correctly, which is significantly above chance. Yet these subjects were unable to report the rules they had apparently learned and then used in the grammaticality task.

Reber's (e.g., 1967; 1989a) account of such grammar-learning results, endorsed by many other investigators since then, proposed that subjects use an unconscious, or implicit, rule-induction mechanism. This mechanism creates a knowledge-base of rules that may be used in a grammaticality task but that is inaccessible to conscious report. As with the other unconscious learning paradigms, we believe that there is another way to interpret the data. We can raise two questions. The first (the Sensitivity Criterion) is whether retrospective verbal report is sufficiently sensitive to test for conscious knowledge of the rules. More sensitive measures of subjects' knowledge, such as concurrent thinking-aloud protocols and recognition tests might reveal marginal or uncertain knowledge. The second question (the Information Criterion) concerns what the subjects are learning from the training strings. If subjects have learned something other than rules, then asking them about rules may lead to erroneous conclusions. On the other hand, if we ask the subjects questions about what they did in fact learn, we may get reasonable answers. It may be that usable knowledge is always both consciously learned and consciously

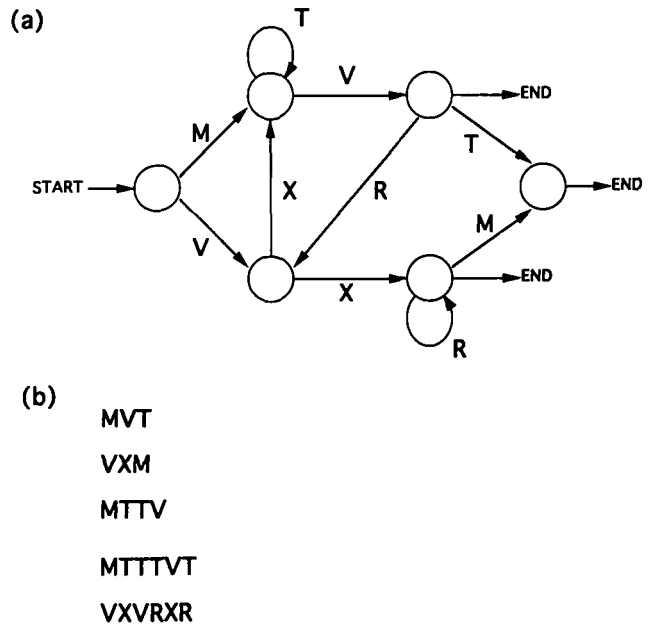


Figure 2. String generator and example strings. (a) Diagram of a finite-state grammar. Strings are generated by selecting one of the possible routes through the network, commencing at "start" and continuing until one of the "end" symbols is reached. (b) Several example strings generated by the grammar.

applied. The experimenter's job is to discern what information subjects are aware of during training and whether that information is used to perform the grammaticality task.

2.5.1. Types of knowledge. The literature has identified three types of knowledge that might be acquired by subjects: rules, memory for whole strings, and knowledge of the frequency and position of substrings, such as pairs of letters. There are several problems with rules. First, it is not really clear what a "rule" would be like: Is it a rewrite rule or a transition graph? How complex can it be, and how many are there? Second, such rules would be very difficult for any but very sophisticated subjects to articulate even if they did explicitly acquire them. Third, it is not clear what sort of mechanism is capable of acquiring such rules, particularly since it must *ex hypothesi* operate outside consciousness. In the face of these questions, it seems sensible to consider other types of knowledge first, and to determine the extent to which they can account for subjects' performance. We return to the evidence for knowledge of rules in artificial grammar learning tasks in section 2.5.3.

The picture with regard to memory for whole strings and knowledge of substrings seems reasonably clear. Such knowledge is easy to articulate and there is ample evidence that subjects do acquire this information, because they do articulate it. These types of knowledge are also consistent with a variety of contemporary memory models, such as chunking (Servan-Schreiber & Anderson 1990), distributed memory (Cleeremans & McClelland 1991), and memory-array models (e.g., Estes 1986; Hintzman 1986; Nosofsky 1986). In addition, these models have been shown to approximate subjects' grammaticality test performance. For example, Dienes (1992) compared a number of these memory models on a set of grammaticality judgment data and was able to achieve good

fits, particularly with distributed memory models. We return to this topic in section 3.3.

With these different knowledge types in mind, we can now ask what sort of information subjects in artificial grammar learning tasks actually acquire, and whether they are conscious of it. A number of studies have asked these questions using several methods and have asked them at various points during training and testing. Mathews et al. (1989) interrupted subjects periodically during training and asked them to instruct an imaginary confederate how to distinguish the grammatical strings. The trained subjects performed better on the grammaticality test than did the yoked subjects, suggesting that not all of the trained subjects' knowledge was explicit and reportable. This verbal report procedure, however, is essentially uncued recall, and so is unlikely to evoke all of the subjects' knowledge of the grammar. More interesting, though, is that the verbal instructions that subjects did report consisted mainly of legal bigrams and other short sequences, sometimes coded by their positions in legal strings.

In a study by Servan-Schreiber and Anderson (1990), subjects were trained on grammatical strings using a recall task. For training, the strings were divided into substrings using gaps (*T PPP TX VS*). Servan-Schreiber and Anderson hypothesized that subjects in all grammar-learning tasks encoded the strings into substring chunks, and the gaps were used to ensure consistent chunkings across subjects. Subjects' written recall preserved these gaps. Servan-Schreiber and Anderson suggested that this phenomenon demonstrates that subjects were in fact encoding the strings as sequences of short strings in accord with the gaps. The subjects' later grammaticality judgments supported this contention as follows. Servan-Schreiber and Anderson constructed ungrammatical strings that consisted of illegal sequences of legal substrings (e.g., *PPPTXTVS*). If subjects were learning just the substrings then these strings would be falsely accepted as legal strings. Indeed, 50% of these strings were mistakenly accepted. On the other hand, test strings that violated specific substrings were correctly rejected; only 26% of these strings were mistakenly accepted. Both subjects' written protocols during training and their test performance, then, support the hypothesis that subjects learn simple substring information in grammar-learning tasks. That only 50% rather than 100% of the strings containing illegal sequences of legal substrings were accepted does not imply that knowledge of substrings is insufficient to account for performance completely. Compared to grammatical strings, these nongrammatical strings (by definition) still contain illegal bigrams (e.g., *XT* in the example above). In addition, subjects' knowledge at test time is clearly incomplete: previously seen grammatical strings were only accepted 70% of the time.

Moreover, Servan-Schreiber and Anderson (1990) went on to build a model that acquired chunks and then used them to evaluate the grammaticality of test strings. The model performed at the level of trained subjects ($r = 0.935$). This result supports their claim that subjects are learning and using chunks by demonstrating that chunks are learnable and sufficient to account for the level of performance of subjects on the grammaticality task.

It is possible that Servan-Schreiber and Anderson's

presentation technique, placing gaps in the training strings, biased subjects' learning procedure. A similar experiment by Perruchet and Pacteau (1990), however, used the standard (no gap) format during training and found similar results. Subjects were trained on strings generated from the same grammar that Reber and Allen (1978) used. To test for awareness of simple substrings, trained subjects performed a recognition test on letter pairs present in the training strings. Subjects performed quite well: only 3 out of 25 old pairs were judged less familiar than any new pair. The correlation between recognition scores and the frequency of occurrence of pairs in the training strings was 0.61. According to the results of the recognition test, then, subjects were aware of the relative frequencies of letter pairs. Similarly, Dulany et al. (1984) concluded that a recognition test of awareness could elicit as much knowledge as was projected in the grammaticality test.

Perruchet and Pacteau (1990) also constructed test strings that contained either (1) illegal orders of legal pairs, or (2) illegal pairs. If subjects only had information about legal pairs on which to judge the grammaticality of test strings, then the illegal pairs should have been rejected, but the illegal orders of legal pairs should have been mistakenly accepted as grammatical. This is the pattern of results Perruchet and Pacteau obtained. Discriminability, measured in *D* scores (zero indicates random responding), was 22 for illegal pairs but only 7 for illegal orders. These results therefore further support the hypothesis that subjects are aware of and make use of only simple substring information.

Perruchet and Pacteau then considered a model that used pair frequency information to make grammaticality judgments. The model produced the same level of performance as subjects, except in one instance. Subjects were sensitive to the beginnings and endings of strings, but the model was not. Perruchet and Pacteau concluded that subjects primarily knew letter pairs, but also which pairs could legally start and end strings. Together with the behavior of Servan-Schreiber and Anderson's (1990) chunking model, these results show that simple fragment-memorization systems can be sufficient to account for subjects' imperfect performance on grammaticality tests.

Dienes et al. (1991) also found evidence that subjects were sensitive to more than just pairs. Following training and a grammaticality task, subjects were given incomplete letter sequences varying in length from zero letters upwards (e.g., *VXT . . .*) and asked to judge which single-letter continuations (*M? V? X? R? T?*) were acceptable at the next location in the string. In this sequential letter dependencies (SLD) task, which was hypothesized to be sensitive to conscious knowledge of the grammar, subjects were sensitive to illegal orders of legal pairs even in the middle of strings. Dienes et al. showed that the knowledge that subjects demonstrated in the completion task correlated with their grammaticality judgments and could be used to model the grammaticality judgment data. They found in addition that knowledge gleaned from subjects' free reports also correlated with their grammaticality judgments, but that less knowledge was reported in the free report task than in the continuation task. These correlations suggest that a single knowledge

source is tapped by both tasks, but that the free report task, uncued recall, is less sensitive.

Reber and Allen (1978) asked subjects to describe retrospectively their learning experience and, concurrently, to justify their grammaticality judgments. Overall, subjects justified their classifications on 821 out of 2,000 test strings. Subjects reported using a variety of information in making their grammaticality judgments. The violation or nonviolation of bigrams was the most common justification, especially concerning the first bigram of a string. String-initial bigrams accounted for fully 30% of the justifications. Violations of single letters, particularly the first or last letter of a string, and violations of trigram or longer sequences were also reported, as well as recognition of and similarity to whole training strings. The grammaticality responses to the remaining unjustified cases presumably consisted of guessing or of knowledge that could not be elicited by verbal report.

So much for substring knowledge. Vokey and Brooks (1992; Brooks & Vokey 1991) have argued that subjects can encode whole-item information in addition to substring information. They found that the similarity of test strings to specific whole-study strings is an important factor in subjects' grammaticality judgments. When the grammaticality and the similarity of the test strings were varied independently, they were shown to be additive factors on grammaticality judgments. Vokey and Brooks argued that such a result indicates that subjects have encoded the whole strings and can determine similarity relationships between strings.

Brooks and Vokey's evidence for whole-string information raises no particular problems for our interpretation of the artificial grammar learning data, since subjects are clearly aware of their whole-string knowledge just as they are aware of the substring knowledge; the study task, after all, requires the subjects specifically to memorize whole strings. However, as Brooks and Vokey (1991, p. 321) themselves concede, their results can at least in principle be explained without reference to whole-item knowledge. Just as grammatical test strings tend to contain more studied bigrams than nongrammatical strings (Perruchet & Pacteau 1990), so also a test string that is highly similar to a study string will contain more studied bigrams than one that is less similar. In fact, Vokey and Brooks' results have been challenged by Perruchet (1994), who has shown that both the effect of similarity and the apparently independent effect of grammaticality that Vokey and Brooks obtained can in turn be reduced to substring knowledge. Grammatical test strings tend to contain more substring components that were part of the training strings than do nongrammatical items. The same is true for similar and dissimilar test items, with similar items tending to contain more substring components from the study strings.

A final piece of evidence supports the view that grammaticality judgments are controlled by comparison to memorized substring or whole-item information. On such a view, but not on an abstraction account, it is likely that judgments would be relatively susceptible to changes in the superficial characteristics of the studied strings. To test this, Whittlesea and Dorken (1993) required subjects to pronounce the training strings from one grammar and to spell the training strings from another grammar. At

test, subjects were asked either to pronounce or to spell test strings and to judge their grammatical status. Subjects were more likely to assign test strings to grammars when the encoding task matched the task for the test string than when they differed. Test strings that were equally similar to strings in both grammars were assigned to the grammar where the encoding and test tasks matched. Such results, although consistent with the idea that judgments are based on a comparison with a set of items in memory that represent the study items in a relatively unanalyzed form, would clearly not be anticipated if what was encoded were the underlying abstract rules of the grammar.

Our conclusion from this section, then, is that subjects use their memory system to acquire knowledge of (possibly) whole strings and (certainly) their parts, and that this simple information is conscious both during acquisition and testing. The results reported by Dulany et al. (1984), Perruchet and Pacteau (1990), and Dienes et al. (1991) show that the knowledge that subjects can consciously retrieve in a recognition test is sufficient to explain their grammaticality judgments. From the evidence we have considered, we do not need to assume the existence of an additional implicit knowledge base, and conclusions to the contrary have arisen because of failures to meet the Information Criterion.

Our interpretation rests on the results of a variety of tests of conscious knowledge that have attempted to address the Sensitivity Criterion. Dienes et al.'s (1991) SLD test, for instance, which required subjects to judge which continuations of a sequence of letters were legal, was actually found in a signal detection analysis to be *more* sensitive than the implicit grammaticality test itself. Thus, if such a test is accepted as a measure of explicit knowledge, no evidence of a dissociation between learning and awareness emerges. Of course, an alternative (see Reber et al. 1985; and the reply, Dulany et al. 1985) is to argue that performance on these explicit tests is contaminated by unconscious influences; subjects may choose a correct continuation on the SLD test as a result of some implicit knowledge to which they do not have conscious access.² The problem with this interpretation, however, is that it means we would have to abandon the test as an index of conscious learning and rely instead on verbal reports, in which case it is hard to see how the Sensitivity Criterion can ever be met. And if that criterion cannot be met, then how are defenders of unconscious learning ever going to unconfound test type from sensitivity, and hence establish the existence of unconscious learning?

We believe it is rather unlikely that unconscious influences play a significant role in the SLD test. Presenting subjects with a letter sequence (e.g., VXT . . .) and asking them to judge, under no time pressure, whether a given letter (e.g., M) could continue the sequence would seem to be a prototypical example of a task requiring conscious reflection, even if it involves mere conscious recollection of studied strings. Nevertheless, to claim that the SLD test is only sensitive to conscious information does require adopting what Reingold and Merikle (1988) call the "exclusiveness" assumption: the assumption that performance on a test of awareness is only affected by conscious influences. This, of course, is a very strong assumption and one that may well be incorrect.

2.5.2. Learning systems. In addition to the question of awareness, a second issue concerns whether whole item, bigram, and possibly rule information are acquired by a single learning system or by separate systems. If they are acquired by separate systems, perhaps those systems interfere with each other's operation? To examine this possibility, Reber and Allen (1978) manipulated the training task. Subjects either observed the strings without any explicit task (observation training), or they performed a paired-associate task, where each string was paired with a different city name. The idea was that the paired-associate task would require better item encoding, thereby facilitating item knowledge but potentially inhibiting other learning processes.

The paired-associate task produced several significant differences from the observation task. Overall, paired-associate subjects were less accurate on their grammaticality judgments: 72.4% versus 81.2% accurate for observation subjects. Paired-associate subjects produced twice as many recognition justifications as did observation subjects (77 vs. 40), and paired-associate subjects' probability of making consistent errors suggested they were more likely to develop unrepresentative knowledge than were observation subjects.

Clearly, the two training tasks affected the quantity of whole-item and substring knowledge that was acquired, but the underlying learning processes do not appear to be in opposition. The verbal reports show that both groups justified their responses with the same knowledge sources, but to differing degrees. It appears, then, that whole-item learning is compatible with substring learning. Vokey and Brooks (1992) examined a range of encoding tasks that produce differences in the extent of item knowledge, but they also found no reliable interference between item knowledge and substring knowledge.

Finally, Dienes et al. (1991) required subjects to generate random digits during training. Their goal was to test whether this task would interfere selectively with subjects given explicit instructions to search for rules that describe the study strings, but not with subjects given implicit instructions simply to observe the study strings. Instead, Dienes et al. found equivalent reductions in learning for both implicitly and explicitly instructed subjects.

2.5.3. Implicit rule induction. Although the considerable evidence presented above supports the conclusion that subjects' knowledge consists of simple substrings (or whole strings), there are two further pieces of evidence that support the conclusion that subjects learn rules. The first piece of evidence supporting rule learning was reported by Reber and Lewis (1977). Subjects were trained on a subset of strings and then solved "anagrams" based on the remaining strings generated from the grammar – that is to say, they took strings of letters and rearranged them to make grammatical strings. The frequencies of bigrams produced by subjects in the anagram task were tabulated and compared with the frequencies of the bigrams in the training set and in the full set of grammatical strings. If subjects were learning bigram frequencies from the training strings, the correlation between the frequencies of bigrams in the training strings and in the solved anagram strings should be high. While this was the case, Reber and Lewis found that the correlation between the frequencies

of bigrams in the solved anagrams and in the whole grammar was actually higher. This result suggests that the subjects went beyond the training set to learn the rules of the grammar.

Perruchet et al. (1992), however, argued that Reber and Lewis's (1977) result must hold on statistical grounds alone. The anagrams demand the production of certain bigrams and not others, in fact, exactly those bigrams that are underrepresented in the training set. Suppose, for example, that *VT* is a bigram in the grammar that is underrepresented among the training strings. *VT* must then be overrepresented among the solved anagram strings since the training and correctly solved anagram strings together constitute the complete set of grammatical strings. It is no wonder, then, that the correlation between the frequencies of anagram bigrams and training bigrams is low and that the correlation between anagram bigrams and the full grammar bigrams is higher. Perruchet et al. went on to demonstrate this fact empirically by training subjects only on the individual bigrams from the training strings. Under these circumstances, subjects could not be learning rules because they only saw bigrams, yet as with Reber and Lewis's subjects the frequencies of their anagram bigrams also correlated better with the full grammar bigrams than with the training string bigrams. The original conclusion, therefore, that subjects go beyond the training strings to learn rules appears to have been an artifact of the experimental design.

The second and more compelling piece of evidence for abstraction is the fact that subjects show some degree of transfer to strings governed by the same underlying grammar, but formed from a new set of letters or from a completely new set of stimuli such as tones. Reber (1969) trained subjects to recall grammatical strings, and when he switched to a new set of letters, subjects showed no increase in recall errors. This result suggested that subjects had learned abstract rules that were easily instantiated with different letters. More impressively, Altmann et al. (in press) required subjects to observe a set of letter strings, generated from the grammar shown in Figure 2, prior to making grammaticality judgments concerning sequences of tones. Some of the tone sequences could be generated from the grammar by substituting a tone for a letter (e.g., middle *C* for the letter *M*). Altmann et al. found that exposure to letter sequences allowed grammatical and nongrammatical tone sequences to be discriminated at better-than-chance levels. Although the improvement was generally small (about 5% increase in correct classifications), this result strongly suggests that at least some aspects of the abstract structure of the letter sequences had been isolated and were available to aid classification of the tone sequences.

It is important to note that the change of stimulus set did have a detrimental effect on performance, however. Compared to a situation in which the study and test items were from the same set (both letters or both tones), classification performance was significantly impaired when the study and test sets differed. Thus, abstract knowledge was plainly not the sole source of information that subjects were relying on – specific memorized fragments or strings must also have been playing a role. A study by Mathews et al. (1989) confirms this conclusion. In Mathews et al.'s study, over a series of training sessions

subjects were trained either on a single-string set or on different sets generated from the same underlying grammar. Subjects in the same set condition learned better, and a final switch to a new set doubled the error rates in the single-set training condition. Such a result would not be expected if an abstract set of underlying rules were the sole factor guiding classification, because the rules would apply equally to the new and to the original letter set.

What is the significance of these results for unconscious learning? To the extent that subjects might be poor at describing what they have abstracted, such results may imply that unconscious learning is taking place. But given the rather small improvement in classification performance that results from training and testing on different sets of items, it is quite likely that what is abstracted is fairly limited (e.g., only two initial symbols are legal, the first two symbols of a string cannot be the same, etc.), and it is quite possible that subjects, if asked, would be able to report such simple regularities. In sum, although the data from transfer studies do suggest that some aspects of the underlying structure can be abstracted, from the point of view of unconscious learning the significance of these findings has yet to be established.

2.5.4. Conclusions. These studies indicate that relatively simple information is to a large extent sufficient to account for subjects' behavior in artificial grammar learning tasks. In addition, and most important, this knowledge appears to be reportable by subjects. Appreciable knowledge of the grammar does not seem to be acquired by explicit hypothesis testing or other complex analytic processes (although we return in sect. 3.2 to consider some rather different cases where grammars appear to be learned explicitly). Instead, knowledge seems to be mainly accumulated over training by simple memory mechanisms that collect frequency statistics on bigrams, slightly longer sequences, and possibly whole items.

2.6. Awareness in instrumental learning tasks

In contrast to the conditioning and artificial grammar studies described above, which arrange relationships between external cues, instrumental tasks establish some contingency between an action the subject performs and an associated outcome. Learning is measured as a change across trials in the propensity to perform the action. Naturally, the question we may again ask is whether such learning can occur without awareness. As in his review of Pavlovian learning studies, Brewer (1974) concluded that the answer to this question is no. There have recently been some further investigations of the role of awareness in instrumental learning: we consider results separately from tasks in which the instrumental contingency is simple or more complex. By "simple" we mean any task in which there is ostensibly just one action available to the subject.

2.6.1. Simple instrumental learning tasks. Svartdal (1989; 1991) has reported a number of studies in which subjects are led to believe that there is a relationship between a reinforcer and one aspect of responding, when in fact the critical variable is some other aspect of responding. For example, Svartdal (1991) presented subjects with brief trains of between 4 and 17 auditory "clicks." Subjects

immediately had to press a response button exactly the same number of times and were instructed that feedback would be presented when the number of presses matched the number of clicks. In reality, however, feedback was contingent on the rate of responding: for some subjects, it was given when the interresponse times (IRTs) were lower than in a baseline phase, while for others it was given when IRTs were higher.

Svartdal (1991) obtained evidence of learning, in that IRTs adjusted appropriately to the reinforcement contingencies, but subjects seemed to be unaware that it was the rate of responding that was important. A structured questionnaire revealed no evidence of awareness of the contingency between response rate and feedback in subjects whose response rate had adjusted appropriately.

Such demonstrations appear at first glance to be quite compelling, especially as the contingency to be learned is such a simple one. It is unclear, however, that the Information Criterion is met in these and similar studies, because it is very difficult to rule out the possibility that subjects acquire "correlated" hypotheses about the reinforcement contingency that are incorrect from the experimenter's point of view but happen to produce response profiles that are difficult to distinguish from those generated by the correct hypothesis. For example, suppose subjects learn that resting their hand in a certain position increases reinforcement rate. This could be a true experienced contingency if that hand position was conducive to a fast or slow response rate. Such an "incorrect" hypothesis would generate behavior that was very similar to what would be produced by the correct hypothesis, yet a subject who reported hand position as the crucial variable would be regarded by the experimenter as "unaware" of the reinforcement contingency.

Although such a criticism is undoubtedly *post hoc*, there is good evidence of subjects' behavior being under the control of such correlated hypotheses. In the 1950s, a number of studies asked subjects to generate words *ad libitum* and established that the probability with which they would produce, say, plural nouns was increased if each such word was followed by the experimenter saying "umhmm" (e.g., Greenspoon 1955); as with Svartdal's (1991) experiment, this result occurred in subjects apparently unable to report the reinforcement contingency. However, in an elegant study, Dulany (1961) proved that subjects were hypothesizing that reinforcement was contingent on generating a word in the same semantic category as the previous one. Although incorrect, this hypothesis was correlated with the true one, in that if the subject said "emeralds" and was reinforced, then staying in the same semantic category meant they were more likely to produce another plural noun ("rubies") than if they shifted categories. Thus the subjects were perfectly aware of the contingency that was controlling their behavior, namely, the contingency between staying in the same semantic category and reinforcement.

In sum, even ignoring possible insensitivity in the test of verbal awareness, results such as Svartdal's (1991) cannot be taken as conclusive evidence of unconscious learning. Subjects may learn a rather different contingency from that explicitly programmed by the experimenter, and the Information Criterion may hence fail to be met. The problem is particularly worrisome in operant studies because, by definition, the experimenter has little

control over the subject's behavior and therefore over the contingencies that may be present. In nonoperant tasks, the problem can be avoided because the experimenter can in principle eliminate all reinforcement contingencies except the one of interest. For this reason it seems that clear evidence for unconscious learning is likely to be difficult to establish in instrumental learning tasks.

In contrast to such apparent dissociations between learning and awareness, Shanks and Dickinson (1991) have argued that there are a number of variables that seem to have rather similar effects on performance assessments of learning and on awareness. In two studies, subjects performed a simple operant-learning task in which pressing a key on a computer keyboard was related, via a schedule of reinforcement, to a triangle flashing on the screen. Subjects were exposed to a reinforcement contingency in which they scored points whenever the triangle flashed but lost points for each response, so that they were encouraged to adapt their response rate to the reinforcement schedule. Learning was demonstrated by changes in subjects' rates of responding. As a measure of awareness, subjects were asked to report on a scale from 0 to 100 what they thought was the relationship between the response and the reinforcer.

Shanks and Dickinson (1991) found that response rate was sensitive both to the degree of contiguity between the response and reinforcer and to the degree of contingency between them. At the same time, subjects' judgments were equally sensitive to these factors. Furthermore, certain judgmental illusions likewise manifested themselves in performance measures. For example, subjects frequently judge an action and an outcome to be related when in fact they are not. Shanks and Dickinson found that this effect appears in performance measures such as response rate as well as in verbal judgments. Of course, the appearance of a bias in two behavioral measures strongly suggests that they are mediated by a common underlying process.

The notion that learning and awareness proceed in tandem is corroborated to the extent that they are affected in similar ways by various manipulations. Shanks and Dickinson's results indicate that – at least for the two important factors of contingency and contiguity – this is exactly the case. Shanks (1993) discusses some further apparent concordances.

The human operant-learning literature provides perhaps the most convincing evidence that learning and awareness are *associated* in simple learning tasks. A wealth of data shows concordances between response rate and verbal reports under different schedules of reinforcement (e.g., Catania et al. 1989; Rosenfarb et al. 1992; see also Skinner 1984b, and accompanying commentaries). For instance, Rosenfarb et al. required subjects to press a button either on a differential-reinforcement-of-low-rate schedule, in which reinforcers were delivered for a response provided that 5 sec had elapsed since the preceding response, or on a fixed-ratio schedule in which eight responses were required to earn a reinforcer. Rosenfarb et al. found that subjects' verbal reports concerning the programmed contingency accorded very well with the actual contingencies. Furthermore, there was a strong correlation between the time at which responding became appropriate for a schedule and the time at which

verbal reports indicated awareness of the reinforcement contingency operating in that schedule.

2.6.2. Complex instrumental control tasks. Several experiments have investigated the relationship between learning and awareness in more complex instrumental learning tasks where the subject has to learn to control an interactive system. Again, the basic idea is as shown in Figure 1, with some learning episode followed by an assessment of awareness. In most of these tasks awareness at time t_2 is measured by verbally questioning the subject.

Berry and Broadbent (1984) conducted an influential and widely cited experiment in which there was an apparent dissociation between learning and awareness. As in Hayes and Broadbent's (1988) study, one of the tasks they used required subjects to interact with a computer "person." On each trial, the subject entered an attitude (e.g., polite) to the computer, which then responded with its attitude (e.g., unfriendly). The subject's task was to try to get the computer to be friendly. The computer's attitude on each trial was a simple numerical function of the subject's input on that trial and the computer's previous attitude. Inclusion of the computer's attitude on the previous trial makes the task quite a difficult one to learn.

Berry and Broadbent (1984, Experiment 1) found, not surprisingly, that performance improved with practice: significantly more trials on target occurred during a second block of 30 trials than during the first block. However, scores on a structured questionnaire designed to assess the subjects' reportable knowledge of the task were no better after the second block than after the first one. Hence, here we have apparent evidence that learning to perform a task can take place without any change in awareness of the underlying structure of the task. Similar results have been obtained in a number of other studies (e.g., Berry & Broadbent 1987; 1990; Broadbent et al. 1986; Hayes & Broadbent 1988; Stanley et al. 1989).

On the other hand, a detailed examination by Sanderson (1989) found evidence of associations rather than dissociations between performance and reports. Sanderson argued that because subjects often have complex prior beliefs about the interactions within a large system, and because these beliefs may be erroneous in a laboratory version of the system, it is possible for their mental models to undergo considerable revision without yielding an overall improvement in accuracy. It is only with prolonged practice that mental models, and hence the verbal reports based on them, begin to show noticeable improvement. Consistent with this, Sanderson was able to obtain significant performance improvements at the same time as weak improvements in the overall accuracy of verbal reports in a complex transportation task, but she showed that the detailed nature of the verbal reports was changing very considerably.

A further experiment by Berry and Broadbent (1984) found the converse of the previous dissociation, namely, reportable knowledge improving without corresponding improvements in task performance. One group simply completed two sets of trials, while between the two sets another group received detailed verbal instructions about the nature of the input-output relationship. These instructions essentially represented a verbal description of the equation governing the computer's attitude. When ques-

tioned at the end of the experiment, subjects who had received instructions outscored those who had not, yet the groups were indistinguishable in terms of number of trials on target. Thus a change in "awareness" (or at least a change in reportable knowledge of the task) was not accompanied by a change in task performance.

What are we to make of such dissociations? One possibility is that it is not only possible for learning to proceed without awareness, but in addition the system responsible for implicit learning is quite independent of another (explicit) system in which learning is accompanied by awareness. Such a "systems" account would then be able to explain why we can obtain double dissociations of the sort reported by Berry and Broadbent (1984): learning to perform the control task involves the implicit system, and proceeds without awareness, while a change in awareness involves the explicit system and can proceed without any benefit in task performance.

While double dissociation results are certainly consistent with the notion that there are two learning systems, one conscious and the other unconscious, we feel that an alternative account is equally feasible: there may be two systems, both of which are conscious, but which encode different types of knowledge. The basic problem is that we do not know that the sort of knowledge the subjects in Berry and Broadbent's experiments acquire when learning to perform the task is at all the same as the knowledge they require to score well on the test of reportable knowledge (i.e., the results may fail to meet the Information Criterion). Suppose, for the sake of argument, that good task performance simply depends on learning an unrelated set of stimulus-response (S-R) pairs or instances (evidence for such a possibility certainly exists; see Cleeremans 1993b). It is then not hard to imagine that although practice provides the subjects with more and more knowledge of this sort, they might be hard pressed to use such knowledge when faced with questions about possible structural rules underlying the task. At the same time, giving the subjects detailed instructions about the task may improve their knowledge of the rules, and hence their questionnaire scores, but might not transfer to better performance on the task itself since S-R knowledge is required for that. But of course the subjects' inability to describe the rules underlying the task does not imply that the S-R learning occurred without awareness: if they were asked to report *that* knowledge, perhaps subjects would be able to do so. In sum, there are ways of interpreting such data that do not appeal to unconscious learning (see Stanley et al. 1989, for an examination of some of the alternative types of knowledge that subjects may encode).

A second problem concerns the sensitivity of the test of awareness. Can we be certain that the questionnaire procedure exhausts the subject's knowledge of the task? Can we be confident that failure to express on the questionnaire any awareness of the nature of the task means that the subjects were unaware at the time they were learning? For example, one alternative strategy would be to ask each subject to instruct a yoked "partner" in how to perform the task. If the partners could then perform the task as well as the original subjects, we would conclude that the original subjects were in fact able to articulate all their task knowledge. Such procedures have been used with other learning procedures (e.g., Mathews et al.

1989, for grammar learning) and have proven highly sensitive.

2.6.3. Conclusions. Instrumental learning experiments arrange some relationship between the subject's actions and certain outcomes. Implicit learning would be demonstrated if learning, as indexed by changes in instrumental behavior, occurred in the absence of awareness of the reinforcement contingencies. Although some studies have found that subjects seem unaware of the relevant contingencies, reliance on verbal report means that the Sensitivity Criterion is unlikely to have been met. Furthermore, because the experimenter necessarily relinquishes a certain degree of experimental control in an instrumental learning task, it is difficult to rule out the possibility that the subject is responding on the basis of a correlated hypothesis, in which case the Information Criterion is violated. Finally, even ignoring these considerations, a surprisingly large number of studies have documented impressive concordances between behavior and awareness.

2.7. Learning and awareness in serial reaction time tasks

Nissen and Bullemer (1987) and Lewicki et al. (1987) introduced a simple and ingenious technique, the serial RT task, in an attempt to demonstrate unconscious learning. In Nissen and Bullemer's version, a stimulus is presented on each trial in one of four locations (A-D) and the subject simply has to press the button corresponding to that location as fast as possible. The subject is given instructions appropriate for a typical choice RT task, but in fact there is a sequence underlying the selection of the stimulus on each trial. The question is, can subjects learn the sequence without being aware of it? With respect to Figure 1, the subject is presented with a series of learning trials in which there are predictive relationships between stimuli. These are accompanied by both a concurrent assessment of learning (RT) and a later assessment of awareness.

Some of the most compelling evidence for unconscious learning using this technique comes from a later study by Willingham et al. (1989); this study is worth considering in some detail because of the heavy reliance placed upon it in recent discussions of conscious and unconscious processing (e.g., Velmans 1991). In their first experiment, Willingham et al.'s subjects performed a 4-choice RT task. The actual sequence of signals was DBCACBDCBA, which repeated many times with no break between cycles. Subjects' RT improved across a total of 400 trials. To see whether this improvement represented knowledge of the sequence or general nonspecific speedup, Willingham et al. compared the speedup of their subjects to that obtained in a group of subjects from the earlier study by Nissen and Bullemer (1987) for whom there had been no structured sequence; for these control subjects, target location was random from trial to trial, except that the same location never occurred on consecutive trials.

The improvement in RT was significantly greater in the sequence group than in the control group, apparently indicating that sequence learning had occurred. Furthermore, this was still true for subjects who subsequently

reported no awareness of the existence of a sequence during the RT trials.

2.7.1. Problem of suitable control group. Although such results suggest the possibility of unconscious learning, there are a number of significant problems with such experiments. First, the demonstration of sequence learning has typically involved one of the following two comparisons: (1) a comparison (e.g., Willingham et al. 1989) between a group exposed to the sequence and one for whom the stimulus on each trial is chosen at random (with the constraint that stimuli never repeat on consecutive trials), or (2) a within-subjects comparison (e.g., Hartman et al. 1989) between performance at the end of a long period of exposure to the sequence and performance on a subsequent block of trials where the stimuli are chosen at random, again with the constraint that stimuli never repeat on consecutive trials. The problem with both of these comparisons is that performance can differ between the sequence and random trials without the subject having any knowledge – implicit or otherwise – of the sequence.

As a moment's reflection reveals, faster responses on the DBCACBDCBA sequence compared with a random sequence might simply be due to response biases developing during exposure to the sequence. The stimuli are not equally frequent (B and C occur three times, D and A twice) in the 10-trial sequence. Thus in the sequence, but not in the random conditions, the subject is to some degree able to predict which stimuli are most probable, a fact that – as has been demonstrated extensively (see Broadbent 1971) – will allow fast responses to develop.

Clearly, the appropriate comparison is with a group of subjects who receive a "pseudorandom" series constrained to have the same number of each of stimuli A, B, C, and D per 10 trials as appear in the sequence proper, and in which stimuli never repeat on consecutive trials. Such an experiment was reported by Shanks et al. (1994). One group of subjects was presented with the normal sequence, another with the pseudorandom series, and a third with a "truly random" sequence, in which again there was the constraint that stimuli never repeated on consecutive trials. The stimuli were dots arranged in a horizontal row and the general procedure followed that of Willingham et al. (1989).

After 400 RT trials, subjects in the sequence group were classified on the basis of a structured interview as having no knowledge of the sequence, some knowledge, or full knowledge. The prediction was that if the no-knowledge subjects had indeed learned something about the sequence, they should have speeded up more than the pseudorandom subjects. In all but the truly random group the RT difference between the first and fourth block of 100 trials was significantly greater than zero. The normal-sequence/full-knowledge group speeded up more than any of the others; the difference between the normal-sequence/full-knowledge and pseudorandom groups confirms that the normal-sequence/full-knowledge subjects had indeed learned something about the sequence. However, there was no significant difference between the normal sequence/no-knowledge and pseudorandom groups, though both speeded up more than subjects exposed to the truly random series. Thus, we suggest that with Willingham et al.'s stimuli and proce-

dures, most if not all of the supposedly implicit learning in the normal-sequence/no-knowledge group was simply due to the development of response biases reflecting knowledge of the frequencies of the different stimuli.

As a consequence, Willingham et al.'s experiment fails to satisfy the Information Criterion. The subjects may have been unable to articulate information about the sequence verbally simply because they were not learning (in any sense) about the sequence. Instead, they were learning about the frequencies with which the different stimuli occurred. This is information they may, if asked, have been able to report.

2.7.2. Prediction tests as measures of awareness. The second problem is that, even ignoring the above considerations, we cannot rely just on the subjects' informal reports as assessments of their state of awareness some seconds or even minutes previously. Two somewhat different strategies have been advocated with regard to using more sensitive tests of awareness, namely, recognition and prediction tests. We discuss recognition tests in the next section. Prediction tests, introduced by Nissen and Bullemer (1987), require the subject to try to predict the next element of the sequence; such a test was used by Willingham et al. (1989) in addition to their verbal report test. After the RT phase of their experiment, Willingham et al. instructed subjects to try to predict on each trial where the stimulus would next appear, with no requirement for rapid responses. Subjects simply chose response keys on each trial until they picked the correct one, at which point they would then try to predict the next stimulus. Across many blocks of this prediction task, the subject again has the opportunity to learn the sequence. Evidence for explicit knowledge of the rule appears as savings (compared with the control group) in the number of trials required to learn the sequence in the prediction task.³

The rationale behind the prediction task is that if subjects are instructed to try to predict events and are able to do so with above-chance accuracy, this is evidence of conscious knowledge, because their predictions must be based on conscious expectancies. As this task requires the subject to act on a conscious expectancy concerning which stimulus will appear next, it is apparently a test of awareness of elements of the sequence. This contrasts with the RT task, in which they have to respond as fast as possible to the current target. The prediction task is a good one in that it is irrelevant whether or not the subjects believe that their performance in the RT phase was being affected by the sequence (indeed, they may not even be able consciously to report having detected a sequence). All that matters is whether any evidence of savings emerges in the prediction task, for, according to the reasoning behind the task, such savings must be due to conscious information about the sequence.

More to the point, failure at the prediction task would demonstrate a subject's inability to draw consciously on information about the sequence, thereby supporting the contention that the information indeed is implicit. It is important to drawing such a conclusion that the prediction task satisfies the Sensitivity Criterion where verbal reports did not. The retrieval cues for the prediction task are virtually identical to those of the reaction time task. Hence we now have two tests that are almost identical,

but in one the subject's performance (i.e., RT) is measured and in the other awareness is assessed. This very much follows the rationale of recent experiments on unconscious perception (e.g., Merikle & Reingold 1992), where the test of awareness and the test of perception are designed to differ in little more than the instructions given to subjects. Although the temporal arrangement of stimuli and responses is different in the two tasks, and the response metrics are quite different, the prediction task nonetheless represents an interesting new procedure for assessing awareness.

What are the results obtained from studies using the prediction task? Willingham et al. (1989, Experiment 1) discovered that subjects they had classified as unaware on the basis of their verbal reports not only speeded up in the RT phase but also, according to Willingham et al., showed no evidence of awareness as assessed by the prediction task. Such a result appears to provide quite compelling evidence of implicit learning, even if Shanks et al.'s (1994) data suggest that learning probably involved frequency rather than sequence information. It is important to note that this dissociation of RT speedup and prediction performance only applies to subjects who have been selected as unaware on the basis of their verbal reports. Across all subjects (regardless of their verbally reported awareness), RT speedup and prediction performance tend to be closely associated, as experiments by Cleeremans and McClelland (1991) and Perruchet and Amorim (1992) have shown.

Willingham et al. compared the performance of their normal sequence/no-knowledge subjects in the prediction task with that of a "no-training" group who had not received the RT phase at all. This comparison was the critical evidence that the sequence learning in the normal-sequence/no-knowledge subjects was implicit. However, there are three problems with these results. First, although Willingham et al. claimed that there was no evidence of savings on the prediction task in their "no-explicit-knowledge" subjects, close inspection of their data reveals that these subjects did perform at a better level than naive subjects, though not significantly so. Over each of the first six sets of 10 trials of the prediction task, performance was better in the normal-sequence/no-explicit-knowledge group than in the control group by about 5% in each set (Willingham et al. 1989, Fig. 3). On the first block of trials, the normal-sequence/no-explicit-knowledge group scored 42.6% correct and the control group 38.7%. Although small, this trend is as much evidence for savings as it is for a dissociation between awareness and learning. A similar conclusion may be drawn for the data reported by Hartman et al. (1989), where small but consistent savings are also apparent.

Second, Perruchet and Amorim (1992) pointed out that Willingham et al. did not instruct their subjects that the stimulus sequence in the prediction phase would be the same as that in the RT phase. Hence subjects may not have been maximally motivated to show transfer savings in the prediction task. The third and final problem is that Shanks et al. (1994), in their replication and extension of Willingham et al.'s study, obtained savings that were of a statistically significant magnitude. Shanks et al.'s normal-sequence/no-knowledge subjects performed much better (mean 5.7 correct predictions) than the no-training con-

trol subjects (mean 2.7) across the first 10 trials of the prediction phase, indicating that at least some of the knowledge they had acquired in the RT phase but were unable to report verbally was available for transfer to the prediction task. In sum, we conclude that the Willingham et al. study has failed to establish unconscious sequence learning.

A number of other studies have also used the sequence-learning task. Several of these have adopted Willingham et al.'s procedure of classifying some subjects as unaware on the basis of their verbal reports and then examining their prediction task performance. Others have sought to obtain different dissociations between RT speedup and prediction performance. Whatever the strategy used, we suggest that claims for implicit learning in these studies (e.g., Cohen et al. 1990; Hartman et al. 1989; Howard & Howard 1989; Knopman 1991; Lewicki et al. 1987; Lewicki et al. 1988; Nissen & Bullemer 1987; Nissen et al. 1987; Stadler 1989) are difficult to interpret for many of the reasons we have raised concerning Willingham et al.'s experiment. These other studies either (1) fail to show that subjects have acquired any *sequence* knowledge in the RT phase, (2) show small but consistent trends toward savings in the prediction task in supposedly unaware subjects, (3) present control subjects in the prediction task with random rather than pseudorandom sequences, or (4) do not provide feedback in the prediction test and hence run the risk of inducing forgetting of the sequence, which will lead to an underestimation of conscious knowledge. Caution suggests that these studies do not warrant the conclusion of reliable sequence learning in the absence of awareness.

Rather than reviewing all of these studies, we consider two widely cited ones (Lewicki et al. 1988; Stadler 1989) that illustrate some of the problems. Lewicki et al. (1988) presented subjects with blocks of trials that were arranged into sequences of five trials. On each trial a target appeared in one of the four quadrants of the computer screen and the subject had to respond by pressing the key appropriate to that quadrant. RTs were collected from a total of 4,080 experimental trials experienced by each subject. On the first two trials target location was random, except that the target was never displayed twice in the same place. Target location on trial 3 was determined by what had happened on trials 1 and 2. If the movement on the first two trials had been horizontal, then the movement from trial 2 to trial 3 was vertical; if it was vertical, then the next was diagonal; and if it was diagonal, the next was horizontal.

Similarly, target location on trial 4 depended on target locations on trials 2 and 3, and target location on trial 5 depended on its locations on trials 3 and 4. The net effect was that target location on trials 3, 4, and 5 was entirely predictable from the underlying rules, but locations on trials 1 and 2 were random. Hence if the subjects were indeed learning something about the rules, this should have manifested itself in a significantly greater reduction in RTs across blocks on trials 3, 4, and 5 than on trials 1 and 2, and this is exactly what Lewicki et al. found (in fact, they took as their dependent measure the number of correct responses with latencies less than 400 msec). Also, when the rules were changed toward the end of training, reaction times increased on trials 3, 4, and 5 but not on trials 1 and 2.

Lewicki et al.'s (1988) subjects could apparently report next to nothing about the rules determining target location. "None of the subjects mentioned anything even close to the manipulated pattern of exposures" (p. 33), although eight of the nine subjects did seem to be aware that their performance had dropped when the rules were changed. Lewicki et al. concluded that the subjects had learned the rules determining target locations on trials 3–5 implicitly or unconsciously.

Perruchet et al. (1990), however, disputed Lewicki et al.'s (1988) conclusions. The criticism is that the set of possible events that could occur on trials 3–5 was more constrained than the set of events that could occur on trials 1 and 2. By analyzing the rules that determined the permissible transitions from one trial to another, Perruchet et al. were able to show in their replication of Lewicki et al.'s experiment that speedup in RT on trials 3–5 relative to trials 1 and 2 was mainly due to relative speedup only on trials 4 and 5; furthermore, it was almost entirely due to two factors. First, on trials 1 and 2, but not trials 3–5, there were some occasions when the stimulus moved back to a location from which it had just come; these backwards movements led to a slowing of RTs simply because they increased the unpredictability of the movement. Second, on trials 1 and 2 there were infrequent horizontal movements, which again led to a slowing of RTs. On trials 4 and 5 horizontal movements were not permissible. Rather than learn rules such as, "If the movement from trial 1 to trial 2 was horizontal, then the next movement will be vertical," subjects need only have learned that the possible transitions had widely different overall probabilities. Low-probability transitions tended to occur on trials 1 and 2 and hence led to slower RTs.

Clearly, Lewicki et al.'s data fail to show that the knowledge that subjects could articulate in their verbal reports was in any way poorer than the knowledge that underlay their RT speedup. This fails to meet the Information Criterion and hence Lewicki et al.'s belief that knowledge of the composition rules was necessary for RT improvement is almost certainly not correct. Instead, relative RT speedup is simply due to subjects' learning that certain movements of the target occurred with low probability. As Perruchet et al. say, "The fact that subjects do not articulate any of the composition rules no longer applies if improvement in performance turns out to be unrelated to this kind of knowledge" (1990, p. 497). Furthermore, "Subjects' reports on the frequency of occurrence of particular target transitions would have been rejected as irrelevant to the actual manipulation" (p. 512).

Perruchet et al. did not assess their subjects verbally, but on the basis of Lewicki et al.'s results we may assume that on such a test they would also be classified as unaware. However, a verbal report test assessing experimenter-defined sequence knowledge may fail the Sensitivity Criterion. Perruchet et al. therefore used a prediction test in their study to see whether the subjects were indeed unaware. Toward the end of the experiment, they required one group of subjects to predict where they thought the target was due to appear. These predictions only had to be made intermittently, because there were between 4 and 11 normal RT trials between successive prediction trials. On such a trial, a question mark appeared in the center of the screen and subjects pressed the button matching the location of the target that seemed most probable.

Perruchet et al. found that subjects performed at significantly better than chance levels on the prediction trials. On trials 3, 4, and 5 they averaged 55.6% correct predictions, against a chance value of 33.3% (which assumes that subjects have learned – explicitly – that the target never appears in the same location on consecutive trials). On trials 1 and 2, only 29.7% of predictions were correct. If we take prediction responses to be conscious "reports" by the subjects of their expectancies about target location, then these results contradict Lewicki et al.'s (1988) claim that their study supported a dissociation between implicit and explicit knowledge. On the contrary, subjects seemed just as able to "report" target location as they were to show selective RT improvements.

We have already described Lewicki et al.'s (1987) experiments in section 2.3.2 and the ways in which they failed to meet either the Information or the Sensitivity Criterion. Briefly, the verbal report is likely to be an insensitive measure, and subjects were likely to be using a simpler form of knowledge to perform the task than Lewicki et al. considered. Stadler's (1989) replication is worth considering, though, because it made a concerted attempt to meet these criteria. Each trial consisted of the presentation of a target item in one of the four quadrants of a computer screen, designated A–D. The choice of target location on each trial was nonrandom, and the question was whether the subjects would be able to detect this nonrandomness. Subjects were presented with sequences of seven trials, with rules constructed so that target location on the seventh trial could be predicted from its location on trials 1, 3, 4, and 6. On each of the first six trials, the digit 6 appeared on its own in one of the quadrants of the screen, but on trial 7 (the "complex" trial), it was embedded in a display containing 36 digits. Reaction time on the seventh trial was the measure of interest. Again, the rules specifying target location were deterministic: thus, if the target appeared in locations C, A, D, and B on trials 1, 3, 4, and 6 respectively, then on trial 7 the target would be in location A.

Stadler (1989) repeated Lewicki et al.'s (1987) finding that RTs on the target trials decreased significantly across a large number of complex trials. To assess awareness, Stadler used a prediction task in which subjects were free to use whatever knowledge they had acquired, be it of the rules or of fragments of sequences, to predict target location. The task thus meets the Information Criterion, and because it reinstates much of the learning context, should also go a good deal of the way toward establishing equal sensitivity to conscious information in the awareness and performance tasks. In the prediction task, Stadler presented each of the four subjects with 48 sequences of seven trials, with target location on the seventh (complex) trial being determined by its location on the first, third, fourth, and fifth of the simple trials. On the complex trial, instead of presenting the target and distractors and requiring the subjects to locate the distractor as quickly as possible, Stadler presented a question mark in each possible location, and the subjects had to guess in which quadrant the target would appear. No feedback was given.

Stadler found that the four subjects made correct predictions on, respectively, 13, 11, 11, and 11 of the 48 sequences, where chance performance was 12/48. Clearly, there is no hint here of transfer to the prediction

task. But in the absence of feedback on the prediction sequences, it seems to us that there is a very substantial likelihood that over the 48 test sequences, subjects would have forgotten a large part of whatever knowledge underlay their RT performance. Without feedback, 48 sequences (including 288 simple trials) represents a vast amount of interfering information. To be fair, Stadler could not, with his design, have given feedback, because better-than-chance performance would have been ascribable equally to savings from the RT stage and to new learning of the sequences. For this reason, designs in which relative savings in the unaware group are compared to those in a novel control group are much to be preferred; they allow a sensitive savings test to be administered without the problem of the forgetting of the sequences. Alternatively, Stadler could have interspersed his prediction trials throughout further RT trials in order to offset any forgetting on the former. At any rate, acceptance of the null hypothesis (of no transfer to the prediction trials) would be better warranted if we knew that there was no hint of transfer on the first few trials, but Stadler did not present these data.

In the meantime, we suggest that cautious readers should not interpret these results, or any of the sequence learning results we have reviewed, as convincing evidence for learning without awareness.

2.7.3. Recognition tests. As an alternative to prediction tests, several researchers have argued that an appropriate and sensitive test of the subject's explicit knowledge – particularly if that knowledge is fragmentary – is to use a recognition memory test. Using the Nissen and Bullemer (1987) task, Perruchet and Amorim (1992), for instance, presented subjects in the test phase with 4-trial sequences such as DBCA, which had either been part of the training sequence or not, and required them to respond exactly as they had in the study phase prior to making recognition decisions about the sequences. The results indicated that the old and new sequences could be discriminated, and furthermore, the recognition scores correlated extremely highly ($r = 0.821$) with RTs. On this basis, there is little evidence that sequence learning was unconscious. However, Perruchet and Amorim did not divide their subjects into aware and unaware groups, and it remains possible that their subjects included an unaware subgroup in whom recognition performance was at chance.

More interesting are data reported by Willingham et al. (1993). Using the standard 4-stimulus task, one-half of the subjects saw a random sequence and one-half saw a 16-trial repeating sequence, equated for overall stimulus frequencies. After the learning phase all subjects were given a detailed verbal questionnaire consisting of five questions and were also given a recognition test, in which five 16-trial sequences were shown, and the subjects had to rate how likely it was that each sequence had been the one used in the RT phase. The mean rating for the distractor sequences was subtracted from that given to the target sequence to yield a recognition measure.

Consistent with the idea that sequence learning was unconscious, Willingham et al. (1993) found only small and nonreliable correlations between the awareness measures (the verbal and recognition measures) and RT speedup in the study phase. This result also confirms, of

course, that the recognition measure indexes something different from the performance measure. However, when Willingham et al. examined their individual subjects' data, they found that only two out of 45 sequence subjects could be classified as genuinely unaware of the sequence: the remainder all scored better than the median of the random subjects on one or more of the awareness measures. Thus it is only for these two subjects that we have any evidence of unawareness. When the RTs of these subjects were examined, they did indeed improve significantly more across trials than the random subjects, but interpretation is complicated by the fact that they started with abnormally slow RTs and ended the training phase with RTs similar to the random subjects. Finally, as Perruchet and Gallego (1993) point out, in a sample of 45 subjects, two of them might be erroneously misclassified as unaware simply because of unreliability in the awareness test. Obviously, a much larger sample is needed before strong conclusions can be drawn.

2.7.4. What is learned in sequence-learning experiments?

Perhaps it should come as little surprise that, in general, RT performance and awareness tend to correlate (see Perruchet & Amorim 1992). Indeed, it is well established that in choice RT tasks, when subjects make a correct prediction about which stimulus will appear on the next trial, their RT on that trial will be much faster than if they had made an incorrect prediction (e.g., Simon & Craft 1989). But what exactly is the nature of the knowledge that subjects acquire in a sequence-learning experiment? In attempting to answer this question, perhaps we can better understand why there is such a temptation to regard the learning as unconscious.

In the studies we have described, the experimenter arranges that the location of the target is governed by a complex rule or set of rules. For example, in Lewicki et al.'s (1987) and Stadler's (1989) experiments, one of the rules says that if the target appeared in locations C, A, D, and B on trials 1, 3, 4, and 6, respectively, then on trial 7 the target will be in location A. In some studies (e.g., Lewicki et al. 1988) it appears that the experimenter is assuming that significant learning in the RT stage must occur because subjects learn those rules in their entirety. But the Information Criterion cautions us to examine closely whether RT speedup can be due solely to learning of the entire rule: Might performance not simply be due to learning of more fragmentary information? In the Lewicki et al. (1987) study, might RT improvement not arise just from the subjects' learning contingencies between, say, target locations on trials 4 and 6 and target location on trial 7?

The best evidence to date suggests that learning fragments of the training sequences is probably sufficient to explain the available sequence-learning data (see Perruchet, in press). Cleeremans and McClelland (1991, Experiment 1) were able to compare RTs to targets that could only be predicted by knowing the previous three elements of the sequence. They found a reliable difference between RTs to targets that conformed to the rules compared to those that did not, indicating that subjects could indeed maintain three items of temporal context. However, no evidence emerged that they could maintain four items of context. If such a result is generally valid – though the size of the temporal context that can be

maintained is likely to be influenced by the exact experimental procedure – it would be very unlikely that Lewicki et al.'s (1987) subjects could learn rules requiring knowledge of six items of context.

Cleeremans and McClelland (1991, Experiment 2) obtained more direct evidence of the constraints on the amount of context subjects can maintain. They set up a task in which the location of the stimulus on a target trial could *only* be predicted by knowing where the stimulus had appeared four (and sometimes more) trials previously. Confirming the results of their first experiment, no evidence emerged that subjects could learn this long-range contingency, even when presented with a massive 60,000 trials. Instead, subjects appeared to be able to predict target location by reference only to the last one, two, or three locations. Thus, asking subjects to report entire rules risks falling foul of the Information Criterion.

2.7.5. Objections to prediction and recognition tests.

Given the tentative conclusion we have reached, namely, that prediction tests do in general reveal savings and that recognition tests so far have not yielded clear dissociations from RT speedup, defenders of unconscious learning might say that these tests are not truly tests of awareness. In all probability, they might argue, subjects do not “know” why they press certain keys in the prediction task; perhaps their fingers just get pulled toward certain keys. According to some of Lewicki et al.'s (1988, p. 33) subjects, describing their RT performance, “their fingers were doing the job by themselves.” Perhaps this happens as much for prediction as it does for the RT task. Similarly, perhaps subjects say “old” in a recognition test not because they are aware that the test sequence was part of the study sequence but because of perceptual fluency (see Perruchet & Amorim 1992). As we have mentioned already, there is no inherent reason why any behavioral measure should be influenced exclusively by conscious processes. In the language of the subliminal perception literature, maybe these tasks are not *exclusive* (Reingold & Merikle 1988).

If the prediction and recognition tests are not pure measures of awareness, then any conclusion based on them – e.g., that supposedly unaware subjects are in fact aware of the sequence – is called into question. But there are at least three reasons to doubt that unconscious processes do play a significant role. First, remember that in the prediction task, a response is required that is *different* from the one that was performed in the RT task. In the prediction task subjects respond to the *next* stimulus, whereas in the RT task they respond to the *current* stimulus. So, if the claim that prediction responses are under unconscious control is correct, subjects would make erroneous responses on the prediction task. This does not seem to be the case to any significant degree. Second, if some unconscious process is contributing to recognition performance, it seems that RTs should be faster in the recognition test for sequences that were part of the original sequence than for those that were not. Perruchet and Amorim (1992) were unable to find any evidence for this. Finally, Willingham et al. (1993) used three different versions of the recognition task. In one, subjects responded to the test sequences just as they had in the study phase, and then gave recognition ratings. Other subjects merely observed the stimulus sequence

prior to making a recognition judgment, and a third subgroup saw the sequence presented in the form of the digits 1–4 rather than in terms of screen locations. The latter two procedures should rule out perceptual fluency as the basis of judgments, yet Willingham et al. observed no difference between the three tests. We believe this strongly suggests that the recognition test is a genuine measure of awareness of the sequence.

Of course, if the prediction task and recognition tests cannot be treated as tests of awareness, we have no recourse other than to examine the subject's verbal reports as the only available index of explicit knowledge, and this, we have argued, is unsatisfactory because it precludes meeting the Sensitivity Criterion. Instead, some authors have suggested that we should abandon the narrow version of the dissociation paradigm that underlies these implicit learning studies (see sect. 2.1) and try to demonstrate *qualitative* differences between conscious and unconscious learning. It could be argued that such a difference would exist if the information that can be expressed in performance (RTs) increases dramatically across learning trials while the information available to awareness only changes marginally. Awareness may not have to be entirely absent. Thus, presumably, one might say that subjects in the Willingham et al. (1989) experiment were able to project more information in their RTs than in their predictions (where the savings tended to be only about 5%), and thus, even though prediction performance was better than chance, this is still evidence for a distinction between implicit and explicit learning. But without measures of (1) the amount of information that is transmitted when a subject shows an RT speedup of x msec, and (2) the amount of information the subject is transmitting when their prediction performance improves by $y\%$, it is very hard to assess such objections. We have no model for how much information is being conveyed by these different measures (for discussion of the same point, see Nelson 1978; Reingold & Merikle 1988).

It is exactly for this reason that one normally looks for cases where awareness of some variable is absent but performance is significantly affected by that variable, because even if we have no formal description of information, we know that a variable (which represents information) is being conveyed in performance but not in reports. For example, we would know that a variable such as the predictability of a sequence is affecting RT but not reports, and therefore RTs are conveying more information. Thus it is impossible to invalidate the null hypothesis (that performance and prediction convey the same amount of information) unless something can be done to show that more information is being conveyed in the implicit than in the explicit measure.

2.7.6. Conclusions. Sequence-learning studies have used prediction and recognition tasks as indices of awareness. These tasks reproduce the stimulus context of the learning stage (hence addressing the Sensitivity Criterion), and can be performed at above-chance levels, whether the subjects' knowledge is of fragments or of the complete sequence (hence meeting the Information Criterion). However, contrary to claims that sequence learning is unconscious, the results to date suggest that in most cases subjects *are* aware of the relevant knowledge, and that their knowledge consists of fragments of the training

sequence. We believe that no convincing evidence of implicit learning has yet emerged in sequential RT tasks; nevertheless, this is a very promising field of research that may in the future allow more positive conclusions.

3. Encoding instances versus inducing rules

In the introduction we raised two fundamental issues concerning implicit learning: that of consciousness during learning and the type of knowledge acquired. By crossing these factors we obtain four hypothetical learning systems: unconscious instance/fragment learning, unconscious rule learning, conscious instance/fragment learning, and conscious rule learning. This review has shown that the evidence for unconscious learning of any sort is highly questionable; we accordingly conclude that unconscious learning is unsupported in general. It is time, then, to turn to our second dimension for characterizing dissociable learning systems, namely, the content of the acquired knowledge, and to assess the rules versus instances distinction within the domain of conscious, explicit learning.

In the following discussion we use the term "implicit learning tasks" to refer to the sorts of tasks reviewed in section 2.

3.1. Evidence from studies of concept learning

Perhaps the most important conclusion from our discussion so far is that performance in a variety of different tasks can be well accounted for by reference to fragmentary knowledge or knowledge of instances rather than to abstract rules. In the present section we pursue this idea by showing that to a large extent human performance in more traditional concept-learning studies can also be well understood in such terms. However, the concept learning literature leads us also toward more compelling evidence that people can genuinely learn rules. Thus we begin to see a characteristic that does distinguish different learning systems: whether the knowledge acquired is of instances or rules.

A view of concept learning that had been popular for many years prior to the 1960s was that learning a concept involves the acquisition of a rule specifying the features necessary and sufficient for membership of that category. However, when Rosch (1975) argued that for many natural categories it was impossible to specify the necessary and sufficient features, and when Posner and Keele (1970) showed that subjects could learn to classify random dot patterns that had not been generated by deterministic rules, research began to be dominated by the alternative view that concepts are represented by prototypes. A prototype is an abstraction from a set of training stimuli that corresponds to their central tendency.

On the prototype view, category membership is determined simply by computing which of a series of stored prototypes the test stimulus is closest to. As in the rule-based account, in prototype theories the subject is assumed to abstract something from the training stimuli and not to retain information about those specific instances.

3.1.1. The role of instances. The view that the learning of a concept could be based on little more than the encoding

of the separate instances that fall under the concept began to emerge in the late 1970s in the work of researchers such as Lee Brooks and Douglas Medin. They observed that, contrary to what would be expected on a prototype account, subjects do appear to retain information about training instances, in that studied instances can bias subsequent classification decisions (e.g., Brooks 1978; Homa et al. 1981; Malt 1989; Medin & Schaffer 1978). On the basis of such findings, Medin and Schaffer (1978) proposed that one component of a concept is simply a set of memorized exemplars or instances.

Although Medin and Schaffer assumed that both instance storage and prototype abstraction could occur during the learning of a concept, subsequent studies have shown that performance in a great many category-learning studies can be understood in terms of instance storage alone (for a historical review, see Medin & Florian 1992). We can illustrate the power of instance-storage theories by considering the results of a study by Shin and Nosofsky (1992), who examined category learning with dot patterns. Shin and Nosofsky first performed a multidimensional-scaling analysis on subjects' judgments of pairwise similarity for the patterns, which yielded coordinates in psychological space for each of the patterns. Other subjects were then trained to classify some of the patterns into three categories and, following that, they were tested on their classification decisions for the remaining patterns.

The instance view proposes that subjects memorize the actual exemplars seen during training and base their classifications on the similarity between a test item and stored instances. Since Shin and Nosofsky knew the psychological coordinates of all of the patterns, they were able to compare subjects' decisions to test patterns with the predictions of a model that assumed that classification was determined solely by similarity to memorized instances. They found a remarkable degree of concordance between predicted and observed classifications, with over 95% of the variance in the observed classifications being accounted for in a 1-parameter model. A prototype theory, assuming that the training instances formed the basis for an abstracted prototype, performed much more poorly in predicting responses.

The implications of such results for implicit learning cannot be overemphasized. The stimuli used in typical categorization experiments are every bit as complex and difficult to label verbally as are the stimuli used in implicit learning experiments, so the powerful evidence for instance storage that emerges from categorization experiments should encourage us to take very seriously the possibility that the encoding of instances is a major factor in implicit learning experiments as well.

3.1.2. Evidence for rule induction. Despite the wealth of evidence in favor of instance theories of concept learning, we argue that it is also possible for people to classify objects on the basis of a rule or hypothesis. As a particularly dramatic example, consider the evidence for a difference between instance learning and rule learning in the sexing of day-old chicks (Biederman & Shiffrar 1987). It has been estimated that professional sexers, trained with feedback on instances, require 2.4 months of solid practice to reach 95% accuracy. However, naive subjects trained on one simple rule immediately achieved 90%

accuracy. Of course the simple rule misses the rare and subtle exceptions that instance learning can provide, so further accuracy gains will be difficult. On the other hand, the initial difference in training time is immense.

Laboratory demonstrations of contrasts between rule and instance learning have been provided in a number of studies (e.g., Allen & Brooks 1991; Kemler Nelson 1984; Nosofsky et al. 1989; Regehr & Brooks 1993; Smith & Shapiro 1989; Ward & Scott 1987). Consider, for instance, Nosofsky et al.'s (1989) Experiment 1. In this experiment, the stimuli were semicircles with an interior radial line: 16 stimuli were constructed from the combination of four sizes of semicircle (1 . . . 4) with four angles of inclination of the radial line (1 . . . 4). In one condition, subjects learned to classify three of the stimuli into category 1 and four into category 2, and then were tested for their transfer performance across the remaining nine stimuli. The dependent measures were the overall probabilities with which subjects placed each of the 16 stimuli into categories 1 and 2.

Nosofsky et al. (1989) found that 97.9% of the variance in these combined classification probabilities across subjects could be accounted for by a quantitative model that assumed that stored instances were the only basis on which the decisions were made, an impressive fit that confirms the conclusion of the Shin and Nosofsky (1992) study. In addition, however, Nosofsky et al. compared their model to several rule-based descriptions. Specifically, it would have been possible for subjects to learn the categorization problem by inducing a rule for partitioning the stimulus space into response regions. For instance, one rule that would have correctly classified the training stimuli was the rule: *a stimulus is in category 2 if the value of angle is 1, or the value of size is 1, or the value of angle is 4; otherwise the stimulus is in category 1*. Nosofsky et al. constructed a variety of such set-theoretic rules, but found that none of them fitted the overall classification performance of the subjects nearly as well as their instance model. From such data it would appear that subjects rely on nothing more than stored representations of the training items in classifying stimuli.

When they looked at the behavior of individual subjects, however, Nosofsky et al. (1989, Experiment 1) found that some subjects' classification responses conformed to patterns that were quite unlikely according to an instance theory, and yet they matched fairly simple rules. Thus, at the level of individual subjects, some evidence for rule following rather than generalization to stored instances did emerge. In their second experiment, Nosofsky et al. chose two rules that could be used to classify the stimuli accurately, and gave two different groups of subjects explicit instructions to follow one of the two rules in classifying the stimuli. Here, the instance theory failed dramatically, accounting for only 82.4% and 40.9% of the variance in responses for the two groups. In contrast, 99.5% and 93.6% of the variance in the subjects' classifications in the two groups were accounted for by the rules themselves. Similar results were obtained in their Experiment 3, using a different set of stimuli. Thus, here we have clear evidence that subjects are able to learn an abstract rule and hence need not rely just on stored instances.

An even more compelling demonstration of the inadequacy of pure instance storage comes from recognition

memory data that Nosofsky et al. collected during the test phase of their experiments. Although in Experiment 1 subjects had no difficulty recognizing the old training stimuli and discriminating them from new test stimuli, when subjects were explicitly instructed to use a rule to classify the stimuli in Experiment 2, no evidence emerged that the subjects could remember which test stimuli had been training stimuli. The implication of this result is that these subjects had encoded nothing in the training stage except the rule: they had not encoded any of the instances. This remarkable finding shows that when appropriate conditions are established, subjects can indeed learn an abstract rule from exposure to instances. Of course, just because subjects had to be given rule-following instructions in Nosofsky et al.'s studies before they would actually engage in rule following does not mean that this will always be necessary. In fact Nosofsky et al.'s stimuli make rule following difficult in that the stimuli do not readily lend themselves to verbal descriptions.

Other compelling evidence for rule learning, from tasks using more complex stimuli, has been reported by Allen and Brooks (1991) and Regehr and Brooks (1993). The rationale of the experiments was as follows: suppose that subjects learn to classify stimuli in a situation where a simple, perfectly predictive classification rule exists and they are then tested on transfer items that vary in similarity to the training stimuli. Observed behavior with the transfer items can be of two contrasting types: (1) transfer items similar to old items from the opposite category (bad transfer items) may be classified as quickly and as accurately as items similar to old items from the same category (good transfer items), or (2) the bad transfer items may be classified much less rapidly and accurately than the good transfer items.

The first case above would be consistent with classification being determined by the speeded application of a rule. In this case all that matters is whether the rule assigns the transfer item to one category or the other. Whether the item is similar or not to a training instance, and whether that instance was in the same or a different category, should be immaterial. On the other hand, the second outcome described above would be consistent with categorization on the basis of similarity to training instances, and there would be no need to cite a rule as being part of the classification process.

Allen and Brooks (1991) and Regehr and Brooks (1993) obtained evidence that both sorts of outcome can occur, depending on the type of stimuli used and the precise nature of the task. They trained subjects to classify animals into two categories. The animals varied in terms of five binary-valued dimensions: body shape, spots, leg length, neck length, and number of legs, but only three of the dimensions were relevant. In some experiments subjects were explicitly told the classification rule (e.g., category 1 is defined by the conjunction of long legs, angular body, and spots). Evidence for rule learning – no difference in latency or accuracy in classifying a new item similar to a training item and in the same category versus a new item similar to a training item but in the opposite category – was related to a number of factors. For example, rule learning was more likely to be the controlling process when subjects were actually told the rule prior to the task, although this was not a necessary condition. It

also depended on the nature of the stimuli used: highly individual stimuli tended to elicit instance-based rather than rule-based classification, and stimuli composed of interchangeable features tended to elicit more rule-based classification behavior.

Additional factors that appear to determine the balance between rule learning and instance learning have been investigated. Smith and Shapiro (1989) found that rule learning was less likely when subjects had to perform a secondary task during the learning stage and when the training stage was conducted as an incidental learning task. Smith and Kemler Nelson (1984) found that rule learning was less likely in a speeded than an unspeeded learning task, and also that there seems to be a developmental trend in rule learning: in situations where adults classify according to a rule, children often do so on the basis of similarity to training instances. In sum, laboratory studies have established the reality of rule- or hypothesis-based concept learning and have begun to identify the circumstances that determine when it predominates.

3.2. Rule induction from artificial grammars

In our earlier discussion of artificial grammar learning, we suggested that performance could be well understood on the assumption that subjects principally learn about whole strings (instances) and about legal substrings. Although some evidence for abstraction exists (e.g., Altmann et al., in press), such a view denies that knowledge of abstract grammatical rules plays a major part in performance. However, when subjects are told prior to seeing them that the strings are rule governed, and particularly if they are taught something about finite state grammars (Mathews et al. 1989; Reber et al. 1980), a pattern of results rather different from that seen under normal conditions occurs. Under these conditions, subjects appear to engage in a very explicit form of rule induction or hypothesis testing that produces reportable rules. The dissociation between performance and verbal reports does not occur under these conditions.

Reber et al. (1980) extensively instructed subjects on the nature of finite state grammars before, during, or after presenting the training strings: When instruction occurred before training, subjects performed well on the grammaticality task but showed evidence of having learned unrepresentative rules: rather than random errors from guessing, their errors were consistent. Nonrandom mistakes fit with the idea that subjects were trying out various hypotheses in turn. At test time subjects will have some hypothesis for evaluating test strings, but it may be incorrect and therefore produce consistent errors. Further supporting this claim, Reber et al. found that a well-ordered presentation of training strings, making the structure of the underlying grammar more salient, assisted learning only in the case where the explicit instructions occurred early in training. The conclusion is that the explicit instructions promote an explicit hypothesis testing strategy, and that well-ordered string presentation was particularly helpful for this learning strategy.

Mathews et al. (1989) found that the nature of the grammar also affected subjects' ability to learn it via an explicit strategy. Finite state grammars are difficult to learn explicitly: instructions that promote hypothesis testing but provide no information about the nature of

grammar rules, such as instructions to "find the rules," fail to produce good learning (Perruchet & Pacteau 1990; Reber et al. 1980). Indeed, Mathews et al. found that subjects given such instructions learned no better than subjects given incidental learning instructions. Other types of grammars, however, are easier to learn explicitly. For example, subjects in Mathews et al.'s study readily acquired a biconditional grammar in which strings consisted of two sets of four letters, with three rules mapping the letters between the sets. Mathews et al. observed far superior learning in subjects instructed to search for rules than in subjects given incidental learning instructions. Moreover, in contrast to subjects given rule-searching instructions, subjects given incidental instructions did not acquire this grammar at all, a result that strongly argues for the existence of independent learning systems.

How does explicit hypothesis testing differ from the memory-based processing that is the norm in implicit learning paradigms? Smith et al. (1992) compiled a set of eight characteristics of rule following behavior. For example, a defining feature is that rules ought to apply equally well to familiar, unfamiliar, and abstract problems or stimuli (as in Allen & Brooks 1991; Regehr & Brooks 1993). Smith et al. review numerous experiments that demonstrate subjects following rules such as *modus ponens* and the law of large numbers.

Turning to rule *learning*, Smith et al. point to subjects' learning protocols and their ability to report intermediate hypotheses as evidence for rule learning. Indeed, explicit hypothesis testing presumes that the hypotheses are conscious and reportable. Protocols indicate that generating and testing these hypotheses is a slow, labored, and conscious process. Lea and Simon (1979) describe a general framework for understanding the cognitive processes involved. Hypothesis testing, they claim, is a form of problem solving involving search through two problem spaces: the space of hypotheses and the space of experiments or instances to test hypotheses.

Studies of series completion, a task that is similar to artificial grammar learning, provide a clear example of the characteristics of the hypothesis-testing strategy (Kotovsky & Simon 1979). In this task, subjects are shown a single string of letters that contains a short repeating pattern (e.g., *MABMCD*) and are asked to continue the series. Their protocols clearly indicate that they cycle through gathering evidence from the string and generating hypotheses to fit the data. In addition, subjects' occasional mistakes are consistent with their last-considered, but incorrect, hypothesis. They will then produce consistent errors as did subjects in the grammar learning studies (Reber et al. 1980). Finally, unusual hypotheses are difficult to generate and therefore unlikely to be discovered; this effect has been shown in several hypothesis-testing paradigms (Bruner et al. 1956; Klahr et al. 1990; Klayman & Ha 1989; Wason 1968). This unsurprising effect may explain the difficulty observed in getting subjects who are instructed to find the correct grammar rules actually to do so. These rules may be unusual enough that they are difficult for subjects to generate without detailed instruction.

How is rule generation and evaluation performed in this dual problem space? Lea and Simon (1979) offer several alternatives. The simplest alternative is to refrain from any feedback from the evaluation stage to the gener-

ation stage except to signal rejection when a hypothesis is disconfirmed by data. Rule generation is then essentially blind to the detailed data of experiments or instances. More sophisticated alternatives allow more information from experiments or instances to inform the generation process. In concept-formation tasks, Lea and Simon found that different subjects' behaviors conformed to different alternatives.

Protocols from Kotovsky and Simon's (1979) series-completion task provide some indication of what alternatives might underlie subjects' behaviors in implicit learning tasks. Kotovsky and Simon's protocols showed that subjects examined the series and discovered periodicities in the letters. Subjects then developed compact rules to describe those periodicities. The explicit hypothesis-testing strategy in grammar-learning experiments might operate similarly. As subjects observe strings, they may discover repetitive sequences. Subjects could then attempt to learn those sequences and develop rewrite rules to describe sequences of those sequences. Unfortunately, we do not know of any protocol data that examine this question.

To conclude, subjects performing explicit hypothesis testing demonstrate in their protocols clear reports of intermediate hypotheses and a relatively slow time course of processing. These phenomena stand in contrast to the obscure and superficial report and relatively fast time course of processing found in protocols of subjects engaged in typical implicit learning tasks. Though we have seen no data, we predict that subjects' behavior in grammar-learning experiments under explicit hypothesis-testing instructions will more closely resemble Bruner et al.'s (1956) and Kotovsky and Simon's (1979) subjects than subjects working under implicit instructions.

It seems clear, then, that there are two separate learning strategies available to subjects, and that these strategies can be invoked by differences in the instructions given to subjects. The rule-induction strategy is characterized by conscious effort to develop and evaluate hypotheses that are often unrepresentative of the actual grammar. This strategy can be invoked by detailed instructions and it can be facilitated by sensibly ordered examples. The "instance" strategy is invoked by instructions simply to memorize or observe the training strings, and it does not seem to be affected by the presentation order of training strings (Reber et al. 1980). Under this strategy subjects encode the whole strings and their basic features – such as pairs and triplets. Both strategies appear to be conscious, but the contents of consciousness vary. The rule-induction strategy trades in hypotheses and is well characterized as problem solving, but the instance strategy trades only in simpler data and seems well characterized as memorization.

3.3. Models for implicit learning tasks

How do computational models fit with the results we have discussed? We first consider the computations involved and then turn to the issue of consciousness.

Artificial grammar research has provided the most detailed analysis of what types of knowledge are acquired during implicit learning tasks and what processes use them at test time. There is memory for whole strings for recognition and similarity judgments, and substring frag-

ments for piecemeal familiarity judgments. Distributed-memory models in general and Parallel Distributed Processing (PDP) models in particular are well suited to these knowledge types and tasks. Networks are capable of learning both types of knowledge simultaneously and in the same set of weights. Dienes (1992) compared a variety of distributed- and memory array models on a set of grammar-learning data and found that an autoassociative network using the delta rule for learning fit the data best. The less competitive alternatives were connectionist models using the Hebb rule, and three memory array models.

The autoassociative network model consisted of a single layer of units that were completely interconnected. Each unit coded for one letter at a particular position in a string. Accordingly, there were five units (one for each letter) for each of six string positions. A variety of other local and distributed encoding schemes were also tested, but no great qualitative differences were found. The model was trained on the set of 20 training strings shown to subjects, and the delta rule (Widrow & Hoff 1960) was used to change the weights until each training pattern could reproduce itself over the input units. In one set of simulations, the model was trained on each stimulus the same number of times as were subjects. Each training trial, therefore, was equivalent for subjects and for the model. The model was tested on the grammaticality task by requiring it to compute a response to each test string. If the model could correctly reproduce the input with minimal error, the response was called "grammatical." If the model reproduced the input with a large error, the response was called "ungrammatical."

The delta rule model produced a high level of performance on the grammaticality test. It also produced the same ratio of random errors to consistent errors as did subjects. Critically, its ranking of the grammaticality of the test strings correlated highly with subjects' rankings. No other model could produce this correlation.

A somewhat similar account has been proposed for sequence learning. Cleeremans and McClelland (1991) developed a simple recurrent network model to simulate their human sequence-learning data. Since much of the benefit in reaction times during the sequence task is presumed to come from predicting the next stimulus, the model was designed to predict the next stimulus at each point in the sequence. The input layer of the model encoded one stimulus of a sequence at a time and used one unit for each of the possible stimuli in the "grammar" of stimulus sequences. As the model stepped through a sequence, the input layer successively encoded each stimulus in turn. At each step, activation from the input layer was fed forward through the model to produce activations in the output layer to represent the model's prediction about the upcoming stimulus. During training, the weights between units were adjusted so as to produce accurate predictions.

The model's behavior corresponded to the human data in three ways. First, after training, the model performed at the same level as subjects on the sequence task. Second, the model learned at the same rate as subjects. Cleeremans and McClelland drew a correspondence between single training trials for subjects and single training trials for the model. The model matched subjects' performance throughout the time course of training. Cleere-

mans and McClelland divided knowledge of the grammar into sets, depending on how many previous stimuli were required to make accurate predictions at each point in a sequence. Like subjects, the model first acquired knowledge allowing it to predict correctly in cases where only one previous stimulus was important. Both subjects and the model then slowly acquired the longer, more complex dependencies.

The third match between subjects and the model was the length of delay between predicted stimuli and their predicting context. Cleeremans and McClelland created a grammar with a set of stimuli that intervened between a predictive beginning stimulus and a predicted ending stimulus (e.g., 15553 vs. 25554). For both subjects and the model, only three intervening stimuli could be tolerated before the initial stimulus was forgotten.

As discussed above, concept learning also often appears to be mediated by the conscious memorization of instances. Recently, PDP models have been developed to simulate human results on these tasks as well (Gluck & Bower 1988; Kruschke 1992; McClelland & Rumelhart 1985; Shanks & Gluck 1994). These models produce close quantitative fits to a variety of concept identification, classification, and recognition data. Like the other PDP models, they learn in a memorization-like fashion by encoding individual stimuli and modifying their weights in response to each stimulus to process that stimulus better.

What about more fragmentary information? Perruchet and Pacteau (1990) found that subjects recognized frequent bigrams better than infrequent bigrams. This result requires some representation of the frequency of bigrams. An important feature of distributed memories (Cleeremans & McClelland 1991; Dienes 1992) is that they produce frequency statistics, in the form of strengths of encodings, as a by-product of the memorization process. Such a process has two good points: it does not require any covert computations to produce useful knowledge, and the computations it does require – memorization – fit with subjects' reportable experience of the training tasks.

These PDP models, therefore, capture a variety of data from sequence-learning, grammar-learning, and concept-learning tasks. Their details differ, but their basic representational abilities and modes of processing are the same. In this sense, they provide a unifying mechanism for learning both whole item and fragment knowledge and for simulating a wide range of cognitive phenomena from conditioning to sequence and grammar learning. Cleeremans (1993b) discusses at length some of the correspondences between human performance in implicit learning tasks and the behavior of connectionist models.

These models do not, however, say anything about consciousness. A standard idea in the psychological literature is to equate consciousness with the states of a processor, rather than with the processes themselves. In a PDP model, the states are the transitory activations of the units. "We assume that responses and perhaps the contents of perceptual experience depend on the temporal integration of the pattern of activation over all of the nodes" (McClelland & Rumelhart 1981, p. 381). In the models of instance memorization, these activations encode the features of individual stimuli. There is no representation within the models of rules or of the testing of

hypotheses. These representations, therefore, fit well with the idea that what subjects are doing in implicit learning tasks is memorizing the stimuli.

The ability to generalize and perform above chance on a grammaticality task arises from the fact that memorization takes place in a distributed system, where every memory contributes to every response. In a connectionist model such as Dienes's (1992), memorization occurs when the weights of the network are changed to encode a stimulus. Similar stimuli produce similar weight changes that establish strong connections over time, whereas dissimilar stimuli, or parts of stimuli, produce dissimilar weight changes that wash out over time. In this way, the central tendency and underlying structure of a set of stimuli is slowly captured by the weights. In both the Dienes (1992) and Cleeremans and McClelland (1991) models, the system is fully "aware" of the stimulus on each trial. In the latter model, the input layer encodes the stimulus one letter at a time, and that letter is fully activated. The input layer in Dienes's model encodes an entire string at once. While the stimulus on each learning trial is fully registered, the weights of the network start out very weak and slowly grow stronger with training, as knowledge is acquired. The models say, therefore, that subjects are fully aware of each stimulus, and that they slowly grow more aware of the underlying structure of the stimulus set as the weights grow stronger. But just as awareness of the structure may be very limited early in training, so will the extent to which that structure actually controls performance: "awareness" and performance will be correlated.

In contrast, connectionist models do not fit well with our understanding of the explicit hypothesis-testing strategies also found in the grammar-learning literature (sect. 3.2). Lea and Simon (1979) describe a possible mechanism, along the lines of the "general problem solver," to perform hypothesis testing. This learning strategy is different enough from the memorization strategy in process and results so that it is not surprising that it calls for a different sort of mechanism to execute the hypothesis-testing computations.

4. Learning and amnesia

In our evaluation of learning systems we have not considered at all the evidence from patients suffering from the classic anterograde amnesic syndrome (Squire 1992). Because these patients are generally considered to suffer from a learning or memory deficit, it is worth considering briefly how the data from this population relate to the data from normal subjects.

We know of no convincing data that would suggest that amnesics are capable of unconscious learning. It is very important to note that most studies have used episodic memory tests and hence fail to meet our condition for inferring unconscious learning from unconscious memory. For example, amnesic patients show normal or near-normal responding to a previously conditioned stimulus although they are apparently unable to recall the conditioning episode (Weiskrantz & Warrington 1979). But this is an example of case (2) rather than case (3) from section 2.3.1 (Fig. 1) and is therefore insufficient to establish that learning had been unconscious: that is, subjects may have been conscious of the reinforcement contingency but

unable to remember the episode in which it was learned. The amnesics, if asked, may well have been able to report a conscious expectancy of the US (unconditioned stimulus) given the CS.

It is true that a number of studies have shown that amnesics, like normals, can sometimes acquire information but be unable to report it verbally. Thus Nissen and Bullemer (1987; see also Knopman 1991) found that Korsakoff amnesic patients could speed up on a sequential RT task but report neither the sequence nor awareness of the existence of a sequence. However, as our discussion in section 2 revealed, that subjects are unable to report verbally some information they can otherwise be shown to have learned does not prove that the learning was unconscious, and this is as true for amnesics as it is for normal subjects. Nissen and Bullemer did not use a prediction test for conscious knowledge, and we know of no other studies that have attempted to use more sensitive tests of awareness.

This is not to imply that amnesics' poor performance on explicit memory tests is due to the insensitivity of such tests. On the contrary, this is unlikely to be the case because compared to normals, amnesics appear to be *selectively* impaired on explicit tests such as recognition (but see Ostergaard, 1994, for a contrary view). For example, Knowlton et al. (1992) reported that amnesics can perform as well as normals in judging the grammaticality of new strings generated from an artificial grammar while being very poor at recognizing the training strings. It is hard to see how such dissociations can be explained by differences in sensitivity between the grammaticality and recognition tests, because in that case the effect should be comparable for normals and amnesics. Instead, the evidence suggests that amnesics have a genuine problem with a certain class of memorial experience (i.e., episodic memory; see Humphreys et al. 1989, for a relevant computational model). However, there is no reason to believe that the problem involves awareness at the time of learning, or more specifically that results from amnesics provide any evidence for unconscious learning.

With respect to the learning of rules versus instances, there seems little doubt that amnesics can perform well (possibly at the same level as normals) in tasks that can be mastered by learning instances or fragments. Thus, in addition to Nissen and Bullemer's (1987) demonstration of learning in a sequential RT task, Knowlton et al.'s (1992) report of excellent artificial grammar learning in amnesics suggests that instance or fragment learning is intact. In contrast, evidence for genuine rule learning is sparse, which is not surprising given the problems involved in distinguishing instance from rule learning. Nevertheless, there is some evidence of conscious rule learning in amnesia. For example, Wood et al. (1982) report that although amnesics have some difficulty learning the Fibonacci rule, they are able to do so.

5. Summary

A variety of strategies have been used to assess, more or less directly, the content of a subject's awareness during a learning episode. Evidence for implicit learning would come from (1) demonstrations of learning with subliminal

stimuli, and (2) dissociations between task performance and measures of awareness such as verbal reports. The latter dissociations appear at first sight to provide evidence for implicit learning, because the inability to report the relevant stimulus relationship licenses the inference that learning may have occurred without awareness.

We have argued that there is little convincing evidence of learning with subliminal stimuli. On the other hand, with respect to the unconscious learning of stimulus relationships, we have documented a number of dissociations between performance and reports. There are nonetheless two reasons to question whether they establish implicit learning. First, there may be a relatively uninteresting explanation of such dissociations, stemming from the experimenter's failure to address the Information Criterion. If learning involves the acquisition of information *I*, but the experimenter is focusing on information *I**, then subjects may appear to be unaware of the relevant knowledge when in fact they are aware of it. Second, if the test of reportable knowledge fails to meet the Sensitivity Criterion, it is impossible to know whether a dissociation is genuine or merely reflects inadequate sensitivity to conscious information in the awareness test. Our review suggests that when attempts are made to use sensitive tests, dissociations do not emerge. Finally, if one is unprepared to accept recognition or prediction tests as measures of conscious knowledge, then it is difficult to see how the Sensitivity Criterion can ever be met. It is simply a fact of life that tests of verbal recall tend to be less sensitive to small amounts of knowledge than other behavioral measures. Perhaps alternatives to the simple dissociation logic adopted in almost all experimental tests of unconscious learning need to be explored, as they have in studies of unconscious perception (Reingold & Merikle 1988).

Our evaluation of the results that have emerged is similar to Holender's (1986) conclusion in this journal concerning unconscious semantic activation. Although there are some interesting pieces of evidence, a cautious approach would suggest that unconscious learning has not yet been satisfactorily established. Instead, there is substantial evidence for more than one conscious learning strategy and knowledge type. People certainly can learn and use rules, and they can also memorize instances and fragments. Researchers have begun to identify the factors (e.g., study time, stimulus properties) that are conducive to rule learning and instance learning.

Proponents of implicit learning, which is hypothesized to involve the unconscious learning of rules, have failed to demonstrate that it accurately describes a class of human learning abilities. On the contrary, human learning is almost invariably accompanied by conscious awareness, and in tasks such as artificial grammar learning, where learning is frequently thought to involve rule abstraction, performance is most often based on the acquisition of instances or fragments from the training stage.

ACKNOWLEDGMENTS

The first author was supported by the UK Medical Research Council, and by a SERC/NATO fellowship held at the Department of Cognitive Science, University of California, San Diego. We would like to express our considerable gratitude to colleagues who have commented on this work, especially Axel

Cleeremans, Anthony Dickinson, Zoltan Dienes, Don Dulany, Celia Heyes, Jonathan Kolodny, Phil Merikle, Pierre Perruchet, and Arthur Reber.

NOTES

1. The mere exposure effect may provide a paradoxical example of how instructions that seem to encourage the subject to rely on available conscious knowledge can result in their doing so to a lesser degree than instructions for a performance task. Several well-known studies (e.g., Kunst-Wilson & Zajonc 1980; see Bornstein 1992, for a review) have given subjects very brief presentations of geometrical figures prior to testing them with pairs of stimuli consisting of one old and one new figure. For each pair, the subjects had to indicate which one was old and which one was preferred. Results indicate that subjects will choose the old stimulus for their preference judgment although their ability to discriminate old from new stimuli in the recognition test is at or close to chance. On the face of it, this provides a powerful dissociation between results obtained from closely matched performance (preference) and awareness (recognition) tests. Ignoring the possibility that the procedure may fail to meet the requirement for inferring unconscious learning from unconscious memory (sect. 2.3.1), we suggest that the result is still not evidence for unconscious learning because, relative to the preference test, the recognition test may encourage subjects to *discount* deliberately a conscious source of information (viz., stimulus familiarity), because they know that familiarity can be a poor index of whether a stimulus has recently been seen. Evidence for this interpretation comes from a recent experiment by Merikle and Reingold (1991), who found that as testing continued, the recognition test gradually became more sensitive to the old/new distinction than the preference test (i.e., hypermnnesia occurred). This result is consistent with the idea that subjects began to rely on familiarity in making their recognition judgments because they realized that when they discounted it, they had no other cues on which to base their recognition decision.

2. A possible empirical way to determine whether unconscious influences do play a role would be to adopt Jacoby's (1991) process dissociation technique of asking subjects to provide letter continuations but to avoid any continuation that they had seen in the study phase. If some studied continuations were given, that would indicate the presence of unconscious influences.

3. The prediction task bears a striking but presumably unintentional resemblance to the task used by the parapsychologist Schmidt (1969) in his tests of precognition. In Schmidt's experiments, target selection on each trial was determined by random particle emission from a strontium-90 source, yet subjects were apparently able to predict at better-than-chance levels where each target would appear! In our discussion of data from the prediction task, we ignore the rather distressing possibility that subjects' performance may be influenced by such precognitive abilities.

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

Is learning during anaesthesia implicit?

Jackie Andrade

Medical Research Council Applied Psychology Unit, Cambridge CB2 2EF, England; jackie.andrade@mrc-apu.cam.ac.uk

Shanks & St. John (S&S) argue that learning during clinically adequate anaesthesia will be implicit, since the patient will be rendered unconscious by the anaesthetic. Unfortunately, this is not necessarily the case: adequate anaesthesia ensures that the patient will have no explicit memory for intraoperative events, but there is evidence that amnesia sets in prior to the abolition of awareness. For example, Andrade et al. (in press) assessed cognitive function in anaesthetised volunteers. They found that there was a dose at which subjects were still able to respond (by raising their arm) to target exemplars in a categorisation task or repeated words in a working memory test and yet scored at chance on a subsequent forced-choice recognition test for information presented during anaesthesia. It is therefore possible, at least in principle, for someone to be conscious during surgery and yet not remember being so on recovery.

Levels of arousal vary considerably during an operation, according to the ongoing surgical procedure, but there is not yet a reliable method for monitoring these changes, let alone for determining whether a given level of arousal is accompanied by awareness. The use of neuromuscular blockers to relax patients means that the traditional clinical signs, such as changes in breathing or muscle tone, are no longer effective for detecting fluctuations in arousal. The problem of detecting awareness is highlighted in a recent paper by Moerman et al. (1993). They investigated a number of cases in which patients had explicit recall of intraoperative events. Experienced anaesthetists could not reliably distinguish the anaesthetic records of these aware patients from the records of patients who had remained unconscious throughout surgery.

Research using Turnstall's (1977) isolated forearm technique suggests that up to 44% of surgical patients may experience some intraoperative awareness (Russell 1986). This technique involves tying a cuff around the patient's arm at the time of injection of the neuromuscular blocker so that the arm remains free to move. Patients are assumed to be aware if they raise that arm in response to a command to do so. This task meets S&S's Information Criterion, since patients are responding to the stimulus itself and not just to one (possibly irrelevant) aspect of it. The stumbling block is that anaesthesia – like drunkenness or sleepiness – reduced motivation, so there are cases where patients have reported, on recovery, that they were aware during anaesthesia but could not be bothered to let anyone know (Parkhouse 1960). Russell (1989) even reports a case in which the patient said she was aware of the commands but wanted to play a game with the anaesthetist by doing the exact opposite of what she was commanded. The isolated forearm technique therefore fails to meet the Sensitivity Criterion.

Another approach to monitoring awareness during anaesthesia is to record EEG (electroencephalogram) responses to auditory stimulation, and to infer changing levels of awareness from changes in the pattern of this evoked responding. One measure of evoked responding that may prove useful for monitoring awareness is known as "coherent frequency" (see

Munglani et al. 1993). The coherent frequency of the auditory evoked response represents the frequency of auditory stimulation that elicits the maximal response in the EEG. In the aforementioned study by Andrade et al. (in press), changes in the coherent frequency correlated highly with changes in cognitive performance brought about by varying the dose of anaesthetic administered. Assuming that awareness is a function of cognition, this is preliminary evidence that this measure of auditory-evoked responding can be used to predict changes in awareness.

However, it is a big step from measuring fluctuations in awareness to determining the point at which awareness ceases. EEG measures indicate a gradually declining level of awareness as anaesthesia progresses, yet there is presumably a precise point at which consciousness of a *particular stimulus* ceases. To quote a colleague, the crux of the problem is that "we don't have ear-lids" – loss of awareness of visual stimuli may occur when patients can no longer open their eyes, but the persistence of the auditory-evoked response suggests that at least a low level of auditory processing continues throughout anaesthesia. Defining the point at which patients lose awareness of auditory stimuli is therefore a far trickier problem.

Most studies of implicit memory for intraoperative events have not monitored awareness at all. Had they done so, one of two possible conclusions might prevail: either learning may occur regardless of the level of awareness, suggesting that awareness is not important for the sort of learning assessed by implicit memory tests in these studies, or learning may be observed only when awareness reaches a certain level. A recent study by Schwender et al. (1992) supports the latter position. They found that when the early cortical component of the auditory-evoked response was preserved, 5 out of 10 patients showed evidence of implicit memory for a story played to them during surgery, whereas only 2 out of 20 did so when the early cortical component was abolished.

To conclude, monitors of awareness during anaesthesia are currently being developed, and there is some evidence to suggest that subjects learn information only when their level of awareness is above a certain point. However, there is not yet a reliable way of determining whether a patient is aware of a particular stimulus. Since this is crucial to the current debate, implicit memory for intraoperative events cannot be taken as evidence for implicit learning.

Awareness inflated, evaluative conditioning underestimated

Frank Baeyens, Jan De Houwer, and Paul Eelen

Department of Psychology, University of Leuven, B-3000 Leuven, Belgium; fpaas21@blekul11.earn.bitnet

Even though a broadly accepted operational definition of awareness does not exist, Shanks & St. John (S&S) do not even attempt to formulate explicitly their own working definition. However, the claim that there is no convincing evidence for unconscious learning is mainly due to S&S's implicit conception of awareness, and more particularly, a consequence of their interpretation of the Sensitivity Criterion. Throughout the paper, S&S argue against reliance on verbal reporting as a valid awareness index. They propose other behavior indices, such as sequential letter dependencies (SLD) tasks (for artificial grammar learning) or prediction tasks (for serial reaction-time studies) as more exhaustive and sensitive assessment tools. S&S would like us to accept that *any* verbal or nonverbal behavior which might *reasonably* be thought of as reflecting a judgement based on a belief-that-*X* should be treated as an index of awareness-of-*X*. However, no formal criteria are provided to determine whether or not a particular behavior can "reasonably" be considered a judgement based on a belief-that-*X*.

This approach is troublesome, because in the limit the distinction between aware/unaware knowledge becomes meaningless. Consider the following study by Baeyens et al. (1993). While performing a distraction task, Ss were first exposed to a covariation between a subtle difference in the line-thickness (*A*) (thick/thin) of geometric figures and the valence (*B*) (positive/negative) of subsequently presented slides. Next, Ss evaluated a series of stimuli which differed in line-thickness. Ss (1) were not able to report the crucial *A-B* covariation, not even when explicitly instructed to search for it; (2) performed at chance level on a recognition questionnaire, (3) but demonstrated having acquired some knowledge about the crucial covariation, as evidenced by the evaluative ratings of the test-stimuli. Given (1), (2), and (3), we inferred learning without awareness. S&S might argue that awareness tests (1) and (2) are insensitive in relying on verbal reporting, and that an assessment of awareness type prediction test could reveal awareness which would otherwise be undetected. The results obtained in Experiment 3 (Baeyens et al. 1993) are of crucial importance in this respect. After the acquisition phase, Ss were led to believe that either a positive or a negative picture would be presented subliminally following the thick/thin line slides. Actually, the positive and negative pictures were never shown. Ss were required to indicate on each trial whether they believed they had seen the positive or the negative slide. This test was conceptualized as a (valence-free)-*implicit* knowledge test, and Ss actually performed better than at chance. However, according to the S&S criteria, these same data could equally well be interpreted as demonstrating *awareness* of the covariation, since judgements (or predictions) "must be based on conscious expectancies"! Apparently, S&S tend to equate a theoretical construct – such as "expectancy" – which might be invoked to explain performance, with the phenomenological experience of "having a conscious expectancy."

We should add that S&S are aware of this type of problem, but their arguments against the idea of prediction/SLD tasks being influenced by unaware knowledge are not at all convincing. For example, in the context of serial reaction tasks, they depend on a very narrow concept of the content-of-learning, namely, specific motor responses, rather than, for example, the spatiotemporal organisation of the stimuli. We do not see any compelling reason to accept this interpretation.

That the evidence for implicit learning apparently disappears in these examples is due to a conception of awareness which departs strongly from the standard conception, grounded in phenomenology. As long as one believes it is useful to distinguish between learning with or without awareness, one cannot but accept that the ultimate criterion for (un)awareness-of-*X* is the presence (or absence) at t_2 of a *phenomenal state* of having-the-subjective-impression-to-consciously-know-that-*X* or have-noticed-*X* at t_1 . This state should be evidenced by (1) a correct verbal statement at t_1 of *X* (recall) or by an endorsement at t_2 of a statement expressing *X* (recognition), whereby (2) subjects indicate they are not merely guessing, and (3) the statement *X* is about the state-of-affairs experienced at t_1 . In addition, (4) t_2 should be as close in time as possible to t_1 , and (5) the performance and the awareness test should be as similar as possible in terms of retrieval context. It should be noted that this set of formal criteria is in line with Merikle and Cheesman's (1986) emphasis on phenomenal experience as the ultimate ground to distinguish between perception with or without awareness.

When this set of criteria is applied, for example, to the results obtained in evaluative conditioning, the evidence of unaware evaluative learning is substantial. S&S's claim that "unconscious evaluative learning has not yet been adequately established," however, arises not only from their inflated awareness concept, but also from a somewhat flawed presentation of the actual data. First, S&S present a strange interpretation of the results of the Baeyens et al. (1990a) study. Namely, S&S agree that concurrent awareness measurement (CAM) is technically impossible with-

out *creating/manipulating awareness*. Next, however, the very fact that CAM resulted in the majority of Ss becoming aware of the CS-US contingencies is used to demonstrate that the recognition questionnaire did not meet the Sensitivity Criterion! Also, the nonsignificant result for the concurrently unaware pairs may not be surprising, given the very limited number of data available for this comparison. The important observation in this study was that evaluative conditioning was *not* influenced by the absence or presence of CAM, while this same manipulation resulted in low versus the creation of high awareness. This argues for the orthogonality of evaluative conditioning and explicit learning of contingencies.

Second, S&S consider the results of the Baeyens et al. (1992) study as an instance of association rather than dissociation between awareness and learning. However, they only tell part of the story. Increasing the number of conditioning trials indeed, resulted – at least in the 2/5/10 trial range – in an increase in awareness and in conditioning, but the absence of any correlation between level of awareness in a S and level of conditioning (over the levels of the number of acquisition trials) also suggested that the increase in awareness and conditioning were not necessarily located in the *same* Ss. Also, a dissociation was observed between conditioned valence shifts and awareness when proceeding from 10 to 20 acquisition trials: whereas awareness continued to increase from 10 to 20 trials, conditioning actually decreased. Similar dissociations were observed in Baeyens et al. (1989), in which it was demonstrated that perceptual CS-US similarity enhanced awareness but not evaluative shifts; and in an evaluative flavor conditioning study (Baeyens et al. 1990b) in which we observed that when a flavor-CS was correlated with a highly aversive US-flavor, none of the Ss demonstrated contingency awareness (recognition test), but they did evidence clear conditioning. When a color-CS was used, however, a substantial number of Ss were contingency-aware, but no conditioning was observed. All these findings are instances in which evaluative learning and awareness clearly did not “proceed in tandem,” but related in an orthogonal manner.

Third, S&S do not discuss evaluative conditioning studies in which the US, but not the CS, was presented subliminally (De Houwer et al. (1994); Krosnick et al. 1992; Niedenthal 1990). The power of these studies is that it appears to be extremely unlikely that to not-subjectively-experienced USs might ever support the conscious detection of a systematic *relation* between US and CS. Nevertheless, reliable evaluative conditioning has been demonstrated under these degraded input conditions.

We should finally stress that it is not the demonstration of unaware learning per se but further evidence for functional differences between instances of conscious or unconscious learning that will be necessary to substantiate the claim that research concerning the conscious/unconscious distinction involves more than a naive fascination for a causally noneffective dichotomy (Merikle & Reingold 1992).

ACKNOWLEDGMENT

Frank Baeyens is a Postdoctoral Researcher and Jan De Houwer a Research Assistant at the National Fund for Scientific Research (Belgium).

A step too far?

Dianne C. Berry

Department of Psychology, University of Reading, Reading RG6 2AL, United Kingdom; sxsberry@reading.ac.uk.

To start my commentary on a positive note, Shanks & St. John (S&S) suggest that we need to be cautious when accepting evidence in favour of implicit (or unconscious) learning. They posit two criteria which they argue need to be applied when

considering studies in this area. This is a useful step forward as it is clear that a number of studies that have commonly been cited as evidence in favour of implicit learning would fail the authors' Information Criterion and Sensitivity Criterion. However, I believe that S&S go a step too far in their rejection of the evidence. I have four main objections to their treatment of the evidence.

1. Conscious processing as a default assumption. S&S adopt the position that unless researchers can prove conclusively that learning is unconscious then we should assume that it is conscious. However, they never justify why conscious processing should be the default assumption. Not all researchers would accept this position. Reber (1990), for example, has argued strongly for the primacy of the implicit. This default assumption gives rise to the following three objections.

2. “If in doubt throw it out.” In reviewing the evidence, S&S seem to adopt the (implicit or explicit) rule of “if in doubt throw it out.” In line with this, they seem to apply a different criterion in discussing studies claiming to provide evidence in favour of implicit learning and studies purporting to provide evidence against it. For example, in considering Perruchet and Pacteau's (1990) artificial grammar learning study (sect. 2.5.1), S&S do not mention Perruchet and Pacteau's failure to use an appropriate control group, nor the other criticisms of this study raised by Reber (1990) and Mathews (1990). Similarly, in discussing Perruchet and Amorim's (1992) sequence learning study (sect. 2.7.3), S&S point out that the researchers did not divide subjects into aware and unaware groups, but they still go on to accept Perruchet and Amorim's conclusions. However, in discussing the sequence learning study by Stadler (1989), which could be interpreted as providing evidence in favour of implicit learning, the authors mention a possible problem with Stadler's procedure and conclude that “readers should not interpret these results . . . as convincing evidence for learning without awareness” (sect. 2.7.2).

3. Selective sampling of experimental evidence. Clearly S&S could not hope to review all relevant implicit learning studies. However, in places I felt that they had ignored some studies that were particularly relevant to their arguments. For example, after discussing Berry and Broadbent's (1984) study with the sugar production and person interaction tasks (sect. 2.6.2), they state that a detailed examination by Sanderson (1989) found evidence of associations rather than dissociations between performance and reports. However, Sanderson's work was based on a very different task (Broadbent's city transport task). S&S do not describe the findings of the more relevant Stanley et al. (1989) study, which used the same tasks as Berry and Broadbent. Stanley et al. also found evidence of verbalisable knowledge increasing with extensive practice, but they still found a significant dissociation between task performance and verbalisable knowledge. Failure to report the Stanley et al. findings is even stranger because S&S go on to state that “one alternative strategy [to questionnaires] would be to ask each subject to instruct a yoked “partner” in how to perform the task. . . . Such procedures have been used with other learning procedures (e.g., Mathews et al., 1989, for grammar learning) and have proven highly sensitive.” But this is just the procedure used by Stanley et al.

On a related note, S&S refer to “questions about possible structural rules underlying the task” in discussing Berry and Broadbent's (1984) type questionnaires. Although I agree that these questionnaires might not pass S&S's Sensitivity Criterion, the questions do not ask about possible structural rules. Given a certain history of responses, they ask subjects either to predict the next value of the output variable or to state which level of the input variable would be needed to bring the output level to target (i.e., in line with what subjects have to do when controlling the tasks). Furthermore, even if we accept S&S's argument that these questions would not be expected to elicit the relevant S-R knowledge, why were subjects not able to report the

knowledge (until they had received 570 trials of practice) using Stanley et al.'s (1989) teachback procedure?

4. Knowledge is assumed to be explicit even when this has not been tested. Throughout the target article the authors assume that certain types of knowledge are likely to be explicit, even when this has not been tested. Hence, in discussing the Altmann et al. (in press) finding of transfer from letter strings to musical tones (sect. 2.5.3), S&S suggest that "it is quite possible that subjects, if asked, would be able to report such simple regularities." Similarly, in discussing Willingham et al.'s (1989) sequence-learning experiment (sect. 2.7.1), S&S state that "this is information they may, if asked, have been able to report." Although these statements are qualified, S&S subsequently distinguish between unconscious learning of fragments and instances and unconscious rule learning, and they argue that "this review has shown that evidence for unconscious learning of any sort is highly questionable" (sect. 3). The problem is that it is not really valid to draw a conclusion about unconscious learning of instances and fragments because in most studies the awareness issue has simply not been addressed.

5. A final note. Taking these four objections together, it could be argued that S&S are too premature in concluding that "human learning is almost invariably accompanied by conscious awareness." It could also be argued, however, that they have focused too closely on the issue of conscious awareness. Many proponents of implicit learning do not equate the term with unconscious learning (e.g., Berry, 1994; Berry & Dienes 1993; Mathews et al. 1989). Rather, they argue that implicitly and explicitly acquired knowledge might be better distinguished on other grounds, for example, by looking at whether they have different properties of storage and retrieval. It is a pity that Shanks & St. John did not devote more of their article to considering attempts to establish other qualitative differences between implicit and explicit learning.

Are subliminal mere exposure effects a form of implicit learning?

Robert F. Bornstein

Department of Psychology, Gettysburg College, Gettysburg, PA 17325;
rbornste@gettysburg.edu

Although Shanks & St. John's (S&S's) review of the implicit learning literature includes extensive discussions of research on Pavlovian conditioning, artificial grammar learning, instrumental learning, and learning and awareness in serial reaction time tasks, it devotes virtually no attention to studies of stimuli perceived without awareness. In fact, early in their analysis of the implicit learning literature (sect. 2.2), S&S assert that "although there have been some positive results [in studies of perception without awareness], a corresponding number of negative findings leads us to suggest that unconscious learning with subliminal stimuli has not yet been conclusively demonstrated."

S&S's assertion is certainly correct with respect to certain areas of subliminal perception research (see Bornstein 1989b). However, their statement is clearly incorrect with respect to at least one area of research in this domain. Specifically, studies of Kunst-Wilson and Zajonc's (1980) subliminal mere exposure (SME) effect (i.e., subjects' increase in liking for stimuli that have been repeatedly exposed for very brief durations, but that cannot be recognized or discriminated from other briefly presented stimuli at better-than-chance levels) have produced robust, replicable results (see Bornstein 1992). SME experiments provide compelling evidence for the existence of perception without awareness, both within the context of traditional dissociation models of unconscious perception (Erdelyi 1985)

and within the context of Reingold and Merikle's (1988) more recent direct/indirect framework.

The purpose of this commentary is to discuss research on SME effects, focusing on S&S's assertion that SME effects do not represent an example of implicit learning. In the following I will (1) discuss studies demonstrating that SME effects are robust and reliable; (2) review studies demonstrating that SME effects influence subjects' affective, cognitive, and behavioral responses; (3) describe recent findings indicating that SME effects can last for up to one week; and (4) note the methodological limitations in SME studies that do not permit us to conclude unequivocally that SME effects are an example of implicit learning.

SME effects are robust and reliable. To date, there have been more than a dozen published experiments demonstrating that stimuli perceived without awareness produce robust mere exposure effects (see Bornstein 1992, Table 1, for a summary of these results). Although the majority of these investigations used simple geometric stimuli (e.g., Mandler et al. 1987), at least one study (Bornstein et al. 1987) has obtained SME effects with more naturalistic social stimuli (i.e., photographs of college students). Not only do stimuli perceived without awareness produce robust exposure effects, but recent meta-analytic reviews and laboratory studies have demonstrated that the magnitude of the exposure effect produced by stimuli perceived without awareness actually exceeds the magnitude of the exposure effect produced by stimuli that are recognized by subjects at better-than-chance levels (see Bornstein 1992 for a detailed discussion of these findings).

SME effects influence affective, cognitive, and behavioral responses. Although Zajonc's (1968) initial conceptualization of the mere exposure effect focused exclusively on affective responses to repeated, unreinforced stimulus exposures, evidence collected during the past two decades indicates that mere exposure effects may also influence subjects' behavioral and cognitive responses. With respect to behavior, at least one study (Bornstein et al. 1987, Experiment 3) has demonstrated that repeated subliminal exposure to a photograph of a college-age student influenced subjects' behavioral responses to that student when the student was later encountered in a laboratory negotiation task. With respect to cognition, several studies (e.g., Mandler et al. 1987) have demonstrated that repeated subliminal exposure to a stimulus influences subjects' cognitive evaluations of that stimulus. Clearly, SME effects influence subjects' reactions to a stimulus on a variety of rating dimensions.

SME effects are relatively long lasting. Most SME experiments assess subjects' reactions to a stimulus after only a brief period of delay between stimulus exposures and ratings (Bornstein 1989a). One experiment, however, (Seamon et al. 1983b) assessed systematically the effect of delay between exposures and ratings on the magnitude of the exposure effect. Seamon et al. examined affect ratings of briefly presented polygon stimuli collected immediately after stimulus exposures, one day later, or one week later. Not only did Seamon et al. find that SME effects persisted for one week following stimulus exposures, but these researchers actually found a slight increase in subjects' affect ratings of merely exposed stimuli between the one-day and one-week delay periods. This result dovetails with recent meta-analytic findings demonstrating that there is a positive relationship between the amount of delay between stimulus exposures and ratings and the magnitude of the exposure effect (see Bornstein 1989a, Table 6).

Are SME effects a form of implicit learning? Although SME effects are robust and reliable, affect a variety of outcome measures, and last for up to one week following stimulus exposures, it does not necessarily follow that they are a form of implicit learning. Such a conclusion would require researchers to demonstrate that subjects' responses to merely exposed

subliminal stimuli generalize to other, similar stimuli that are encountered at some later date (see Reber 1969; 1989a). Unfortunately, such a study has never been conducted. To be sure, there is ample evidence that supraliminal mere exposure effects generalize to other similar stimuli, and that these effects have many of the characteristics of implicit learning (see Gordon & Holyoak 1983). However, until a mere exposure effect experiment such as that conducted by Gordon and Holyoak with supraliminal stimuli is conducted using subliminal stimuli, it will be impossible to draw firm conclusions regarding the generalization of SME effects to other related stimuli. Until such a study is conducted, researchers will not be able to determine whether SME effects fulfill the criteria required to demonstrate the existence of implicit learning effects.

Conclusion. S&S argued that studies of perception without awareness do not provide evidence for unconscious learning because these experiments have been difficult to replicate and have never provided evidence that stimuli perceived without awareness produce measurable effects on behavior (sect. 2.2). A review of the literature on SME effects suggests that these conclusions are unwarranted. Although SME effects have not yet been used to demonstrate implicit learning effects, the SME paradigm has the potential to provide important information regarding this issue. Thus, it would be premature and counterproductive to ignore the SME paradigm as a tool for studying implicit learning. SME experiments might well prove useful in increasing our knowledge regarding the parameters of implicit learning phenomena while simultaneously providing important information regarding the range of affective, cognitive, and behavioral effects produced by stimuli perceived without awareness.

ACKNOWLEDGMENT

Preparation of this comment was supported in part by National Institute of Mental Health Grant MH46824-01A1.

Of what are we aware?

Nathan Brody and Michael J. Crowley

Department of Psychology, Wesleyan University, Middletown, CT 06459

Shanks & St. John's (S&S's) target article persuasively demonstrates the power and importance of the Sensitivity Criterion as a test for determining the nature of a subject's conscious experiences. We recently completed a study that illustrates the importance of the Sensitivity Criterion.

In our study, Ss were presented with a standard artificial grammar learning experiment. In addition to the usual procedures, Ss were asked to respond to a grammatical knowledge questionnaire that included a variety of questions about their knowledge of the rules of the grammar, including questions about abstract features of the grammar such as the rules governing recursions as well as knowledge about permissible letter sequences. Ss were assigned to conditions in which their knowledge of the grammar was assessed before or after performing the standard grammatical classification task.

We obtained the following results: (1) Ss demonstrated explicit grammatical knowledge on the questionnaire. Ss had a mean score on the grammatical questionnaire that was 2.48 standard deviations higher than the mean of a control group that had not been exposed to the grammatical strings. The mean score of the experimental group on the grammatical classification task was 2.19 standard deviations higher than the mean of the control group. (2) The correlation between performance on the grammatical knowledge questionnaire and the grammatical classification task was .54. The test retest reliability obtained by correlating scores for the two different random orders of the

same items in the grammatical classification task was .58. (3) Whether Ss responded to the questionnaire before or after performing the classification task did not influence their classification performance.

These results demonstrate that Ss acquire explicit knowledge from exposure to grammatical strings and individual differences in the amount of knowledge acquired are related to performance on the grammatical classification task. More generally, these data illustrate the power of the Sensitivity Criterion and indicate that what Ss are aware of may be a function of the way in which awareness is assessed.

Methods of assessing awareness may induce awareness that was not previously present. The awareness of the explicit rules of the grammar that is demonstrated in the grammatical questionnaire in our study may or may not have been present at the initial exposure to the grammatical strings and may have been induced by the act of filling out the questionnaire itself.

Methods of assessing awareness may obscure a subject's knowledge. Consider the typical findings of studies demonstrating preferences for previously encountered geometric stimuli that Ss do not know they have encountered (Seamon et al. 1983a; 1983b). In these studies, Ss are tachistoscopically exposed to geometrically similar stimuli. They are then asked to discriminate between previously presented stimuli and new stimuli. If the conditions of viewing are properly arranged, Ss will demonstrate chance levels of performance on the discrimination task. If they are asked to choose which of a pair of old and new stimuli they like, they will prefer the previously encountered stimuli. Discrimination by preference occurs as long as one week after the initial presentation of the stimuli, when Ss' ability to discriminate between stimuli previously encountered and new stimuli is exactly at chance. It is possible that the question "Which of these stimuli have you encountered?" elicits a search of memory traces. The tachistoscopic conditions of viewing of the stimuli should create poorly articulated memory traces that are subject to intralist interference if Ss engage in a search process attempting to discriminate old from new stimuli. The question "Which stimulus do you like?" may not induce a search of memory traces. Ss may respond immediately and therefore the access to the encoded memory traces may be more direct and less subject to intralist interference. If this analysis is correct, then the different questions asked of Ss may access memory in different ways. Different results for the different questions do not necessarily provide evidence for separate and independent encoding of affective and cognitive knowledge.

Awareness must be inferred from responses. Methods of assessing awareness may each be subject to method-specific sources of variance that change the relationship between awareness and the response indicators used to infer the presence of awareness. One of the reasons that signal detectability theory provided a major methodological advance in the study of subliminal perception is that the theory described the relationship between responses and awareness. Changes in responses might or might not be indicative of a change in a subject's experience. Howe et al. (1993) argued that it is necessary to develop formal models of the relationship between responses and the hypothetical processes that are related to them in order to demonstrate independence of processes. So too, formal models relating states of awareness to response indicators may be required to demonstrate unconscious influences on behavior. In the absence of such models, S&S's criteria appear to have considerable validity. If the sensitivity requirement is satisfied by appropriately exhaustive procedures, and it is possible to demonstrate the Ss are neither aware nor capable of being made aware of something, and it is possible to demonstrate that what Ss are not aware of has an influence on behavior that is different from what it would have when it is in awareness, then it will be necessary to assume that there are unconscious influences on behavior. Are there such influences?

Is implicit learning about consciousness?

Richard A. Carlson

Department of Psychology, Pennsylvania State University, University Park, PA 16802-3106; cvy@psuvm.psu.edu

I heartily agree with Shanks & St. John's (S&S's) central point – that “unconscious” implicit learning has not been adequately demonstrated. I am generally enthusiastic about their summary of the logic necessary for such a demonstration – their Information and Sensitivity Criteria. I also find convincing their argument that the learning systems or modes that *have* been experimentally dissociated are conscious instance/fragment learning and conscious rule learning. Of course, my agreement and enthusiasm are based at least in part on the compatibility of their conclusions with those of my own efforts (e.g., Carlson & Dulany 1985; Dulany et al. 1984). But I have the nagging feeling that the central point has been missed – that this review fails in a fundamental way to do justice to the literature on so-called implicit learning.

I believe the problem is that very compelling intuitions, supported by commonsense observations, lend credence to the notion of implicit learning. In teaching and discussing this literature, I find that sufficiently detailed experimentation and argumentation may silence – but rarely if ever *convince* – students and colleagues who begin with a belief in unconscious or implicit learning. Furthermore, I suspect that many thoughtful people will continue to think that this literature is somehow about what is and is not conscious, rather than being convinced that it concerns instead (merely?) the distinction between instance and rule learning.

For me, the central questions thus ought to be – What are the sources of these intuitions concerning implicit learning, and how are we to encompass them in our scientific understanding of the mind? Part of the answer is certainly suggested by the concept of *correlated hypotheses* (Dulany 1961; S&S, sect. 2.6.1., para. 3). Students of implicit learning might do well to explore the links between this concept and *action identification theory* (Vallacher & Wegner 1987), which outlines some systematic sources of variation in how individuals will identify their own actions – and by extension, their own knowledge. More fundamentally, we cannot make progress in this area so long as *consciousness* and *awareness* remain largely primitive, unanalyzed constructs. Rather than arguing about, for example, whether recognition tests do or do not assess conscious knowledge (e.g., Dulany et al. 1985 vs. Reber et al. 1985), we ought to ask: What constitutes consciousness in these learning and test situations? For example, what *information* specifies an experiencing self during learning and at test (e.g., Carlson 1992; Natsoulas 1992)? How does the “perceptually given” self (Neisser 1991) differ from learning to test? What difference might this difference make in subjects' abilities to report their experiences?

I have space only to suggest an answer to the question posed at the beginning of the previous paragraph. That answer goes something like this: when learning and test situations provide overlapping *stimulus-specifying* (or, more precisely, but more awkwardly, *domain-object-affordance-specifying*) information, but differ substantially in the *self-specifying* information they provide, evidence for implicit learning can be obtained. Stimulus-specifying information simply comprises the perceptible properties of the displays presented to subjects. The notion of self-specifying information is perhaps less familiar: just as perspective information in pictures specifies a visual standpoint for the depicted scene (Gibson 1971) and the “ambient optic array” specifies a point of observation within the environment or an “ecological self” (Neisser 1991), the information available in experimental learning and test situations specifies an experiencing (learning, judging, etc.) self.

As presented by S&S, the Information Criterion refers only to

stimulus-specific information (perhaps expressed at various levels of abstractedness). But correlated hypotheses, as freely expressed by subjects in operant conditioning experiments such as those analyzed by Dulany (1961) in an earlier literature on “learning without awareness,” include self-specific information about such things as intentions: for example, “I am trying to.” And Dulany's fully developed theoretical analysis of learning in this paradigm demonstrated that performance could be predicted by the *combination* of subjects' intentions and their hypotheses about which behaviors were being reinforced (Dulany 1968).

My suggestion, then, is that two aspects of implicit learning situations must be considered. First, as S&S note, accuracy of test performance is (given appropriate subject goals) determined largely by the overlap of stimulus-specific information available to subjects' awareness between learning and test. For example, for a particular test situation we can determine the performance to be expected on the basis of information that might be paraphrased as “the presence of these features makes this letter string grammatical” – regardless of whether “these features” is an instance-based or abstract specification (e.g., Dulany et al. 1984; 1985). Second, the intuition that learning is “implicit” or “unconscious” is supported by cases in which there is a lack of overlap or correspondence between *self-specific* information available during learning and at test. Here, Vallacher and Wegner's (1987) observations about how individuals identify their own actions may provide some clues to understanding. To understand how consciousness is involved, we must ask not just what information guides performance, but *how* it does so in combination with subjects' intentions and attention.

Appropriately interpreted, then, evidence from situations in which learning seems to be implicit might tell us about the nature of consciousness and its role in learning rather than marking points to be labeled “conscious” and “implicit.” I hope that the careful analyses presented by Shanks & St. John will help shift discussion in that direction.

The aware pigeon

A. Charles Catania

Department of Psychology, University of Maryland, Baltimore, MD 21228-5398; catania@umbc7.umbc.edu

I have no trouble with the proposition that human verbal behavior can determine the subsequent verbal or nonverbal behavior of the same individual. The phenomenon has been well documented and is especially convincing when the shaping of verbal behavior is followed by a change in nonverbal behavior that is described by the newly shaped verbal behavior (e.g., Catania et al. 1982). With respect to such procedures, it may be useful to point out that recent research by Svartdal (1992) showed that response force became more sensitive to operant contingencies with increases in cognitive load in a complex discriminative task. That research obtained verbal reports by asking subjects what they would tell someone else about how to succeed in the task; subjects reported only strategies for attending to various properties of the task stimuli rather than response properties (e.g., hand position) that might have been correlated with differential force.

But Svartdal's experiment, of course, like so many others discussed by Shanks & St. John (S&S), involves only a dissociation between a verbal report and a response measure and therefore fails to meet S&S's Information and Sensitivity Criteria. Their concern, as expressed even in their title, is with human learning systems, but the problems posed by human verbal behavior for the assessment of awareness suggest that it might be better to turn to nonhuman organisms. With such organisms, we are less likely to make inappropriate assumptions

about the behavior the subject comes into the experiment with. Furthermore, as for awareness or other private events (cf. Skinner 1945; 1984a; 1984c), when we try to teach others when it is appropriate to speak of awareness, we have no more access to what is going on in their heads than to what is going on in the head of the pigeon. (I regard both kinds of heads as full, not empty, but I do not expect to find out much about what is in them just by asking.)

So here goes with some pigeon thought experiments. Suppose I present symbol sequences to a pigeon and teach it to peck a left key after it sees any one of a set of symbol sequences generated by a finite-state grammar and a right key after it sees any one of a set of symbol sequences not generated by that grammar (this would probably not be hard to do; pigeons can learn even large arbitrary stimulus sets at fairly high levels of accuracy if procedures are set up properly: e.g., Vaughan 1988). Once I have taught the pigeon to discriminate the symbol sequences in this way, I can now set out to discover the particular features of the sequences upon which this discrimination depends by presenting probes of various new legal and illegal strings. Except that the reinforcement contingencies will be appropriate for the behavior of a pigeon rather than a human, the methods closely resemble those described by S&S in their discussion of the research that followed upon Reber's (e.g. 1967; 1969; 1976) experiments. One possible outcome is that the pigeon might demonstrate that it had learned some features of the finite-state grammar; another is that it might demonstrate that it had learned some other set of features, such as particular substrings.

Consider yet another experimental arrangement. On one of four pigeon keys I present some stimulus, such as a red spot, to a pigeon. Over successive trials the position of the spot varies in some systematic way on the array of four keys and I teach the pigeon to peck the key on which the red spot appears. Once the pigeon's responding is well established, it is likely that its reaction times will be shorter than usual when the stimulus appears in positions that are predictable in some way from the algorithm that determines the successive positions of the spot, much as in the serial reaction time tasks described by S&S. (This is probably even easier to do than the first proposed experiment; much more complex responding with respect to locations and response timing has been studied with pigeons: e.g., Shimp et al. 1989.)

Next, I complicate the procedure a little. On each trial, all the keys are first lit white; I then allow the pigeon to produce the red spot only by pecking the key to which the red spot has been assigned on that trial by the algorithm. I would not find it surprising if pecks on an appropriate key were more likely when the spot was scheduled to appear in a position that was predictable in some way from the algorithm than otherwise. On the other hand, I could also imagine that to make the experiment work I would have to refine the procedure (e.g., by embedding occasional prediction trials within the sequence of reaction time trials rather than presenting prediction and reaction time trials as separate sequences at different times, or maybe even by finding some way to obtain both a prediction and a reaction time on the same trial). After all, a failure to obtain a correlation between reaction times and accuracy of predictions over trials within the sequence might only mean that the procedures had not yet adequately met S&S's Sensitivity Criterion, so these experimental moves are consistent with their strategy (sect. 2.3.2) of making performance and awareness tests as alike as possible, except for differences in instructions.

For the present purposes, it hardly matters whether in the first experiment the pigeon showed us that it had learned the finite-state grammar or some set of features, such as particular substrings, or that in the second experiment its reaction times and correlated predictions depended on sequence learning or response bias; it also makes no difference whether what is learned is characterized as particular associations, relationships,

or concepts. The question is whether it is appropriate to speak of the pigeon's awareness, because according to S&S's criteria, the evidence from the pigeon should be as acceptable as that from a human. They have ruled out verbal reports as one of their criteria for judging a learner's awareness, and therefore their Information and Sensitivity Criteria should be as applicable to nonhuman organisms as to human ones.

If that is so, however, then they must conclude that the pigeons are as aware of whatever features our probes identify as the humans would have been under parallel procedures. In fact, by their criteria (sect. 2.7.3) differential responding to events in recognition tests is equivalent to awareness of them. If learning about events is synonymous with being aware of them, however, of what use is the concept of awareness? In other words, S&S have designed a set of conditions such that learning cannot occur without awareness, because learning has become their criterion for awareness. Thus, their account seems to have some and perhaps all of the properties of a *reductio ad absurdum*. After presenting their criteria they say, "It is . . . important to recognize that our criteria do not make unconscious learning undemonstrable" (sect. 2.3.2), but by the end they have demolished every possible proof.

This is perhaps to be expected. Terms such as *consciousness* and *awareness* have been around for a long time, but they have yet to acquire consistent technical senses in psychology (cf., Verplanck 1992). Whatever else S&S do in response to this commentary, I hope they will provide an explicit definition of these terms. What sorts of things are they? Are they precursors or consequences of behavior? According to S&S, awareness has content, but does that mean it is not a process? Given that the parallel distributed processing models reviewed by S&S simulate human performances that meet their criteria for awareness, must we assume that the models or the computers that implement them are aware? Is there a difference between consciousness and awareness? What about awareness, if what the subject is presumed to be aware of is actually not the case? When, in their very last paragraph, S&S speak of "conscious awareness," do they really mean to imply that there might be such a thing as unconscious awareness? And what about the evolutionary origins of awareness? There must have been existing behavior before there could be any awareness of that behavior, but what survival advantages can awareness have without something to be aware of? Did the earliest discriminations require that the organism be aware that it was discriminating? Probably not.

Like it or not, S&S have brought us back to verbal reports: "Of course, if the prediction task and recognition tests cannot be treated as tests of awareness, then we have no recourse other than to examine the subject's verbal reports as the only available index of explicit knowledge" (sect. 2.7.5). If we are interested in them, we must tackle the sources of our own vocabularies to understand how they function in our scientific verbal behavior. There are some properties of the phenomena to which the term *consciousness* is applied that are inextricably tied to human verbal reports, but we must not forget that those reports are behavior too. My own view is that awareness of our own awareness (as in distinguishing between being aware of a stimulus and being aware that you are aware of it) is one of the special characteristics of human behavior. One starting place to learn about it is in the study of verbal behavior and of the practices by which we learn the consistent use of these vocabularies from each other, even though our verbal communities can act only upon public behavior and the circumstances under which it occurs. For that purpose, we must look not to the syntactic structure of verbal behavior but to its functional properties (Skinner 1957).

Awareness and abstraction are graded dimensions

Axel Cleeremans

Laboratoire de Psychologie Industrielle et Commerciale, Université Libre de Bruxelles CP 122, 1050 Brussels, Belgium; axcleer@is1.ulb.ac.be

When Reber (e.g., 1967) introduced the notion that learning could proceed without awareness, psychology was very different from what it is today. In particular, the dominant metaphor for information processing was one based on the use of symbol manipulation carried out in rule-following systems – the kind of goal-directed, fully conscious activity subjects engage in when solving complex formal problems (Newell & Simon 1972). Hence Reber's fascination with his subjects' apparent ability to acquire and use information about material generated from finite-state grammars without being able to tell him what the rules were. This phenomenon that Reber dubbed implicit learning has now been replicated dozens of times: we cannot report verbally on every aspect of what motivates our decisions after a learning episode.

Shanks & St. John (S&S) indicate themselves that they have no argument with the existence of such dissociations. However, they dispute the claim that such dissociations necessarily entail the existence of a separate, independent unconscious learning system. Instead, they propose (along with others, see Perruchet & Amorim 1992) the following account of implicit learning. Subjects in typical implicit learning situations acquire a single database of fully conscious, fragmentary information about the stimulus material. This knowledge is demonstrably sufficient to account for subjects' performance. Independent mechanisms have access to that knowledge to sustain task performance on the one hand, and explicit reports on the other. Usually, associations are observed between performance and explicit reports. In cases where dissociations are observed, these stem from failures of the tests of explicit knowledge to be either sensitive enough (hence failing the Sensitivity Criterion) or to ask subjects about what they know, that is, instances rather than rules (hence failing the Information Criterion).

In this commentary I want to suggest that such a "systems" view may not be the best way to understand the data nor to gain new insights into the mechanisms involved. I have three basic problems with S&S's account.

First, I doubt that anyone researching implicit learning would be willing to defend the position that S&S take pains to demolish, namely, that there is an unconscious mind that is just the same as the more familiar conscious one, only minus consciousness. I doubt that even Reber ever believed in such a strong statement, and there is indeed little empirical evidence to support it. Such a "systems" analysis, in which crude dichotomies are used to account for the data, needs to be revised in light of the fact that both availability to consciousness and abstraction appear to be graded and separate dimensions with which to characterize knowledge and learning. New ways of thinking about these issues have emerged, particularly through recent developments in computational modeling. A further point here is that I believe it is misleading to reinterpret the data in terms of variation along one dimension (instances vs. rules) instead of along the other (conscious vs. nonconscious) because these dimensions, as S&S themselves recognize at one point in their target article (sect. 1.1, para. 5), are independent.

Second, this account fails to acknowledge that neither task performance nor tests of explicit knowledge are process-pure (Merkle & Reingold 1991). Explicit knowledge can demonstrably influence performance (e.g., Curran & Keele 1993), and knowledge acquired during performance can similarly influence tests of explicit knowledge. In addition, most tasks are presumably carried out using a mixture of strategies that range from mere memorization to active rule induction. Because of this, associations between performance and explicit reports are in-

deed expected in most cases. Note, however, that some subjects, such as amnesic patients, show no evidence of explicit sensitivity to instances despite remaining able to perform complex discriminations. I simply fail to understand S&S's rather crude dismissal of these data as evidence that learning can be implicit.

Third, and perhaps most crucial, this account is not a process theory. It may be descriptively adequate in the simplest of cases, but it fails to have explanatory power and plausibility in more complex cases. As an illustration, consider the difference between problem solving and learning about sequential structure in a choice reaction time (RT) experiment. Solving a complex problem without the aid of external memory is very hard and fully explicit, because all the relevant pieces of information need to be held in working memory, the contents of which subjects are aware of by definition. By contrast, learning about the sequential structure of the stimuli in a choice RT task is implicit and effortless – indeed, almost automatic – because information about the sequence can modulate performance without being represented in working memory, that is, without being in the "focus of awareness." Relevant relationships between stimuli need not and are not represented in consciousness when learning is taking place. This may be hard to believe as long as one sticks with a warehouse metaphor of cognition, where large databases of facts about the world are accessed by the various mechanisms that need them. Indeed, in such a case the fragmentary information that the database contains has been represented in working memory during learning and can be retrieved during a recall or recognition test. If one makes the further assumption that this knowledge is exclusively responsible for performance, then it indeed becomes hard to understand what is implicit about implicit learning.

I find it rather implausible, however, that subjects' knowledge can be adequately accounted for by such theories. Consider, for example, a subject performing in Cleeremans & McClelland's (1991) experiments. In sharp contrast to the simple, short, repeating, deterministic sequences used in the vast majority of sequence-learning experiments, the stimulus material we used was generated from a nondeterministic finite-state grammar. Hence, almost all permutations between elements of subsequences of any length appear during training, albeit with different frequencies. There is an infinite number of such sequences, and still thousands of them if one only considers subsequences of up to six elements. I find it utterly implausible to assume that subjects who were merely instructed to respond to the current stimulus would somehow consciously encode and memorize all these possible subsequences and use this knowledge to prepare explicitly for the next event. Yet, the reaction time data shows exquisitely detailed sensitivity to the ensemble of constraints resulting from an encoding of all the subsequences (see Cleeremans 1993b). There is no evidence whatsoever that subjects have conscious access to this kind of distributional information about the stimulus material. The fact that subjects can consciously retrieve specific instances does not tell us anything about whether these instances are what performance is based on, nor does it tell us anything about how they are used, if at all, during learning. This is not to say that one could not, in principle, account for these data through pure instance-based theories, but rather than doing so it seems to produce either implausible or purely descriptive interpretations of the data.

As S&S again acknowledge, the crucial distributional information gets represented spontaneously as a "side effect" of processing in connectionist networks. In the rest of this commentary, I would like to show how such models provide better natural primitives with which to understand both processing and representation in implicit learning experiments.

Connectionist networks learn by processing instances. Yet, the internal representations they develop as a result of processing may or may not be more general than what could be derived from a list of the processed instances, depending on a variety of

factors, such as the network's representational resources or the demands of the task. For example, a simple recurrent network (SRN; see Cleeremans 1993b; Elman 1990) trained to process sequences generated from a finite-state grammar will sometimes develop internal representations that are organized in clusters, with each cluster representing a node of the grammar – as abstract a representation as could be. In other cases, however, the network's internal representations tend to be organized in numerous very small clusters that each correspond to one or a few training instances (see Servan-Schreiber et al., 1991, for detailed examples). As a case in point, the SRN has often been described as processing fragmentary information. This is descriptively correct, but it is not how things work inside the network. The network does not develop a database of subsequences that it can consult and ponder about as a result of training. Instead, as it processes each stimulus, the constraints that exist between the successive elements of the sequence are progressively incorporated in the pattern of connection weights so as to allow the network to respond better to the task demands. Note that the fact that subsequences are not explicitly represented in the network does not make it incapable of recognizing such sequences either. The network can indeed be used as a finite-state recognizer (Servan-Schreiber et al. 1991), and could discriminate between old and new sequences just as human subjects can.

In short, there is a representational continuum that extends from raw storage of instances to fully abstract representations. Abstraction is a graded dimension. Hence it may not be surprising to find that subjects end up developing representations that are abstract in some cases and not in others.

Turning now to consciousness, S&S are right in pointing out that current models of implicit learning phenomena are essentially silent on the issue of awareness. I therefore disagree with their interpretation of weight changes in connectionist networks as resulting both in better performance and better awareness of the structure of the training set. On the contrary, the very fact that obviously nonconscious and elementary networks are able to account for human performance in a number of complex implicit learning tasks has always struck me as evidence that awareness is in fact ancillary in these tasks, that is, that it is not required for successful performance, and that additional mechanisms are needed to represent its effects. In addition, S&S seem to misapply what they describe as a "standard idea in the psychological literature" (sect. 3.3., para. 11), that is, that the states of a computational system can be interpreted as the contents of consciousness but not the processes that produce these states. In a connectionist network, activations represent states whereas processing is carried out through connections between units. Hence, if we stick to the standard interpretation, there is no sense in which a network that grows stronger connections in response to its training environment becomes more aware of the constraints represented by the weights on these connections.

Thus, a cautious approach here requires one to admit that the models have nothing to say about awareness per se.¹ However, one can still usefully explore the *effects* of awareness on performance, for example, by assuming that being aware of some additional knowledge about the stimulus material makes that knowledge available for further processing and allows it to modulate performance (see Cleeremans 1993a).

As far as empirical data are concerned, I very much agree with Whittlesea and Dorken's (1993) recent analysis. Briefly, they suggest that knowledge can be implicit for a variety of reasons, such as because one's attention is not drawn to the relevant relationships, because people have "poor theories about what they learn and its implications for future performance" (p. 245), or because "knowledge may have an unconscious influence on subsequent activities" (p. 245) and may be hard to describe explicitly when it is distributional in nature and represented in nonverbal codes.

Hence, it seems that awareness is also a graded dimension. In the case of implicit learning, the key point is that subjects do not need to be aware of all the relevant constraints when making a decision. These constraints are implicit in the sense that they are not all simultaneously represented in working memory when the decision is made.

To conclude, there is indeed no need to appeal to a wholly separate and super-powerful unconscious learning system to understand implicit learning data. Implicit learning involves elementary and familiar mechanisms. Yet, by the same token, there is also no need to assume that knowledge can only modulate performance if it is explicitly represented. Distributional knowledge can be acquired incidentally and used implicitly in different contexts.

ACKNOWLEDGMENT

The author is a Research Associate of the National Fund for Scientific Research (Belgium).

NOTES

1. Note also that even if one assumed that internal representations or connection weights were available for inspection by some other part of the system, it may not necessarily be easy to extract synthetic information from them. It takes sophisticated analysis methods to uncover the structure present in distributed, time-varying representations.

Dissociable definitions of consciousness

Zoltán Dienes and Josef Perner

Laboratory of Experimental Psychology, Sussex University, Brighton, Sussex BN1 9QG, England; dienes@epvax.sussex.ac.uk.; josef@epvax.sussex.ac.uk.

Shanks & St. John (S&S) have provided us with a clear and cogent statement of why we should not believe in implicit learning. They find that all the evidence reviewed is consistent with the claim that whenever subjects learn that *A* predicts *B*, they have conscious knowledge of the *A-B* relationship. S&S use what at first sight seems like a perfectly reasonable definition of being conscious of the *A-B* relationship, namely, subjects have conscious knowledge if they are able to state that they expect *B* in the context of *A* on a forced-choice test. This type of objective threshold indeed captures some aspect of our intuitive notion of consciousness. It misses out on an important aspect, however – an aspect that has probably been inspiring the "believers" in this debate.

Our intuitions about consciousness are not exhaustively specified by the type of objective threshold used by S&S. Cheesman and Merikle (1984) pointed out that one could also define a subjective threshold of consciousness, namely, knowledge underlying discriminative performance could be considered unconscious if subjects believe they are literally guessing on the test. In that case, subjects lack metaknowledge about their knowledge; they do not know that they know. Blindsight patients show by their discriminative responding that, for example, the knowledge about whether the stimulus is a circle or a square is above objective threshold. However, the knowledge could be regarded as unconscious because patients claim they are guessing. Chan (1992) argued that knowledge of artificial grammars could be unconscious in that same sense of subjects lacking metaknowledge about their knowledge: Chan found that subjects were just as confident in their incorrect responses as in their correct ones.

These results were recently replicated and extended by Dienes et al. (submitted), who found that when subjects believed they were literally guessing in classifying test strings of an artificial grammar, they were still classifying significantly above chance. That is, knowledge of an artificial grammar can be below a subjective threshold of awareness, as defined by Cheesman

and Merikle (1984). In addition, a secondary task during the test phase interfered with knowledge above the subjective threshold but not below, indicating that the distinction might be carving nature at its joint.

S&S's "objective threshold" criterion requires using direct and indirect tests that are very similar in order to satisfy plausibly the Sensitivity and Information Criteria. But then it becomes exceedingly difficult to imagine an information processing mechanism that would allow performance on the indirect test but not the direct test. For example, take the claim that (1) the subjects' knowledge of an artificial grammar consists of knowledge of fragments of that grammar, say, predominantly bigrams; and (2) this knowledge is conscious. According to S&S, these are two orthogonal claims. Let us say we accept claim (1) but not (2), that is, we think the knowledge is unconscious. How should subjects perform on a bigram rating task? The problem is that S&S provide no principled constraints for when unconscious knowledge should and should not apply. What sort of mechanism would not apply the knowledge when we give it a task specifically chosen to tap exactly the contents of the knowledge in exactly the situation in which all relevant retrieval cues are present (so that it is a sensitive test)? If the bigram rating test did not elicit the knowledge, we would have no grounds for believing that the knowledge was of *bigrams* and not higher order fragments. Thus, whenever the Information and Sensitivity Criteria are satisfied, it becomes unlikely that an objective threshold would detect different types of knowledge. It remains quite possible that a subjective threshold would.

In trying to establish whether knowledge is unconscious, one needs to ask how to establish whether knowledge is conscious. S&S address this question only cursorily in section 2.7.5. They argue that recognition and prediction are conscious tasks, and that unconscious processes are unlikely to play a role. They put forward two points in their defense:

1. They point out that in the sequence tasks, prediction and RT tasks require different responses: in the prediction task, subjects respond to the next stimulus, whereas in the RT task they respond to the current stimulus. However, the savings in the RT task are based on the previous stimulus anticipating the next. Thus, for example, Cleeremans's connectionist model (Cleeremans 1993b) of the sequence tasks can be used to simulate either RT or prediction performance. Even if they were responses to different aspects of the task, what does this tell us about whether the prediction task requires conscious knowledge of the sequence?

2. S&S argue that if some unconscious process is contributing to recognition performance, it should behave in the same way as RT. Specifically, RTs should be faster in the recognition test for old rather than new sequences, and perceptual fluency should affect recognition. If these arguments have any force, they show that recognition did not get at the same knowledge base as RT, and thus that there was a failure to meet the Sensitivity or Information Criteria. These arguments would only be relevant to the issue of consciousness under the tacit assumption that there are two and only two types of knowledge, and if a certain measure of knowledge does not share certain properties (e.g., perceptual fluency) with measures of knowledge of one type (i.e., the ones purported to be unconscious) then it must be measuring the other type (i.e., be conscious). But there is no reason to make this assumption, so the question of whether or not recognition is purely conscious is not addressed by these data.

We suggest that a powerful way of telling whether or not knowledge is conscious is to ask the subjects. If subjects believe they are literally guessing, we have good evidence the knowledge is unconscious. By not asking subjects for the basis of their recognition or prediction judgments, S&S might be failing to meet their own Information Criterion. Recognition and prediction judgments are always conscious in the trivial sense that subjects are aware that they are making a judgment. However, it

does not follow that subjects have conscious knowledge of the basis of their correct judgments. Making this inference violates the Information Criterion; that is, from knowing that subjects are aware of "making judgments" one cannot conclude that they are aware of the basis of making these judgments.

S&S are aware of difficulties with their criterion and conclude that if one cannot use recognition and prediction tasks as exclusive measures of conscious knowledge, one has to rely inadequately on recall. But this is not the only option: one can give up the criterion and search for a more fruitful one. The subjective threshold is one of the more promising ones (Berry & Dienes 1993; Dienes & Perner, forthcoming; Merikle 1992).

Implicit practical learning

Elizabeth Ennen

Department of Philosophy, McGill University, Montreal, Quebec, Canada
H3A 2T7 czeh@musica.mcgill.ca

Shanks & St. John (S&S) argue that the case for implicit learning in humans has not been made. Their strategy is to catalog the various implicit learning paradigms and show, in each case, why the inference to *unconscious* learning is unwarranted. Given the logical structure of this approach, the strength of their position depends on the degree to which their catalog of implicit learning paradigms is complete. If they have not examined *all* the likely suspects, their claim that no one fits the bill is substantially weakened. The omission of a careful analysis of human nonhippocampal skill learning is problematic for just this reason.

S&S argue that there are two types of learning: instance learning and rule learning. Both instance learning and rule learning consist of the acquisition of information about the world. Learning a piece of information, be it an instance or a rule, is not the same as learning how to perform a task. Information learning, which includes both instance learning and rule learning, can thus be distinguished from the purely practical element of skill learning. Learning a skill may include, but is never limited to, the learning of information. Once we isolate the purely practical element of skill learning, we will have located a genuine example of implicit learning.

The distinction between practical learning and information learning can be illuminated by reviewing Gilbert Ryle's (1949) distinction between knowing-how and knowing-that. Cohen and Squire (1980) in an article on the spared capacities of human amnesics, misappropriate Ryle's terminology when they claim that, "[The] distinction between procedural or rule-based information and declarative or data-based information . . . is reminiscent of the classical distinction between knowing how and knowing that." [See also Eichenbaum et al: "Two Functional Components of the Hippocampal Memory System" *BBS* 17(3) 1994.] This comparison involves a misreading of the nature of knowing-how, at least as it is understood in the philosophical literature. The possession of information always falls into the category of knowing-that; it matters not if the information is data-based or rule-based. Knowing-how is something other than information possession.

Ryle cautions against the temptation of confusing knowing-how with the knowledge of rules. An agent may know that a particular rule obtains without knowing how to apply it. Knowing the rules of chess, for example, does not ensure skillful playing. Knowing-how itself cannot be explained by appeal to knowing-that. Consider the case of knowing how to apply a rule. If we try to explain this capacity by invoking the agent's knowledge of some "rule-application rule," we will be faced with an infinite regress. The agent's ability to apply the special application rule must then be explained, ad infinitum. At some point, talk of knowing-that (information possession) gives out, and we must recognize the distinctiveness of the purely practical com-

ponent of the agent's ability. Knowing-how is not the knowledge of rules. Learning-how is not the acquisition of rule-based information.

Successful information learning, or learning-that, tells us something about the world. It takes the following general form: an agent learns that x , where x is either an instance or a rule. Practical learning, or learning-how, does not share this form. Practical learning is the incremental modification of perceptual-motor connections over time such that the agent acquires a neurophysiological disposition for improved performance. In the case of information learning, the causally relevant element can be variously described as a set of neurophysiological changes or their semantic correlate – the information acquired. When agents learn that x is the case, we can ask whether or not they are aware of x . In the case of practical learning, the neurophysiological changes have no semantic correlate; no "information" is acquired. There is nothing in practical learning that could count as an object of awareness.

For learning to be considered implicit, agents must be unaware of what they are learning while they are learning it. An agent can be aware of a piece of information. An agent cannot, however, be aware of a neural disposition for enhanced performance. (An agent may be aware of the fact of having acquired such a disposition but cannot be aware of the disposition itself.) Since the causally relevant learning is nothing more than a series of neurophysiological changes, and since agents never have conscious access to neurophysiological changes, both the Information Criterion and the Sensitivity Criterion are satisfied on conceptual grounds. Practical learning is therefore a genuine case of implicit learning.

Are rules and instances subserved by separate systems?

Robert L. Goldstone and John K. Kruschke

Psychology Department, Indiana University, Bloomington, IN 47405-1301; rgoldstone@indiana.edu and kruschke@indiana.edu

Although they find little evidence for an empirical dissociation between unconscious and conscious learning, Shanks & St. John (S&S) do argue for a "dissociation between a rule-induction system and an instance-memorization system" (sect. 1.1). Although the distinction between instance and rule processing is intuitive, we find the conclusion that "it seems clear, then, that there are two separate learning strategies" (sect. 3.2) to be premature.

The first reason for questioning a clear dissociation between systems for processing rules and instances is that the distinction itself is unclear. If an entire episode is remembered in full detail, then this would clearly seem to fall in the domain of an instance processor. But, as is commonly assumed in exemplar models of categorization, what if fewer aspects of the episode are stored? If only one aspect is stored (e.g., an object's color), then it is hard to tell whether an impoverished instance or a simple unidimensional rule is being stored. Furthermore, some exemplar models assume that the information stored from an episode depends on the instances that have already been learned (Aha et al. 1991; Skorstad et al. 1988). For example, an exemplar model might only store an object's color because the object's other features are identical to a previous instance, or the model might not store any information from an episode at all if it is already correctly categorized. At the least, there appears to be a gradual continuum from concrete instances to abstract rules.

The presence of a graded continuum does not show that a distinction is without merit, but it does cast doubt on the postulation of completely separate processing systems. Considerations such as these have led Barsalou (1990) to argue that exemplar and abstraction-based representations are empirically

indistinguishable from each other. Any substantive claim that there are separate exemplar and rule-based systems can be tested only when the "rule" and "exemplar" representations, and their processing, are clearly defined.

The second reason for questioning a simple separation between instance- and rule-learning systems is the existence of complex interactions between the systems. Rules often include instance information. Exceptions to rules are often stored within the rule (Schank 1982). In Nosofsky et al.'s (1994) RULEX model, the same system that devises categorization rules also determines exceptions to the rules. There is a large amount of evidence from analogical reasoning and problem solving (e.g., Ross 1987) that learned abstractions are not stored separately from the concrete instances that give rise to the abstraction. Instead, rules and abstractions are typically attached to their concrete instantiations. Even when the abstraction has been stored, it is often not detached from its specific instantiation (Allen & Brooks 1991). This is difficult to reconcile with the view that separate systems are responsible for extracting rules and instance information. Instance information naturally affects and intrudes upon rules.

Reciprocally, rules affect instance storage. Medin (1986) argues that rules and strategies filter and alter the instances that will be stored. Nosofsky et al. (1989) provide evidence (cited by S&S) that as rule use increases, the amount of preserved instance information decreases. Thus, the instance storage system is not automatic or independent of the rule system. If there are two separate systems at all, then at least they seem to be tightly coupled, rather than autonomous.

Pinker (1991) argued that past tense production in English is based on dissociable rule and instance-association subsystems. Part of the evidence he marshaled was that exceptional verbs are strongly influenced by frequency of use, whereas regular verbs are not; he also suggested that regular verbs, unlike irregulars, do not form "attracting clusters in memory" (p. 253) on the basis of similarity. Neither of those differences arises in simple category-learning experiments, as explained above (e.g., Allen & Brooks 1991). Thus, the type of evidence that has been used to justify an instance/rule dissociation in language learning does not seem to be found in generic category-learning paradigms.

Finally, demonstrations that subjects categorize differently when asked to use rules might be just as consistent with parametric variation within a single system due to instructions as it is with the presence of a system solely dedicated to rule processing. In fact, the former alternative has the advantage of accommodating shared processes for rules and instances. It is likely that many of the same mechanisms (e.g., history-dependent storage, sensitivity to context, and attention to novel or informative aspects) would be part of both instance- and rule-learning systems.

The conclusion is not necessarily that S&S are wrong to hypothesize separate learning systems for rules and instances. However, it does not seem to us that a compelling case has been made for two *separate* systems. Determining that two processes are subserved by separate systems is notoriously difficult. Substantive progress in our understanding of learning is more likely to involve tests of specific, formalized models and empirical exploration of the exact interactions between instance and rule learning. As increasingly rich interactions between the two types of learning are discovered, the impulse to view instances and rules as processed in separable, dissociable systems may subside.

Implicit assumptions about implicit learning

Keith J. Holyoak and Merideth Gattis

Department of Psychology, University of California at Los Angeles, Los Angeles, CA 90024-1563; holyoak@locus.ucla.edu and merideth@psych.ucla.edu

The analysis that Shanks & St. John (S&S) offer of the distinction between implicit and explicit learning appears to be guided by some implicit assumptions. In evaluating the possibility that learning can occur without awareness, S&S adopt the implicit default hypothesis that all learning is conscious unless proven otherwise. In addition, S&S propose explicit and highly stringent criteria for evaluating whether learning can occur without awareness. After concluding that the evidence does not meet their strict criteria for rejection of this default hypothesis, they proceed to accept it, concluding that although two basic learning strategies exist, both operate "within the domain of conscious, explicit learning." In contrast, their criteria for evaluating when learning is instance-based are left implicit and appear to be substantially more lenient. Finally, they assume that the distinction is fundamentally defined by a binary cut on some single dimension: either unconscious versus conscious learning, or learning by instances versus rules. The possibility that learning modes could be usefully characterized in terms of multiple converging properties is not seriously considered. In general, the conclusions reached by S&S appear to depend on their implicit assumptions at least as much as on their objective assessment of the empirical evidence.

S&S point out that a serious question facing anyone interested in investigating the role of awareness in learning is, "What counts as a suitable test of awareness?" (sect. 2.1). The most commonly used tests of awareness have been verbal reports of stimuli or of rules governing the relationships within stimulus sets. S&S correctly argue that researchers have often been too lax in their criteria and have underestimated the contents of awareness and too quickly conceded unconscious learning. However, S&S then set up criteria that make tests of awareness essentially coextensive with tests of learning. Their Information and Sensitivity Criteria state that the measure of awareness must test the same information as the measure of performance and must be sensitive to all conscious knowledge. The combination of these criteria is highly restrictive, making it difficult to distinguish tests of awareness from tests of performance. Furthermore, S&S require not only episodic but also semantic knowledge to be inaccessible for learning to be considered implicit: "in order for us to infer unconscious learning . . . the subject must be unaware of the relevant relationship that occurred in the study episode in addition to being unaware of the episode itself" (sect. 2.3.1). Given such stringent criteria, it becomes impracticable to show that learning ever exceeds awareness.

Because verbal reporting fails the criteria set up by S&S, they suggest that a prediction task (Nissen & Bullemer 1987; Willingham et al. 1989) can provide a more complete measure of conscious knowledge. In this task, subjects predict the location of the stimulus on the next trial by pressing a button matching that location, a task marginally different from the reaction time (RT) task in which subjects press a button matching the location of the stimulus on the current trial. Presumably learning leads to a speed-up on the RT task because subjects learn the sequence of locations, knowledge that could also serve as the basis for prediction. S&S argue that "the rationale behind the prediction task is that if subjects are instructed to try to predict events and are able to do so with above-chance accuracy, this is evidence of conscious knowledge *because their predictions must be based on conscious expectancies*" (sect. 2.7.2, emphasis added). But presumably performances on both the RT task and the prediction task are based on covariational knowledge. S&S go on to say that "the prediction task satisfies the sensitivity criterion where verbal reports did not," but they fail to say how any measure of

awareness that can satisfy the Sensitivity Criterion can be distinguished from a measure of performance. What is being measured in each task, and the relationship between measures of awareness and performance, is therefore left unclear. One possibility is that reaction time, prediction, and verbal report lie on a continuum of accessible knowledge, each determining performance in a different way.

The analysis offered by S&S appears especially strained in their assessment of studies showing apparent implicit learning by amnesics in serial reaction time (Nissen & Bullemer 1987) and artificial grammar tasks (Knowlton et al. 1992). Although in these studies amnesics had little or no detectable memory for the training examples, S&S claim that such findings do not provide firm support for unconscious learning (because S&S would take any evidence of semantic knowledge, in addition to or instead of episodic knowledge, as implying that learning was conscious). Moreover, although elsewhere in the article S&S are favorable to the use of recognition tests to assess explicit knowledge, they nonetheless fail to accept amnesics' failures on such tests (Knowlton et al. 1992) as providing any evidence for unconscious learning. S&S instead suggest that prediction tests need to be performed with amnesics. As noted above, however, it is quite unclear how either prediction or recognition tasks differentiate awareness from performance.

Their treatment of the evidence from studies with amnesics reveals an apparent double standard in the way S&S assess hypothesis about how implicit and explicit learning can be differentiated. On the face of it, such evidence for learning without memory for training examples is surely more compatible with the hypothesis that implicit learning can operate without awareness than it is with the hypothesis that implicit learning is necessarily tied to instance memory. S&S first assume that implicit learning depends on instance learning, then note that amnesics perform well on implicit learning tasks, and on that basis conclude that instance learning must be intact in amnesics. They reach this conclusion despite acknowledging that "amnesics have a genuine problem with a certain class of memorial experience (i.e., episodic memory)." S&S apparently view instance learning as quite distinct from episodic memory, whereas memory for specific training instances has generally been considered an example par excellence of episodic memory in operation. The apparent fluidity in the implicit criteria that S&S use to evaluate evidence of instance learning contrasts sharply with the rigid nature of their corresponding explicit criteria for evaluating evidence of learning without awareness.

In general, S&S appear to assume that the implicit/explicit distinction must rest upon some binary cut along a single dimension. Nowhere do they consider the possibility that two modes of learning might be profitably distinguished not on the basis of some absolute, one-dimensional divide, but rather on the basis of several looser but converging properties. For example, Holyoak and Spellman (1993, p. 278) characterize implicit knowledge as "(a) knowledge about covariations in the environment, (b) learned by exposure to stimuli exhibiting the covariations, (c) obtainable without attention or awareness (although in some cases similar knowledge might be obtained explicitly), and (d) demonstrated by improved performance on tasks that seem to require thinking (e.g., generalization and prediction); but it is knowledge that does not have a fully explicit representation in that (e) it is not fully verbalizable and (f) it is not manipulable in the sense that it cannot be re-represented explicitly to serve as input to other procedures." In addition, following Reber (1992b), Holyoak and Spellman note that implicit knowledge may be distinguished by its longer evolutionary history.

Such a mix of partial regularities is much less "neat" than the simple dichotomies considered by S&S. However, it is such tantalizing convergences between multiple properties, some of which would otherwise seem unrelated, that continue to make the implicit/explicit distinction so compelling as a focus for research on learning and knowledge representation. Clearly,

developing a more precise taxonomy of learning modes remains an important goal for the field. This goal is more likely to be achieved if we avoid implicit assumptions about implicit learning that prematurely restrict theory development.

Development, learning, and consciousness

Mark L. Howe and F. Michael Rabinowitz

Department of Psychology, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X9; mhowe@kean.ucs.mun.ca

The target article addresses an important question, namely, is there any reason to suppose that there exist distinct learning systems that operate using fundamentally different laws? Whether independent systems are required to explain cognitive functioning in general is of both historic and contemporary relevance. The recent literature is replete with examples of this concern in areas as diverse as animal learning (Shettleworth 1993), cognition and cognitive development (Howe et al. 1993), and memory (Jacoby et al. 1993).

In their discussion of this problem, however, Shanks & St. John (S&S) becloud two sets of key issues. The first concerns consciousness. Here, the authors fail to distinguish between three possible meanings of consciousness: the organism as conscious while learning, the organism as being conscious of learning, and the organism being conscious of what is being learned. Nonhuman animals and human infants are typically awake and alert (i.e., conscious in the first sense) when they learn, but are not necessarily aware that they are learning (i.e., conscious in the second sense) or of what they are learning (i.e., conscious in the third sense). Thus, evidence bearing on the first use of consciousness does not inform us about the second or third use of consciousness (although the second or third can inform the first).

The second issue concerns whether S&S are talking about the existence of distinct learning systems *per se* or simply distinct strategies subjects use in different learning contexts. This confusion occurs early in their article, when they introduce the idea that "distinct learning systems encode very different sorts of information; one system induces rules, whereas a second system memorizes instances" (sect. 1). Toward the end of the article, the authors conclude that "there are two separate learning strategies available to subjects, and that these strategies can be invoked by differences in the instructions given to subjects" (sect. 3.2). Although there is no doubt that subjects use different strategies as a function of experimental instructions and other task variables, it is unclear how strategic differences relate to the problem of whether there exists more than a single learning system.

Because of these confusions regarding the nature of learning and consciousness, it is difficult to evaluate S&S's conclusion that (1) all learning is conscious and that (2) there exist two independent systems, one for instance-based learning and the other for rule-based learning. We would argue that distinguishing between the three different meanings of consciousness outlined earlier is crucial to understanding the role that consciousness plays in learning, particularly in a developmental context. Moreover, it is our contention that in the absence of formal modeling it is futile to attempt to demonstrate the existence of independent systems, whether for learning or for memory (Howe et al. 1993).

Perhaps one of the most interesting issues raised in the target article concerns the nature of the relationship between consciousness and learning. A particularly interesting ramification involves the role of consciousness in development. That is, is an immature learning system (e.g., a human infant) conscious when it learns and, if so, in what sense? Certainly this immature system, although not verbally aware, is likely to have been awake and alert while learning. In addition, although the system

cannot provide us with a verbal report, it may know that it has learned and may even know what it has learned.

These possibilities raise an interesting series of empirical, philosophical, and theoretical issues. To begin, consider questions related to being conscious of learning. First, what does it mean to say that an immature, preverbal system knows that it has learned? In other words, how does it represent this knowledge? If such a representation exists, perhaps it is one of the earliest forms of metacognition. Second, how can this knowledge be measured? That is, what sort of behavior would unambiguously index this acquisition: recognition, prediction, transfer and generalization, emotional responses (e.g., indices of happiness), or nonverbally communicating the acquisition to others? Third, what are its implications for subsequent development; that is, is it a deterministic element in the development of the system, one that produces an abrupt, stagelike transition? For example, does the immature system show accelerated learning in a variety of contexts following some demonstration that the system knows it has learned.

Next, consider questions related to being conscious of what is learned. First, what does it mean to say that an immature, preverbal system knows what it has learned? Again, this is a question of knowledge representation. Second, how can this knowledge be measured? Here, unlike in measuring knowledge that learning has occurred, only verbal awareness or some other form of symbolic awareness (e.g., signing or drawing) seems sufficient to demonstrate knowledge of what has been learned. This is because nonsymbolic measures are not necessarily related to symbolic awareness in more mature systems. For example, a subject who could predict a sequence would be aware that the sequence was learned but might not be able to verbalize the prediction rule. Using the rock-paper-scissors game, children and adults have been unable to verbalize prediction rules that they successfully used (Rabinowitz 1993). Third, how does the immature system become verbally aware? Perhaps the simple acquisition of language is sufficient to ensure verbal introspection. Alternatively, this may be a later acquisition of the system, one that is mediated by either the general linguistic experience of hearing others describe thoughts, feelings, and learning experiences or by first using language to facilitate problem solving through the use of verbal hypothesis testing. Fourth, what are its implications for subsequent development? It seems obvious that verbal (or symbolic) awareness facilitates cognitive development. Although there is little empirical data relevant to this issue, one recent study has shown that children learn more quickly and transfer concepts to new domains only if they demonstrate verbal awareness of the concept (Rabinowitz & Howe, *in press*).

To summarize, we believe that S&S have raised a number of important issues related to the relationship between consciousness and learning, issues that will both influence and be informed by developmental research, but we find unconvincing their arguments concerning the need for more than one learning system and we believe that the resolution of this issue may not be critical to understanding the role of consciousness in learning. It is our further contention that to understand the relationship between consciousness and learning, the term "consciousness" needs to be more precisely delineated.

ACKNOWLEDGMENT

Preparation of this article was funded by Natural Sciences and Engineering Research Council of Canada grants OGP0003334 to MLH and OGP0002017 to FMR.

Human autonomic conditioning without awareness

H. D. Kimmel

5813 Mira Serena, El Paso, TX 79912; dq00@utep.bitnet

Shanks & St. John (S&S) find no credible evidence that autonomic responses are conditionable in humans who are unaware of the stimulus-reinforcement contingencies administered. Conflicting evidence is either dismissed on methodological grounds (e.g., their Sensitivity Criterion is not met) or simply ignored. This commentary presents a few examples of the latter.

The role of awareness in human conditioning has a long history. Almost thirty years ago, Grings (1965) reported that nonverbal deaf children and severely mentally retarded, nonverbal adolescents (mean IQ = 33) acquire classical autonomic CRs as readily as college students. Grings concluded that conditioning does not require awareness in the usual sense of the word. S&S would probably argue that no adequate sensitivity test was reported. True! No such test could even have been administered to Grings' subjects. Grings also found that subjects shocked for responding autonomically to a cue made fewer responses to that cue, whereas subjects shocked for not responding to the cue made more responses; when questioned after three days of conditioning, these subjects revealed no knowledge about stimulus relations. These early reports of human conditioning without awareness are now forgotten relics of the past. Also ignored are the many studies showing operant autonomic conditioning in humans unable to identify what they might have done to influence the reinforcer, even when told that the reinforcement depended on something they did (summarized in Kimmel 1974). Data from subjects who reported doing (or thinking of) things to influence the reinforcement were explicitly excluded from these reports.

Manifestations of aversive conditioning without awareness may be quite diffuse. Kimmel et al (1983) found that illuminating a room with a colored floodlight (blue or yellow), in the presence of which a few brief mild shocks had been administered a few minutes earlier, influenced evoked and slow brain potentials as well as cardiac and electrodermal reactions to an auditory reaction time stimulus. No shocks were presented during the reaction time task. These effects occurred as frequently in subjects who could not state whether the shocks had occurred during the blue or the yellow floodlight as in subjects who accurately verbalized the color-shock relationship. S&S might claim that the unaware subjects would have been found to be aware if they had been interrogated more effectively (they were university students who stated unhesitatingly that they had seen both the blue and the yellow floodlights – the “unaware” ones were either unable to say which color was associated with shock or said incorrectly that shock was associated with both).

S&S prefer to distinguish between instance and rule learning as opposed to learning with and without awareness. They question whether subjects can learn implicit rules without awareness, arguing that specific fragments or instances are what subjects actually report. Lachnit and Kimmel (1993) found rule learning in a Pavlovian procedure in which two CSs were each separately paired with shock but were presented together without shock (negative patterning), and two different CSs were separately presented without shock but were paired with shock when administered together (positive patterning). Subsequently, subjects were told that shock administration may have followed a special “rule” and were asked if they had experienced such a rule. Subjects who answered affirmatively were asked what the rule was. Half of the subjects described specific instances (e.g., which stimuli were paired with shock) rather than stating abstract rules. In addition, ten subjects reported a rule more or less accurately, while another ten neither stated a rule nor gave specific instances. There was no difference be-

tween the latter two groups in acquisition of patterning or in transferring the rule to new stimuli.

There is ample evidence that informing subjects about CS-reinforcement relationships enhances subsequently observed CRs, but these effects are not always unambiguously associative. Giving other kinds of information about impending events, for example, telling the subject about future tasks (Kimmel 1967) or about an intelligence test that will be administered later (Kimmel & Bevill 1985) can produce similar effects. Even if instructions indeed influence human conditioning associatively, this does not necessarily mean that only subjects who can describe the conditioning contingency verbally can be conditioned. S&S are not convinced that lack of verbal recall is a sufficiently sensitive basis for concluding that conditioning without awareness has occurred. We obviously disagree concerning the determination of awareness. The subjects in the foregoing studies were questioned very closely, and the questioners were motivated to detect any sign of awareness. At the very least, Shanks & St. John should describe what they consider inadequate interview procedures from studies claiming conditioning without awareness, instead of rejecting them peremptorily or acting as if they do not even exist.

Can procedural learning be equated with unconscious learning or rule-based learning?

Zoe Kourtzi,^a Lindsay M. Oliver,^b and Mark A. Gluck^b

^aDepartment of Psychology; ^bCenter for Molecular and Behavioural Neuroscience, Rutgers University, Newark, NJ 07102;

^akourtzi@draco.rutgers.edu, ^boliver@pavlov.rutgers.edu, and ^cgluck@pavlov.rutgers.edu

The aim of Shanks & St. John's (S&S's) target article is to determine whether or not two independent learning systems exist, one based on declarative or explicit learning, the other based on procedural or implicit learning. S&S approach the question of independent learning systems in terms of two distinctions: conscious versus unconscious learning, and instance-based versus rule-based learning. We argue here that these two distinctions do not capture the intentions of the researchers originally making the differentiation between declarative and procedural learning, or the implications of the data that motivated this distinction. The distinction between declarative and procedural learning can be better understood in terms of stimulus recoding in learning. To this end we adopt a model of hippocampal function (Gluck & Myers 1993) in animal learning to restate the distinction between declarative and procedural learning in terms of stimulus representation.

S&S define implicit or procedural learning as learning based on *unconscious rule induction*. They evaluate evidence in favor of implicit learning focusing primarily on studies of intact subjects. They conclude that the existence of unconscious learning has not been satisfactorily established, but that both instance-based and rule-based learning occur dependent on the type of the stimuli and the nature of the task. The definition of implicit learning as unconscious, rule-based learning, combined with the lack of evidence for unconscious learning suggest that implicit learning is learning based on *conscious rule induction*. As noted by S&S, the distinction between declarative learning and procedural learning was originally intended to refer to the difference between the learning of facts and the learning of skills and arose in part from the observation that amnesics are impaired in declarative learning but not in procedural learning (Cohen & Squire 1980). The terms *explicit* and *implicit* were used to replace *declarative* and *procedural* with the discovery that amnesics are capable of more than skill-based learning (e.g., Graf et al. 1984; Squire & McKee 1992). However, the

notion of implicit learning as rule-based learning suggested in the target article does not seem to capture the essence of the kind of learning of which amnesics are capable. Here we describe a model of hippocampal function in animal learning. We use this model to reframe the distinction between declarative and procedural learning in terms of stimulus representations, and give a beginning account of how the model might explain the human data that first motivated the distinction.

Recently, Gluck and Myers (1993) proposed a theory of cortico-hippocampal interaction in classical conditioning. The theory allows for two kinds of learning, learning that depends on stimulus recoding and learning that can proceed on the basis of stimulus response associations alone. We believe that this distinction, as it is incorporated in the cortico-hippocampal model, provides a better basis for reframing the original distinction between declarative and procedural learning. The theory is instantiated by a network model consisting of two modules, in which the hippocampal module forms new stimulus representations that facilitate learning. The cortical module is incapable, by itself, of forming novel stimulus-stimulus-based recodings but uses the new hippocampal representations as a basis for forming new stimulus-outcome representations. Within this framework, a hippocampal lesion is characterized by removing the hippocampal module. When the hippocampal module is lesioned, the system learns associative stimulus-response relationships based on this fixed encoding of the stimulus inputs. In comparison, the intact system learns the same stimulus-response (S-R) relationships but does so based on a flexible and distributed recoding of the stimulus inputs that reflect both predictive S-R relationships as well as sensory-sensory correlations.

The Gluck and Myers (1993) cortico-hippocampal model rests on the assumption that the role of the hippocampus in associative learning is the development of new stimulus representations that facilitate learning. We propose that the hippocampus may play a similar role in human learning, and that the deficits seen in hippocampally damaged amnesics reflect an inability to form new stimulus representations. The cortico-hippocampal model predicts that the hippocampus is required for any learning in which stimulus recoding is necessary. Under the model, failures in declarative learning can be understood in terms of failures to combine in memory those elements of the environment that go together to form a specific declarative memory. The elements of a declarative memory are in many cases encountered frequently in everyday life; it is the unique combination of those elements that makes the memory unique. The inability to recode stimulus representations in a way that takes into account configurations of stimulus elements precludes the learning of declarative information.

The Gluck and Myers model could account for the fact that amnesics are impaired in declarative learning, because declarative learning requires stimulus recoding that depends on an intact, functioning hippocampus. In contrast, learning for which amnesics show little or no impairment would be best characterized as being based on fixed associative S-R relationships and therefore, can proceed without stimulus-stimulus (S-S) recoding. The model predicts that hippocampally damaged amnesics will be capable of just such tasks. For example, Knowlton et al. (1993) found that amnesic subjects performed as well as normal subjects early in training in a probabilistic categorization task. Such a task can proceed on the basis of S-R associations and hence the model predicts no impairment for amnesics. The term *procedural* fails to characterize this kind of learning, because certain tasks that are not skill based can be learned on the basis of associative S-R relationships, without S-S recoding. For example, the priming effects reviewed by S&S (Graf et al. 1984), in which amnesics maintain the ability to complete word stems, can be accounted for without reference to stimulus recoding, and hence the model expects that amnesics will show no impairment.

The above studies show the importance of distinguishing between S-S representations and S-R associations. This distinction is not captured by the narrow definition of implicit learning used in the target article. Gluck and Myers' model suggests that there are two kinds of learning, hippocampus-mediated S-S representational learning and hippocampus-independent S-R associative learning, and animal lesion data support this distinction (Gluck & Myers 1993) in animal learning.

ACKNOWLEDGMENTS

Mark Gluck's research was supported by the Office of Naval Research through the Young Investigator program and by a grant from the McDonnell-Pew Foundation.

Consciousness in natural language and motor learning

Joel Lachter

Department of Psychology, University of Rochester, Rochester, NY 14627; lachter@psych.rochester.edu

In their target article, Shanks & St. John (S&S) propose that all learning is conscious. This proposition raises some interesting issues that the authors do not spell out. In this commentary, I would like to raise such issues in two areas: natural language and the development of motor skills.

In their introduction, S&S acknowledge the importance of language as a *prima facie* argument against their claim that one cannot learn unconscious rules. One presumes that they are referring to the ubiquitous example of natural languages. Yet they completely ignore natural language learning in the rest of their paper, preferring to discuss artificial language learning. There are at least two relevant differences between natural and artificial language learning. First, the evidence for the use of rules is much stronger in the case of natural language learning. Children generally achieve much higher proficiency in their natural language than is obtained in the artificial language learning experiments cited by S&S. In addition, children make spontaneous categorical judgments about the acceptability of sentences (Lachter & Bever 1988; Slobin 1978). This makes a strategy of applying memorized phrases unlikely. Second, children generally learn natural languages, whereas the artificial language experiments cited by S&S were all run on adults. People's ability to learn natural languages drops sharply as they grow older. Current research shows that, as the age of first exposure increases, performance declines and appears to be based on the probabilistic application of set patterns for both first (Newport 1990) and second (Johnson & Newport 1989) language learners. This sort of performance is similar to S&S's description of artificial language learners.

There is much debate surrounding the need for explicit rules in natural language and whether such rules are innate. In the context of such debates it is easy to lose track of two simple but widely agreed upon facts that, taken together, imply that children can learn aspects of language of which they are not conscious: people are not conscious of the reasons for their grammatical intuitions and some of what underlies these grammatical intuitions is learned.

Even proficient language users are usually not aware of why some sentences are grammatical and others are not. For example, why should the sentence, "John was assigned the task that I didn't know to whom they would entrust," be ungrammatical in English, while the sentences, "John was assigned the task that I didn't know they would entrust to him," and "I didn't know to whom they would entrust the task," are acceptable. Of course, if these judgments were based on innate knowledge it would not be a problem for S&S's thesis. However, languages differ in which constructions are allowed. For example, "l'incarico che

non sapevi a chi avrebbero affidato," (the Italian equivalent of "the task that I didn't know to whom they would entrust") is a grammatical phrase in Italian (Sells 1985; Roberto Zamparelli, personal communication 11/29/93). If the same construction is grammatical in one language and ungrammatical in another, children must learn this distinction. But what could they be consciously aware of that would lead them to learn the English rather than the Italian "rule"? Clearly, this thought experiment will never pass S&S's Information Criterion because there will always be information available that is correlated with what we suppose the child to be learning. A child endowed with all the possible human languages might be able to differentiate Italian from English by the presence or absence of trilled Rs; something children might be presumed conscious of. It is likely that one can differentiate all human languages on such a basis. However, although it is easy to find such correlates, it is hard to find ones that make a plausible basis for how children actually learn language. I believe the onus is on S&S to describe some form of information that children might plausibly be conscious of, that could give rise to linguistic intuitions.

Another issue raised by S&S's thesis concerns what people are conscious of while learning. This issue may be clearest in the case of the development of a motor skill. Think of learning to drive a car, learning to play a musical instrument, or learning to control a computer mouse. Typically, people receive instruction on how to perform such a task. They then repeatedly attempt to perform the task with increasing degrees of success. Thus, there is a two-stage learning process: first, one learns explicit rules for carrying out a task; then, one applies these rules as part of a controlled process. With repeated application, performance of the task becomes automatic, to the point where it can be done without conscious monitoring or intervention (James 1890; Schneider & Shiffrin 1977; Shiffrin & Schneider 1977). The learning involved in the second stage of this process differs from the first in that there is no conscious access to what is being learned. Certainly, there is conscious access to the stimulus and the response. But the learner knows what the appropriate response to the stimulus is at the *beginning* of training. What is learned is not the stimulus-response pairs but a new way of generating the response from the stimulus, and there does not appear to be conscious access to that new process (James 1890; Shiffrin & Schneider 1977; Velmans 1991).

The problem with the Shanks & St. John thesis appears to be a failure to appreciate the difference between the content of conscious experience and the content of the memory traces left by that experience. If one consciously attempts to remember that Mary's phone number is 484-2100, but the memory trace laid down by the experience is that Mary's phone number is 848-2100, it is the latter that one has learned. Of course, in the normal case, what is learned is what the person is consciously attempting to learn. What is interesting about natural language and motor learning is that this does not appear to be the case.

Tacit knowledge and verbal report: On sinking ships and saving babies

R. O. Lindsay^a and B. Gorayska^b

^aPsychology Unit, Oxford Brookes University, Oxford OX3 0BP, England;

^bDepartment of Computer Science, City of Hong Kong Polytechnic, Kowloon, Hong Kong; ^arolindsay@uk.ac.oxford-brookes and

^bsgoray@uk.cphk.cphkvx

The target article examines the proposal made by various authors that two distinct learning systems exist, one responsible for "explicit" learning, the other for "implicit" learning. Shanks & St. John (S&S) suggest that this proposal confounds two underlying dimensions: learning with and without awareness, and learning based on rules as opposed to learning based on

instances. Relevant studies are carefully reviewed and it is concluded that "unconscious learning has not been satisfactorily established" and that "in tasks such as artificial grammar learning, where learning is frequently thought to involve rule abstraction, performance is most often based on the acquisition of instances or fragments from the training stage."

The position from which S&S argue is oddly reminiscent of that of beleaguered behaviourists, earnestly examining the inside of their theoretical ship for the slightest sign of metaphysical leakage. The difference is that S&S seem to fear an ingress, not of consciousness but of unconsciousness. Our position is quite the converse of that taken by S&S: tacit knowledge which is unavailable to awareness is both more basic and more frequently encountered in learning systems than explicit knowledge; conscious, verbalizable knowledge is a special case. The methodological caution that S&S evince and commend is highly laudable, and we appreciate and accept the importance of seeking to ensure in the case of particular claims of knowledge without awareness that tests focus on relevant knowledge, and that a reasonable case can be made for comparability of sensitivity in awareness tests. There is a real danger, however, that S&S have unwittingly allowed methodological rigour to slide into logical scepticism. How is a proponent of learning without awareness to demonstrate that no test exists which could – if only it were relevant or sensitive enough – reveal awareness after all?

This demand would be reasonable enough if criteria for relevance were readily available, or if sensitivity scaling were a trusted technology, but when neither of these things is true, the position taken by S&S, no doubt in good faith, can easily become perverted into a formula which justifies the automatic rejection of all claims that unconscious learning occurs. Some might be tempted to say that this would be no bad thing, and that indeed the psychological ship can only be expected to founder if such metaphysical substances as unconscious learning begin to seep aboard. We believe that this scepticism is misguided, and that learning without awareness is not a substance which undesirably intrudes but is part of the stuff and fabric of which the vessel is built.

S&S seek to drive a wedge between verbal reports and awareness. They claim, for example, that "we have little reason to believe that the verbal report test provides an exhaustive index of conscious information, since there are other tests such as recognition that manifestly detect information left undetected by verbal report tests." This is clearly a *non sequitur*, because although the recognition test establishes the presence of information, it does nothing to establish that the subject is conscious of it. Nor indeed is it clear that any criterion could establish this, other than a verbal report, or a response to a verbal instruction such as "only press the button if you are aware that *x*." Accordingly, we continue to maintain the traditional view that awareness can be established only via verbal reports. S&S have provided no grounds for rejecting this criterion.

We can now ask the question "Are there grounds for believing that learning occurs which a subject is unable to report?" The answer is plain: all animal learning and all learning in prelinguistic children falls into this category. We think there is abundant evidence that awareness is associated with symbolic planning processes which cannot become established without nonsymbolic learning, which is probably best described by connectionist models. Such learning is, of necessity, unavailable to awareness at the time it occurs. As the mechanisms for nonsymbolic learning presumably continue to exist even when the capability for symbolic planning is fully established, there is every reason to suppose that learning that is nonsymbolic and unavailable to awareness can occur in adults.

A systematic defence of this position was outlined by MacKay (1974), almost two decades ago. It is worth quoting MacKay at some length, as the target article by S&S suggests that his powerful insights are in danger of being forgotten:

We must distinguish between what *we* can say we know and what a suitably equipped *observer* could say we know. . . . It is scientifically inappropriate to regard knowledge we can express in words as paradigmatic, and tacit knowledge as a special case. What we need from the outset is a methodology that can cope with tacit knowledge, taking verbalizable knowledge as a special case. . . . Since the conditional readiness of a physical information system can be modified by essentially continuous and overtly undetectable changes in thresholds or in coupling elements determining transition probabilities . . . as well as by discrete changes in connectivity, the knowledge represented by such modifications might well be unverbalizable for a special reason; namely that it can manifest itself only when and if an appropriate challenge "interrogates" the state of readiness. In other words, much of the "stored information" in such a system can be revealed only by the way it runs and is inaccessible to any other form of inspection. It is thus doubly unverbalizable. (MacKay 1974, p. 94)

Our conclusion is that there are good reasons for believing that knowledge without awareness is commonplace and inevitable, and that in the absence of any contrary reasons, verbal report should continue to be taken as criterial for awareness. The emphasis which Shanks & St. John place upon methodological caution in evaluating particular claims that implicit knowledge exists is commendable, but can cause bathwater to become confused with babies.

Implementational constraints on human learning and memory systems

Chad J. Marsolek

Department of Psychology, University of Arizona, Tucson, AZ 85721;
marsolek@convx1.ccit.arizona.edu

Shanks & St. John (S&S) argue that the distinction between learning instances versus rules is a more useful way to characterize the architecture of human learning systems than the distinction between conscious and unconscious learning. They emphasize that the content of the stored material illuminates these systems more effectively than considerations of whether subjects are aware of stored material. I applaud this emphasis and agree with the conclusion that learning and memory systems respect the instances-rules distinction. It appears, however, that S&S do not consider a vitally important source of relevant evidence. In theorizing about any systems, what is known about their *implementations* serves to place strong constraints on theory. For example, hemispheric specialization of cortical processing systems places interesting limits on S&S's conclusions.

Recent work suggests that a specific-visual-form (SVF) system operates relatively independently of an abstract-visual-form (AVF) system to underlie visual memory (Marsolek, submitted; Marsolek et al. 1992; Marsolek & Schacter, submitted; Marsolek et al., submitted). These systems store information differing in content. An SVF system effectively stores visual details (e.g., whether a *Y* or *L* vertex appears in the upper left of a form) in order to distinguish specific instances of a type of form (e.g., *p* vs. *P*). In contrast, an AVF system effectively stores invariant features (e.g., the proposition that an enclosed space is upper-left-connected-to a vertical line) in order to recognize types of forms (e.g., a *p/P* regardless of letter case).

This theory is interesting in the context of S&S's target article, because the former system stores specific instances whereas the latter system stores abstract information similar to rules about types of visual structures. As S&S discuss, cognitive research supports both the view that specific instances are learned and the view that abstract rules are learned in memory tasks. I suggest that neuropsychological and neuroscientific evidence is needed to supplement this picture.

Hemispheric asymmetries indicate that a SVF system operates more effectively in the right cerebral hemisphere (RH) than

in the left (LH), whereas an AVF system operates more effectively in the LH than in the RH. After learning new types of visual forms, subjects classify the previously unseen prototypes of the newly learned types more effectively when they are presented directly to the LH (in the right visual field) than to the RH (in the left visual field). In contrast, they classify previously seen exemplars more effectively when they are presented directly to the RH than to the LH. Prototype forms possess the features that do not vary across the different instances in one type of form, and these invariant features are valuable clues for recognizing types of forms. The information that is found in prototype forms and that is useful for recognizing types of forms is therefore stored more effectively in the LH than in the RH (Marsolek, submitted).

In addition, when subjects complete a word stem (e.g., "con") to form the first word that comes to mind, they tend to produce a completion that forms a word viewed earlier in the experiment. This priming effect is greater when stems are presented in the same letter case as previously presented completion words than in the different case, but only when stems are presented directly to the RH. Same-case priming is not greater than different-case priming when stems are presented directly to the LH. Thus, the visual details that distinguish specific instances of a type of form (e.g., lower- vs. uppercase versions of the same word) are stored more effectively in the RH than in the LH (Marsolek et al. 1992; submitted). These findings are consistent with S&S's conclusion that the instances-rules distinction helps to characterize dissociable human learning and memory systems.

One reason why implementation-level constraints are vitally important for further characterizing these systems is illuminated by the following example. In additional word stem experiments, subjects deliberately recollect words presented earlier in the experiment to help them complete stems. In this situation, same-case memory is greater than different-case memory when stems are presented directly to the LH, but not when stems are presented directly to the RH. Thus, when cued recall is measured, information about letter case is stored more effectively in the LH than in the RH, contrasting with the results obtained when priming is measured (Marsolek et al. 1992). These results complicate the conclusion that a simple instances-rules distinction is important when taxonomizing human learning.

One might argue that these results support the hypothesis that four systems underlie human learning and memory. Subjects are not encouraged to recollect consciously information from a previous learning episode when word-stem completion priming is measured, although they are when word-stem cued recall is measured. Thus, an unconscious instance-based system (in the RH), an unconscious rule-based system (in the LH), a conscious instance-based system (in the LH), and a conscious rule-based system (in the RH) may operate to underlie the results in these experiments. If so, it would appear that the instance-rule distinction that S&S support is only half of the story, and the unconscious-conscious distinction is needed to complete the picture.

Alternatively, even without resorting to the very difficult task of assessing whether word-stem completion priming can be considered an example of unconscious learning (for research related to this issue, see Bowers & Schacter 1990), further implementational constraints help to clarify the situation. Research examining functional anatomy through positron emission tomography, impaired abilities following brain damage, and other evidence indicates that relatively independent cortical systems store perceptual versus conceptual information (see Schacter 1990; 1992; Tulving & Schacter 1990). Furthermore, additional studies of hemispheric specialization indicate that precise and focused semantic information is processed more effectively in the LH than in the RH, whereas diffuse and wide-ranging semantic information is processed more effectively in the RH than in the LH (see Beeman 1993; Burgess & Simpson

1988; Chiarello 1988). With these further constraints, one can focus on the content of the material stored in different systems when characterizing them, as S&S attempt to do. Assuming that conceptual knowledge about letter case information can be stored in nonperceptual systems and that priming effectively taps information stored in perceptual systems whereas cued recall effectively taps information stored in conceptual systems, perhaps the best conclusion to draw is the following. A specific perceptual system (in the RH), an abstract perceptual system (in the LH), a specific conceptual system (in the LH), and an abstract conceptual system (in the RH) operate relatively independently to underlie human learning and memory.

To the extent that this conclusion satisfies a large number of constraints from neuropsychology and neuroscience as well as from cognitive studies of learning and memory, it is preferable to the alternative conclusion involving a conscious-unconscious learning distinction. Implementational evidence appears to be necessary to develop fully theory in human learning and memory.

On the futility of attempting to demonstrate null awareness

Philip M. Merikle

Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1; pmerikle@watarts.uwaterloo.ca

The single most important message that can be taken from Shanks & St. John's (S&S's) target article is that it is futile to try to use single dissociations to demonstrate the existence of implicit or unconscious learning. By and large, both the proponents and the opponents of implicit learning accept the legitimacy of single dissociations as a means of deciding whether or not learning can occur without concurrent conscious awareness. However, there is a fatal flaw in the logic underlying this version of the dissociation paradigm. Before it is possible to demonstrate implicit learning, it is necessary to demonstrate that there is absolutely no conscious awareness of task relevant information. The obvious problem is: How can it ever be established that any behavioral measure is a satisfactory exhaustive measure of all potentially relevant conscious information (cf. Reingold & Merikle 1988)?

This weakness in the logic underlying studies of implicit learning is a weakness that is shared with many other studies of unconscious influences. In fact, for more than 100 years, the predominant method for studying possible unconscious influences in perception, memory, or learning has been to establish single dissociations between two different measures. In all of these studies, one of the two measures is assumed to reflect the influence of relevant conscious knowledge, while the second measure is assumed to reflect perception, memory, or learning independent of relevant conscious information. The experimental challenge has been to demonstrate that information is perceived, remembered, or learned even when it is completely unavailable to awareness, as indicated by the null sensitivity of a measure assumed to index relevant conscious knowledge.

The reason that more than 100 years of research with the dissociation paradigm has not led to definitive answers is that any demonstration of null sensitivity for an assumed measure of awareness can always be interpreted in a number of different ways. The exhaustive nature of an assumed measure of awareness can almost always be questioned, as it is difficult if not impossible ever to justify completely that any particular measure is sensitive to all relevant conscious information. However, the exhaustiveness assumption must be made before any simple dissociation provides strong evidence for an unconscious influence. If the exhaustiveness assumption cannot be justified, a dissociation between measures may simply indicate that the

measure of conscious knowledge was insensitive to at least some potentially relevant conscious information.

Rather than seeing the exhaustiveness assumption as a barrier to successful experiments, S&S embody it in their Information and Sensitivity Criteria. Taken together, these two criteria state that an adequate test of awareness must be sensitive to *all* potentially relevant conscious knowledge. S&S imply that it should be possible to satisfy the exhaustiveness assumption and to conduct satisfactory studies. I disagree. I think the exhaustiveness assumption is difficult if not impossible to justify. For this reason, a skeptic will probably always be able to argue that the experimental manipulations in any particular experiment failed to rule out awareness for at least *some* potentially relevant conscious knowledge.

For evidence of the difficulty in ever satisfying the exhaustiveness assumption, one need only consider the conclusions reached by S&S. After reviewing much of the relevant research on implicit learning, they conclude that unconscious learning has not been satisfactorily established. The primary basis for this conclusion is that by and large most studies have failed to demonstrate convincingly that null performance on an assumed measure of conscious knowledge actually indicates null awareness of all potentially relevant information.

Given the problematic nature of the exhaustiveness assumption, I doubt that it will ever be possible to establish the existence of an unconscious influence in learning, memory, or perception simply by demonstrating a single dissociation between measures. I am confident that as long as a demonstration of null awareness for all potentially relevant knowledge is a prerequisite for a convincing demonstration of learning without awareness, controversy will continue to characterize this area of research. The large number of indecisive empirical studies suggests that establishing single dissociations is not a very productive research strategy. I am certain that continued attempts to use single dissociations to document unconscious influences will only serve to continue current controversies.

A better strategy than attempting to "prove" or "disprove" the existence of learning without awareness is to assume that both conscious and unconscious influences exist and then to develop experimental techniques that can distinguish between these two types of influences. This research strategy has proven successful in studies of both perception and memory (cf. Jacoby et al. 1992; Merikle 1992). To date, however, this research strategy has not been adopted to any great extent in studies directed at demonstrating implicit or unconscious learning. The main advantage of attempting to distinguish conscious from unconscious influences is that it directs research toward questions concerning how these two types of influences may differ. If the distinction between implicit and explicit learning can be shown to predict qualitatively different outcomes, then it is both interesting and important. However, if the distinction only predicts quantitative differences in performance, then there is reason to question whether it helps to advance our understanding of human cognition in any important way.

Faulty rationale for the two factors that dissociate learning systems

Hiroshi Nagata

School of Health Sciences, Okayama University, Shikata, Okayama 700, Japan

As Shanks & St. John (S&S) argue, "the aware/unaware and rules/instances dimensions are logically distinct (sect. 1.1, para. 5)" and crossing the two dimensions produces four hypothetical learning systems (see sect. 3, para. 1). Such argumentation is reinforced when Reber (1989a, p. 219) characterizes implicit learning as "the process by which knowledge about the rule-

governed complexities of the stimulus environment is acquired independently of conscious attempts to do so." However, I shall claim that S&S's interpretation of Reber's arguments on the findings in artificial grammar learning studies is flawed and that their rationale for hypothesizing the two factors that dissociate learning systems is not supported. This is particularly true for the rules/instances dimension.

1. Rules versus instances. S&S's instance-storage theory is sufficient to account for performance in previous artificial grammar learning studies only if the theory incorporates a process that Brooks and Vokey (1991) call "abstract analogy." This appears to rule out an alternative rule-induction theory, but I do not think that S&S are actually arguing about something that is qualitatively different from Reber. This is because an ordered set of bigrams is emphasized as an important cue when sorting grammatical from ungrammatical strings (Reber & Allen 1978), and because the functional mental representation of a finite-state grammar is considered to involve not a formal Markovian system as a target rule but the ordered set of bigrams and trigrams (Reber 1989a, p. 226). Reber provides no specific mention of a relationship between this ordered set of bigrams/trigrams on the one hand, and what he often describes as "invariance patterns of the stimulus environment" (p. 226) or "structural patterns of relational invariances that the environment displays" (p. 229), on the other. However, it is safe to presume that these refer to the same thing. Because Reber regards rules as "complex interdependent covariations among events" (p. 231), we must conclude that the ordered set of bigrams/trigrams is exactly what he considers rules. For if rules constitute the abstract knowledge acquired through implicit learning, so does the ordered set of bigrams/trigrams.

S&S's instance-storage theory, too, gives a major role to fragments in a grammaticality judgment task. It is evident, further, that S&S attach more importance to fragments than to instances as a successful cue for grammaticality judgments (see sect. 2.5.1, para. 15; sect. 2.5.2, para. 3). Since for S&S subjects' knowledge "consists of instances or fragments of the training stimuli" (abstract), the critical issue we must examine is whether S&S allow a certain type of order to relate the fragments to each other. The target article is not at all clear on this point, namely, whether or not order information is essential in the grammaticality judgments. To support their instance-storage theory, S&S cite the Dienes et al. (1991) study, however, which demonstrates subjects' sensitivity to illegal orders of legal pairs (bigrams). This suggests that S&S, too, regard order information as no less important. It should also be noted that this information is obviously indispensable if subjects are given, instead of a usual classification task, a production task in which they are required to produce whatever letter strings they think are legitimate.¹ Hence, S&S must include this order information in their instance-storage theory to make it viable.

Thus, there is no substantial difference in the knowledge that S&S and Reber believe subjects develop during a learning phase. The only difference between them concerns which side of the coin they put special emphasis on. S&S favor a more specific side of the coin, attending to instance-based information, while Reber favors a more general ("abstract" is his term) side, attending to rule-based order information. The coin itself consists of ordered sets of bigrams/trigrams. In this respect, S&S are quite right. Yet the problem is whether the type of knowledge acquired (rules vs. instances) constitutes a reasonable dimension in hypothesizing separate learning systems.

2. Implicit versus explicit. Reber's characterization of implicit learning as an unconscious process hints that implicit and explicit processes are mutually exclusive, comprising two branches of a binary alternative. However, Reber (1989b, p. 243) does not define a clear boundary between the two processes when he remarks that his "subjects have some conscious awareness of rules." He further adds that what he insists on is not that $\alpha = 0$ but that subjects' conscious knowledge is "insufficient to

account for their behavior; that is, $\epsilon > \alpha$, even though $\alpha > 0$ " (p. 243), where α is the knowledge available to consciousness whereas ϵ is that available to unconsciousness. This does not imply, of course, that Reber views implicit and explicit learning as having a single learning system in common. Quite the contrary. He favors an implicit learning system that is cognitively distinct from the explicit one.²

The studies S&S cover in the target article provide reasonable support for their position that the implicit learning system has not been satisfactorily established. Notice, however, that although prior implicit learning evidence could be accounted for without reliance on independent implicit processes, a substantial amount of human knowledge and skill is acquired in an incidental or unintentional manner. This state of affairs is particularly evident in language learning.³ We know how to understand and produce sentences in our native language, even though we may not be able to articulate explicitly the abstract principles or rules (where it is reasonable to assume them⁴) used to create or parse such sentences. Even linguists and psycholinguists are far from being able to do that. It is doubtful if every domain of knowledge could be accounted for satisfactorily by S&S's explicit learning system. Their system postulates, by definition, that every aspect of knowledge is known to learners.

I would like to know whether S&S's conclusion remains unchanged when we assume a single, instead of a dual, learning system,⁵ for which the level of awareness ranges from explicit to implicit states on a continuum. This learning system encodes, stores, and represents a variety of objects and events including, for example, their attributes, relations, and covariations. It is accordingly devoted to encoding rule-like properties as well as specific instances involved in a stimulus environment. Some types of knowledge acquired by this system become known to learners relatively explicitly ($\alpha > 0$), while the other types of knowledge remain less explicitly known ($\epsilon > \alpha$). Even in a single knowledge domain, there are varying gradations of awareness. Thus we can offer unsophisticated but rather explicit explanations for some portion of our knowledge, while we are entirely unaware of other portions.

It is worth adding that the knowledge requiring the mechanism of Brooks and Vokey's (1991) "abstract analogies," which S&S defend, is evidently less explicit in awareness than the knowledge not requiring that mechanism. This is typically illustrated in the observation that generalization to same-set (letters unchanged) items is greater than that to novel-set (letters changed) items (e.g., Whittlesea & Dorken 1993).

3. Conclusions. Because there is no substantial difference between S&S and Reber with respect to knowledge acquired, S&S's attempt to make the rules versus instances difference an independent dimension for hypothesizing dissociable learning systems is invalidated. The implicit versus explicit dimension as a state of awareness – as opposed to that of a learning system – should not be dispensed with in accounting for the acquisition of a range of knowledge and skills that is not covered by S&S.

NOTES

1. Note that studies are lacking which have used production as a performance measure.

2. See Reber (1989a) and Reber et al. (1991) for a sketch of his evolutionary model of implicit processes.

3. Recall that Reber (1967) started implicit learning studies in the present style to elucidate the mechanism underlying language learning.

4. Whether native speakers *know* the principles or rules of their language as espoused by generative theorists has not been empirically established, I believe. See Nagata (1988; 1989a; 1989b; 1989c; 1992) concerning the instability of linguistic intuition, and Nagata (submitted) concerning native speakers' insensitivity to Chomsky's (1981) binding principle.

5. Whether this learning system can be defined as exclusively explicit or exclusively implicit cannot be determined.

The intuitive mind

Geir Overskeid

*Institute of Psychology, University of Oslo, 0317 Oslo, Norway;
geir.overskeid@psykologi.uio.no*

Shanks & St. John (S&S) present solid arguments, as do Cleere-mans (1993b) and others, in favor of the view that there are two types of conscious learning. S&S are hard to follow, however, when they claim that there is therefore no *unconscious* way of learning.

Most people feel that their conscious decisions matter – and conscious decisions probably often do, even though our decisions in turn are determined by other events. When humans act, however, it is not always because we (our conscious selves, that is) have decided to act. There is much cited evidence indicating that the decision may often be taken for us by our brain, and that we jump on a running train believing we started it – namely, the conscious “decision” to act comes after the act has been initiated (Libet 1985). When we learn, decisions about when to act are often crucial, whether learning is explicit or implicit. In other words, we cannot even be sure that all explicit learning is conscious.

If learning is a change that takes place in an organism so as to enable it to solve new problems, then many parts of the organism learn. The immune system learns to recognize viruses, muscles learn by getting stronger. Our immune system and muscles hardly learn in a conscious way. Therefore, all learning in humans is not conscious. When it comes to human learning that strongly involves the nervous system, however, such as instrumental or classical conditioning, S&S claim that every aspect of learning is conscious. But what do people do when they learn to play tennis?

When a slightly changed combination of power, body position, and racket angle enables me to win more points, I have learned, but I do not know how, though I sometimes talk to myself when I play. Like most skills, playing tennis seems to depend on a blend of high-level, conscious guidance, and behavior of which I can become conscious only by self-observation.

Perhaps I am fooling myself. Maybe I just *think* I am not conscious of the way I learn to play. Well, there seems to be good experimental evidence, too, that people can learn without consciousness – evidence that S&S do not refer to. For example, people can learn in their sleep.

In their conclusion to section 2.6.2, S&S state that “implicit learning would be demonstrated if learning, as indexed by changes in instrumental behavior, occurred in the absence of awareness of the reinforcement contingencies.” Instrumental learning during sleep should be one of several cases to satisfy that criterion: Granda and Hammack (1961) reported finding avoidance learning in every sleep stage, including what has been defined as “deep sleep” (stage 4).

S&S claim that classical and instrumental conditioning has not been demonstrated in humans in the absence of consciousness. That the two types of conditioning are very well-known phenomena in most species of animals, including primates (Catania 1992) should mean either that animals have the kind of knowledge that S&S would call conscious, or that all learning in humans does not presuppose consciousness.

Consciousness is a key concept of S&S’s target article, yet they hardly discuss the meaning of this difficult word. In using the word “consciousness,” however, they consistently refer to “reportable knowledge” (e.g., sect. 2.3.2; sect. 5, para. 2). So (1) no animals or very few, have the kind of consciousness that S&S discuss, and (2) if people who cannot speak or understand speech can learn, their learning should be called “implicit.” Now, profoundly retarded people, with no vocabulary at all, have been shown time and again to learn through operant procedures (e.g., Barton et al. 1986; Mace & Knight 1986), as have normal preverbal children (Bentall et al. 1985; Lowe et al.

1983). That the behavioral effects of instrumental conditioning have been demonstrated in monkeys even after the spinal cord has been isolated from supraspinal influence by transection (Wolpaw & Carp 1990) should serve to cast further doubt upon the necessary role of consciousness in instrumental learning.

One of the two criteria that S&S use in deciding whether implicit learning has been demonstrated seems difficult to handle, and has obviously contributed to the two authors’ negative conclusion. In formulating the Sensitivity Criterion, they write (sect. 2.3.2) that “we must be able to show that our test of awareness is sensitive to all of the relevant conscious knowledge.” In section 2.7.2, S&S affirm that testing whether a subject is able to predict the next element of a sequence is a measure of awareness that satisfies the Sensitivity Criterion. It is difficult to see, however, how the Sensitivity Criterion can ever be satisfied. It is particularly troublesome to understand how being able to predict something proves that conscious knowledge was behind the prediction. A basic problem in designing expert systems stems from the fact that human experts tend to reach conclusions they “feel” to be correct, but are often unable to state the premises on which a conclusion is based (Dreyfus & Dreyfus 1986; Overskeid 1992). Johnson (1983) confronted a physician with the fact that he did not seem to do diagnosis the way he taught students to do it. And the doctor answered: “Oh, I know that, but you see I don’t know how I do diagnosis, and yet I need things to teach students. I create what I think of as plausible means for doing tasks and hope students will be able to convert them into effective ones.”

Many of Western culture’s traditional assumptions turn out to be wrong when scrutinized by science. However, the distinction between conscious and unconscious learning and knowledge, discussed by Socrates and Democritus around 400 B.C. (Overskeid, submitted), can still withstand attack and leave alive the intuitive mind.

Dissociating multiple memory systems: Don’t forsake the brain

Mark G. Packard

Department of Psychology, University of New Orleans, New Orleans, LA 70148; mgpps@uno.edu

In their interesting article Shanks & St. John (S&S) challenge the notion that a distinction between a conscious/explicit system, and an unconscious/implicit system is meaningful in taxonomizing human learning. Although conceding the existence of two learning systems that encode “different types of knowledge,” S&S propose that both systems acquire knowledge consciously. Having consciously attended to the target article (even considering it while driving, in which my learned behavior was certainly being *expressed* unconsciously), I am neither convinced that unconscious learning does not exist, nor that the terms “instances” and “rules” are particularly useful operating principles in distinguishing between multiple memory systems.

In failing to offer first their definition of consciousness, S&S add to the difficulty of assessing their view that unconscious learning does not exist. Nonetheless, the fundamental flaw in the analysis of S&S can be summed up in their own words: “In our evaluation of learning systems we have not considered at all the evidence from patients suffering from the classic anterograde amnesic syndrome.” In formulating theories about the operating principles that distinguish different types of memory, the literature on learning and memory in animals and human amnesics, and that on learning in normal humans should not be considered in isolation. There are inherent limitations in the interpretation of data from normal human subjects when arguing for the existence of multiple memory systems. The first is simply that without evidence of selective involvement of differ-

ent brain structures in mediating different types of memory, one could always argue that a single brain system exists to encode knowledge (even possibly different types of knowledge), and that the ultimate behavioral expression (which is used to *infer* multiple systems) of such knowledge depends on the content of the information processed. Second, even if it is assumed that multiple memory systems exist, studying memory in normal subjects in which one or more systems may be simultaneously active limits the ability to dissociate them. Findings of brain research suggest that multiple memory systems interact in the intact brain, and at present we have little understanding of the factors that determine the nature of that interaction. In normal subjects in whom both systems are active, there is no a priori reason to assume that a relatively more complex "conscious" learning system could not also acquire simpler learning tasks that may also be acquired by an "unconscious" system. Therefore, the data from normal subjects cited by S&S pertaining to Pavlovian conditioning and instrumental learning do not conclusively argue against the existence of unconscious learning. Rather, the data may be interpreted to suggest that a conscious form of learning is preferred by the brain when both systems are active and the experimental situation *permits* conscious learning. Thus, because of their exclusive consideration of data obtained from normal subjects, S&S cannot convincingly argue for the existence of multiple memory systems, a necessary first step in developing theories about the operating principles that distinguish different forms of memory.

However, a consideration of the extensive literature bearing on this question from research on animals and human amnesics leads inevitably to the conclusion that multiple memory systems do exist. Our own approach to the multiple systems hypothesis has involved comparing the effects of both lesion damage and posttraining stimulation of different brain structures in rodents on the acquisition of pairs of tasks with similar motivational, sensory, and motor characteristics, but with different mnemonic requirements. This approach has led to the observation of double dissociations in mnemonic function following manipulations of different brain areas, in particular, the hippocampal system and the caudate nucleus (Packard et al. 1989; Packard & McGaugh 1992; Packard & White 1991). Other findings support such an anatomical dissociation in monkeys, and recent evidence suggests that the separable roles of the hippocampal system and caudate nucleus in memory are also observable in humans with specific neurological disorders that predominantly compromise the hippocampus and caudate nucleus. Thus, a close correspondence is beginning to emerge in studies examining the anatomical bases of multiple memory systems across mammalian species. Considered together with S&S's suggestion that all forms of learning are conscious, they must be willing to concede animal consciousness (at least in mammalian species). If not, how do they maintain their position *and* explain the evidence indicating evolutionary conservation in mammals of brain structures involved in different forms of memory?

An additional problem with the proposal of S&S concerns the usefulness of "rules" as a meaningful term for defining a type of learning. A rule is a guide or a principle for governing action. In describing overt behavior, almost any task can be construed to involve rule learning, which limits the development of tasks designed to differentiate multiple memory systems. For example, the acquisition of a win-stay simultaneous visual discrimination, in which animals learn to approach a specific stimulus (e.g., a white light) might be described as learning the rule "approach the white light." Similarly, acquisition of a win-shift radial maze task, which requires animals to avoid visiting maze arms in which they have already obtained reward, might be described as learning the rule "do not revisit previously entered maze arms." However, despite the fact that it is relatively easy to describe rules that may underlie task acquisition, lesions to different brain areas produce a double dissociation in the acquisition of these two tasks (Packard et al. 1989). Thus, the proposition that

the use of rules is a unique form of learning does not adequately capture the essential difference in the two tasks. Instead, the development of theories that focus on the *content* of information in a particular task may prove more useful in defining the operating principles that distinguish multiple memory systems.

Of course, no single perspective is likely to provide a complete understanding of the complex issue concerning the nature of multiple memory systems. Rather, thoughtful integration of data obtained not only from normal human subjects (as aptly reviewed by S&S) but also from human amnesics and other mammalian species is undoubtedly the most promising approach.

What about unconscious processing during the test?

Pierre Perruchet^a and Jorge Gallego^b

^aLEAD, Faculté des sciences, Université de Bourgogne, 21000 Dijon, France; ^b Laboratoire de Physiologie, Université de Paris XI, 94276 Le Kremlin-Bicêtre Cedex, France; perruche@satia.u.bourgogne.fr

It may reasonably be anticipated that Shanks & St. John's (S&S's) will be a landmark paper in the research on implicit rule learning. Till now, unconscious rule induction seemed a well-established phenomenon. S&S make clear that most of its supporting evidence has recently been ruled out. As the authors illustrate, subjects' performance in implicit learning settings is essentially based upon the memory for specific events or fragments of study items, instead of being guided by some abstract form of knowledge. This renders *ipso facto* irrelevant all prior evidence for unconscious rule learning from the implicit learning literature. S&S's provocative plea passes the ball to those who advocate the reality of the phenomenon: it is now up to the latter to collect new experimental evidence for their claim. S&S provide clear criteria for them to meet.

It may be disputed, however, whether this line of reasoning (that we fully endorse) alone justifies S&S's strong claim against the existence of any sort of unconscious learning. First, a more cautious conclusion about the unconscious registration of specific events during the *training phase* may be in order. S&S are right to point out that specific knowledge of the study items has been shown to be available to consciousness in virtually any conventional implicit learning paradigm. But it is fair to acknowledge that only a limited number of paradigms have been explored; moreover, that these may have been far from optimal for revealing unconscious detection of the events under study. Indeed, because researchers basically intend to prevent subjects from rule searching, they take no care to discourage subjects, (and often even encourage them) to memorize their specific experiences. Under these conditions, the actual findings are not really surprising. Future work will have to make more extensive use of specially devised methods, such as dual-task conditions of learning, to give a real opportunity to subjects to acquire specific information without consciousness.

There is a second, much more fundamental reason, for some skepticism with regard to S&S's radical rejection of any intervention of unconscious processes during learning. This reason is rooted in the fact that their paper, mirroring the available literature, passes silently over one major component of the learning process. As evidenced in their Figure 1, S&S confine the occurrence of learning to time t_1 , that is, during the study episode. The only objective of the test tasks provided at t_2 is to reveal what has been acquired at t_1 . We believe that this widely held conception is undermined by a serious flaw. What needs to be explained in the end by a learning theory is the improved performance of subjects in their natural environment. This requires including, as part of learning, the processes occurring during what is artificially isolated, in most laboratory situations,

as the *test or transfer phase*. This aspect can of course, be disregarded without detrimental consequences insofar as one conceives that the study session triggers the elaboration of a ready-to-use set of general purpose rules, whose application in subsequent situations is somewhat trivial. In this perspective, the test phase of a learning session may be conveniently viewed as a window opening on the subjects' mind, which does not elicit specific adaptive mechanisms *per se*. However, if, as S&S themselves contend, people memorize piecemeal information in incidental learning conditions, then their performance during a subsequent test no longer proceeds from the straightforward application of ready-to-use knowledge. Performance can no longer be accounted for by a simple "retrieval" process, as claimed by S&S. One must now account for how a large amount of specific, fragmentary knowledge influences the processing of a subsequent situation. We still do not know most of the presumably complex mechanisms involved in this "on-line computation" (to borrow the terms of Smith [1989]).

The failure to consider the processes involved in the test phase of a learning session may have contrasting (although not mutually exclusive) consequences, as a function of the postulated nature of these processes. Because the knowledge base that subjects use during the test is available to conscious awareness (insofar as the current experimental evidence has demonstrated), it makes sense to assume that subjects engage in controlled forms of analytic reasoning to process this information. Disregarding this possibility may lead to an overestimation of the sophistication of the (presumably unconscious) processes engaged during the study phase. For example, in dealing with transfer to a new letter-set paradigm in artificial grammar settings, S&S concede that according to available data, some sort of abstract knowledge has been gained during the study episode, under incidental conditions. Transfer may instead be due to the abstracting operations that subjects deliberately engage in at test time to deal with the test items, on the basis of specific recollected information (see Perruchet & Pacteau 1991, for more detailed arguments).

That an amount of knowledge sufficient to account for improved performance is amenable to consciousness under direct, explicit requests, however, in no way implies that this knowledge is consciously retrieved and processed when subjects are faced with a test situation. Although direct experimental evidence is sparse, there are some hints that the use of knowledge to perform adaptively in a new situation may proceed without intention or concurrent awareness even though this knowledge can be made explicit if necessary (see Perruchet, *in press* for arguments). If this is correct, S&S's rejection of an independent unconscious learning system may be unwarranted, although this system would differ radically from the hypothetical unconscious rule-induction mechanism imagined until now. Rather than being aimed at the elaboration of an abstract knowledge base applicable to a large range of future, virtual situations, unconscious processes should be engaged when subjects use their prior experience to respond to the immediate adaptive pressures of their environment.

On the representational/computational properties of multiple memory systems

Russell A. Poldrack and Neal J. Cohen

Beckman Institute and Department of Psychology, University of Illinois, Urbana IL 61801; poldrac@s.psych.uiuc.edu and ncohen@s.psych.uiuc.edu

The literature on "implicit learning," as cogently reviewed by Shanks & St. John (S&S), revolves around the two issues of whether such learning occurs unconsciously, namely, without the subject's awareness, and whether it entails the abstraction of

rules rather than the storage of specific instances or fragments of the stimulus materials. We argue here that attempting to characterize or distinguish among different "learning systems" on the basis of such characteristics is less productive than exploring distinctions based on the representational and computational properties of different underlying memory systems. Also misguided, we believe, is the decision to exclude from serious consideration the remarkable pattern of sparing and loss of memory abilities in amnesia, which, as described by Eichenbaum et al. (this volume), has suggested the existence of multiple memory systems based specifically on distinct representational/computational properties.

1. Representational characteristics rather than qualitative properties of learning systems. The lack of progress on the question of awareness is striking: today's results may differ, but the arguments are basically the same ones that were debated without clear resolution in the literature on learning without awareness more than thirty years ago. Without a theory that explains the functional role of awareness in learning, the importance of the question for the study of learning systems is unclear.

Similar problems obtain for the debate over abstraction. As Barsalou (1990) has shown, arguments in favor of abstract versus exemplar representations of experience are basically undecidable from the empirical data; these schemes cannot be tested without also specifying the nature of processing of the representations. With the appropriate processing assumptions, the predictions of any abstractionist model can be mimicked by, and hence these models are empirically indistinguishable from, exemplar models, and vice versa.

Contrast this approach to the connectionist models of "implicit learning" by Cleeremans (1993b) and Dienes (1992), which have little to say about the question of awareness. Rather, they offer a computational approach that attempts to clarify the nature of the information processing required to perform a given task and the nature of the experience-mediated representations that support it.

Also contrast this approach to the multiple memory systems view offered by Cohen and Squire (1980), and more recently updated, elaborated, and extended by Cohen and Eichenbaum (1993). This view, summarized by Eichenbaum et al. (this volume), has distinguished between functionally separate procedural and declarative memory systems, which are differentiated by their representational properties and computational characteristics: declarative memory supports relational, compositional representations capable of mediating the flexible use of memory in novel contexts, whereas procedural memory supports representations that are fundamentally individual (non-relational) and inflexible, mediating the acquisition and expression of skilled performance in repeated contexts. Articulated in this way, the theory cannot only guide computational modeling efforts, but it can also be applied to studies of memory in animals and can be examined with an eye toward the neural basis of memory. Eichenbaum et al. have taken this a step further, attempting to articulate how the different anatomical components of the hippocampal system (damaged in certain forms of amnesia) give rise to the functional properties of declarative memory.

2. Amnesia and multiple memory systems. S&S largely ignore the ideas about multiple memory systems in amnesia and the preserved versus impaired learning and memory phenomena on which such theoretical interpretations are based, on the premise that these data do not speak to the issue of awareness in learning. Yet it has been well documented that amnesic patients can exhibit normal learning of (motor, perceptual, and cognitive) skills, and can do so in the absence of the ability to recollect the learning experiences, to remember explicitly the specific contents of their learning experiences, or to have insight into the nature of what was learned (see Cohen & Eichenbaum 1993). These preserved skill-learning phenomena are just as robust and reliable as the more widely studied preserved repetition

priming (i.e., preserved memory) effects, and just as independent of conscious recollection or explicit remembering. And it is precisely the discrepancy between the ability of amnesic patients to acquire and express skilled performance and their failure to show any conscious recollection or awareness of their learning or of the determinants of their increasingly skilled performances (even while generating those performances) that fueled multiple-memory-system accounts such as the procedural-declarative distinction.

In the procedural-declarative theory, the amnesic patients' deficits in conscious recollection of their learning experiences, explicit remembering of the contents of their learning experiences, and awareness of the nature of their enhanced skills are tied together in each, requiring the representational flexibility that only declarative memory can provide. In the absence of declarative memory following damage to the hippocampal system, there are only the inflexible representations supported by procedural memory, that, although fully capable of mediating the acquisition and expression of skilled performance in repeated contexts, are completely incapable of permitting conscious introspection.

Returning now to the implicit learning literature, we are forced to conclude that by failing to note the importance of the representational characteristics of the memory systems supporting particular performances, and by restricting its focus to more qualitative properties of putative learning systems, the field seems clearly limited in its ability to develop mature, computational theories of learning and memory.

Learning without awareness: What counts as an appropriate test of learning and of awareness

Sam S. Rakover

Department of Psychology, Haifa University, Haifa 31905, Israel; *rsps742@haifaumv.bitnet*

In a thorough and critical review of recent experiments in learning without awareness or implicit learning, Shanks & St. John (S&S) conclude that learning without awareness has not been established empirically. The main reason for this inference is that tests of awareness used in these experiments did not meet the Information Criterion and the Sensitivity Criterion.

History repeats itself, however. About forty years ago, a research program in verbal learning without awareness reached the conclusion that learning without awareness had been established empirically (e.g., Krasner 1958). However, improvements in the methods of assessing awareness (i.e., meeting the Sensitivity Criterion better) and use of the method of "correlated hypotheses" (i.e., use of the Information Criterion), led researchers to the opposite conclusion: that learning cannot occur without awareness (e.g., Brewer 1974; Dulany 1968; Spielberger & DeNike 1966).

Similarly, research on implicit learning first established empirically that learning does occur without awareness. However, a subsequent wave of experiments, taking into consideration the above two criteria, has shaken that conclusion and proposed that learning without awareness has not been found.

I propose that this historical repetition is not haphazard; rather, it indicates that there are profound problems and ambiguities with regard to the two questions: what constitutes a suitable test of awareness, and what counts as an appropriate task of learning without awareness?

Tests of awareness. S&S proposed that although prediction and recognition tasks generate good indices of awareness, because they address the Sensitivity Criterion by reproducing the context of learning, verbal reports of awareness are likely to be insensitive. I propose the opposite. The first thing to notice is

that recognition is not always more sensitive than recall. Under certain conditions subjects recall material that they cannot recognize (e.g., Flexser & Tulving 1978).

Prediction and recognition, as well as relearning and reaction time, are no more than performance indices of newly acquired information, and as such they provide us with only part of what has been learned. In contrast, verbal report, by its very nature, can furnish us with all the information that has been acquired (e.g., as in the case where a subject describes in full the rule constituting a task). There is a profound conceptual difference between self-reports, in which subjects access their cognitive states *directly*, and other performance tests, in which awareness is inferred *indirectly* from the subject's behavior. Note that in making this distinction I am not committing myself to the thesis that introspection is incorrigible and infallible. I am simply stressing that verbal reports furnish the only method by which one can communicate one's subjective private experience regarding the rule of a task.

If awareness of acquired knowledge is inferred from performance in indirect tests then one can legitimately ask whether or not subjects were aware of the knowledge, and whether or not they were aware of the effects of that knowledge on their performance. The answers to these questions cannot be based on the use of a new set of indirect tests, since these questions can be repeated with respect to the new set of tests as well. This infinite regress has to be stopped by using self-report. In other words, although it is legitimate to ask whether or not a certain performance was acquired consciously, a correct self-report of an acquired new rule does not arouse such a question.

Finally, I propose that the measurement of awareness by, for example, a prediction task, gives rise to the "assessment of awareness" paradox in implicit rule-learning research: consider Stadler's (1989) experiment, in which learning was measured by reaction time and awareness by a prediction task, where a subject was required to predict the position of the target stimulus. As Stadler mentions, the major difference between these two tasks is the response: reaction time versus choice. Apart from this, there is nothing conceptually special in the prediction task to qualify it as a procedure for assessing awareness. Thus, if my argument is correct, we merely have here two performance indices that provide us with certain inferential information. Hence, if the prediction task is a procedure for measuring awareness, reaction time must also be a procedure for measuring awareness, because both tasks are conceptually equivalent. And if this is so, then we do not need a special test for awareness, because reaction time would suffice. The upshot of this argument is that an indication of learning by any performance test is at the same time an indication of awareness.

Tests of learning without awareness. I have proposed that certain learning tasks may not be susceptible to introspective awareness and to awareness effects (Rakover 1993). According to the criterion for *task susceptibility to introspective awareness*, a given task is not susceptible to awareness if a subject, after having been provided with all the information concerning the rule of that task, apart from the specific content of the rule itself, is nevertheless unable to discover even a part of the content of the rule. According to the criterion for *task susceptibility to awareness effects*, a given task is not susceptible to awareness effects if, after having been provided with all the information concerning the rule constituting a task, as well as with the appropriate conditions for performing the task, the subject's performance of the task remains unaffected. These criteria determine (through their experimental utilization) whether a given task is susceptible to introspective awareness and awareness effects, independently of the research attempting to discover how explicitly/implicitly acquired knowledge transfers to subsequent tests. These two criteria generate four types of tasks: aware/aware, aware/unaware, unaware/aware, and unaware/unaware.

Hence, the question of learning without awareness should be divided into several subquestions. One interesting question is whether or not a given unsusceptible task can be learned at all. In order to answer, one must first establish the susceptibilities of a given task by performing an experiment, followed by a new experiment to find out whether or not the task can be learned. As an example, consider the experiment performed by Lewicki et al. (1987). They carried out a pilot study in which subjects were informed about the structure of the rules constituting the task, without being told what the specific content of these rules was. The subjects were also encouraged (by a high cash reward of \$100) to specify at least part of the rules. None of the subjects came even close to describing the correct rules. This task, therefore, can be classified as unaware/aware, because it is not susceptible to introspective awareness, although the subjects would probably have changed the performance in the task had they been informed of the specific rules of the task. Hence, Lewicki et al.'s experiment provides us with a positive answer to the above question. Note that if a given task is not susceptible, it cannot be used to decide empirically the question regarding learning without awareness. In nonsusceptible tasks, by definition, if learning does occur, it occurs unconsciously.

If these two points are correct, then Shanks & St. John's general conclusion about learning without awareness is called into question.

What manner of mind is this?

Arthur S. Reber and Bill Winter

Department of Psychology, Brooklyn College and the Graduate Center of CUNY, Brooklyn, NY 11210; artreber@bklyn.bitnet and wiwbc@kunym.cuny.edu

Our general reaction to Shanks & St. John's (S&S) target article is that it reinforces something we have long believed; in fields such as the cognitive sciences the truth lies not in "mere" experimentation but rather in the wise interpretation of the evidence as a whole. Not surprisingly, we put a rather different gloss on the evidence and have a different framework within which to view it than that presented in the target article. There are many points of disagreement; we shall focus on the most poignant ones.

Put starkly, S&S's conclusion that "human learning is almost invariably accompanied by conscious awareness" (sect. 5, last para.) must be wrong. It is as simple as that. If complex and abstract representations cannot be acquired largely independent of awareness of both the process and the products of acquisition then just how do children acquire natural language, how do we all come to be imbued with the social mores of our societies, how do we induce the complex arrays of categories and concepts with which we parse our environment? Indeed, the evidence of an implicit mode of acquisition of complex knowledge is to be found in virtually every dimension of human perceptual and cognitive activity. Any careful assessment of human action provides rich testimony to the manner in which behaviors, feelings, actions, and modes of reaction were learned without recourse to processes that played themselves out within the spotlight of a top-down, modulating consciousness.

Moreover, S&S's perspective entails a rather odd model of mind. If they are right that all learning in adult members of our species takes place within consciousness then we are going to need to rethink virtually everything we know about the phylo- and ontogenetic processes that lie at the core of evolutionary biology and, by extension, the psychological processes that rely upon them. To put it simply, because nonhuman species show clear evidence of the acquisition of rule-governed knowledge independently of conscious awareness (see any of the contributions to: Ristau 1991; Weiskrantz 1985), and because infants and very young children acquire complex, rule-governed knowledge

(Haith & McCarty 1990; Rovee-Collier 1990), S&S seem stuck with the odd conclusion that somehow something happens to members of our species somewhere before puberty such that all learning suddenly shifts from being synergistically modulated by both implicit and explicit systems to being entirely in the domain of the explicit. We note with interest that S&S neglect the literature on learning in animals and young children, an omission that we feel seriously weakens their argument. We suggest that a framework that is more coherent from an evolutionary biological point of view should recognize that consciousness is a late arrival on the evolutionary scene and was preceded by systems that enabled organisms to acquire complex knowledge about their environment independent of a modulating, self-reflexive consciousness (Reber 1993).

That aside, we certainly acknowledge that reasonable people can dispute over the message contained in the existing literature and, as noted above, wise interpretation is what is needed here. Implicit learning is a complex process and one that is not easy to capture in a controlled laboratory setting, but the conclusion that it cannot be dissociated from an explicit process is simply not supported by the existing literature. In what follows we would like to point briefly to various places where our reading of the literature differs from that presented in the target article.

1. In their discussion of Pavlovian and instrumental conditioning S&S leave out, as noted, the work on animals and infants. That normal adults may be aware of CS-UCS relationships does not mean that all conditioning takes place within conscious awareness. *Aplysia* is surely not conscious of the CS-UCS pairing (at least in the same sense and quality as adult awareness), yet this humble slug is even capable of differential conditioning (Carew et al. 1983). Infants as young as three months old learn to anticipate the location of a face that systematically moves from place to place (Haith & McCarty 1990) and to control aspects of their environment by simple motor responses (Rovee-Collier 1990). A quick look at the classic literature in kinesics (Birdwhistell 1970) shows that we acquire wonderfully complex muscle movement patterns that shape our face, gestures, and gait largely without awareness of the subtle interplay between the muscle groups involved and the manner in which their actions are shaped by the environment. Do S&S wish to suggest that the principles underlying Pavlovian and instrumental conditioning change with the number of neurons an organism has or the point in its development when the data are collected, or, that conditioning without awareness is possible but only in other species or in the earlier stages of ontogeny in ours? None of this is open to introspection; none of these behaviors are acquired with guidance from consciousness.

Part of the difficulty here is that in laboratory settings Pavlovian and simple instrumental conditioning involve relatively simple patterns of covariation that are easily noticed by nonimpaired adult humans; that is, these experiments tend to use stimulus materials whose structure is highly penetrable cognitively, leading to misrepresentation of the true roles that Pavlovian and instrumental conditioning play in "the real world." In any event, if the findings S&S cite to support the role of awareness in conditioning are robust (e.g., Grings et al. 1973), why can't clinicians eliminate phobias simply by making the client aware that the UCS is no longer going to occur?

2. S&S's juxtaposition of the encoding of instances and the abstraction of rules is an unfortunate one, because it sidesteps the deep issues of representation. Rules can, indeed must, be the result of abstractions across instances. The manner in which such abstractions are induced, moreover, can easily be thought of as beginning with the encoding of the fragmentary elements that made up the initial instances. As far back as the 1970s (Reber & Allen 1978; Reber & Lewis 1977) the case has been argued that implicit learning of artificial grammars results in an abstract representation, but the abstract representation itself is founded upon two- and three-letter chunks that are present in the learning array. As Whittlesea and Dorken (1993) have

shown, the nature of the memorial representation that subjects develop in implicit learning experiments is functionally determined by the conditions of learning and testing. The "either-or" nature of the issue that S&S present misrepresents matters. But the most interesting point, the one that S&S never come to grips with satisfactorily, is that the memorial representation must be of a form that permits for generalization or transfer to novel stimulus displays with different physical instantiations. Altmann et al. (in press), Brooks and Vokey (1991), Manza and Reber (1992b), Mathews et al. (1989), and Reber (1969) have all shown that knowledge acquired with stimuli instantiated in one physical form can be used to make decisions about stimuli instantiated in a different physical form, even to the point of using a different modality (Altmann et al., in press; Manza & Reber 1992b) – provided, of course, that the underlying structural regularities remain the same. Abstract memorial representations are required for such transfer.

3. S&S argue at length that in many of the studies on implicit learning the measures used to assess awareness were not sufficiently sensitive. In many cases, this is indeed correct and their point is well taken. However, two additional points need to be recognized. First, most of the measures that S&S suggest do not tap a naturally occurring awareness of knowledge held but rather specifically refocus the subjects' cognitive processes so that they "squeeze" metaknowledge out of them. Second, no one is arguing that tacit knowledge is totally opaque to consciousness; the point is that having conscious awareness of either the process or the products of implicit learning is not the *natural* state of affairs. The fact that labored and intensive quizzing and probing of subjects who have completed implicit learning experiments can occasionally get them to provide a verbal report of what they learned misses the point. The process of making implicit knowledge conscious is a post hoc, top-down, secondary process whereby individuals laboriously struggle to gain metaknowledge. It is, moreover, typical of the kinds of functions we expect from the operations of a human consciousness – top-down, modulating, probing attempts at self-reflection and the building of the metaknowledge base for self-awareness. Metaknowledge can be achieved, but it is a struggle. This pattern is shown clearly in the work of Mathews and his colleagues (Mathews et al. 1989; Stanley et al. 1989), where subjects' capacity to utilize implicit knowledge peaks rather quickly although their ability to explicate to others what they know takes much more experience and practice before it becomes efficient. Perhaps the best essays on this discoordination between the implicit and explicit were written years ago by the physician-chemist-philosopher Michael Polanyi (1966), who argued that the essential feature of complex knowledge systems was that we typically "know more than we can say."

4. S&S also maintain that there has been little or no work showing *qualitative* differences between implicit and explicit learning. This is true, but what has been done clearly points to a dissociation. Both Reber et al. (1991) and Aaronson and Scarborough (1977) found dramatically smaller individual differences between subjects when working in an implicit mode rather than an explicit one. Reber et al. (1992) also found that implicit learning, unlike explicit learning, was largely uncorrelated with IQ. Abrams and Reber (1988) found that psychotic and nonamnesic alcoholics differ dramatically in their capacity to function in implicit and explicit modes of acquisition, appearing normal when engaged in implicit learning but showing severely impaired explicit problem-solving ability. Rathus et al. (in press) report a similar pattern of dissociations when working with a group of highly anxious subjects. In all these cases, clear qualitative differences emerge between implicit and explicit modes of cognitive function.

5. Finally, S&S make scant reference to the amnesia literature. This is a serious error of omission; some of the most important findings relevant to the issue of implicit learning have emerged from that source. S&S suggest that amnesia studies

cannot be applied to the question of unconscious learning, because the possibility that subjects learn explicitly about stimulus relationships at the time of study cannot be ruled out. This comment reveals a basic misunderstanding about amnesic functioning and its importance to cognitive research. Although it is certainly true that amnesics are capable of explicit cognitive analysis at the time of study, the resulting information is, given any appreciable delay, lost to them by the time of testing. The extreme volatility of newly encountered explicit information in amnesia is one of the most robust and reliable findings in psychology, documented by literally scores of studies. The fact that amnesics nevertheless show normal implicit learning across delays indicates that the information necessary for successful performance can be acquired without the participation of explicit learning mechanisms.

This point is made a fortiori by Knowlton et al.'s (1992) artificial grammar learning study, where implicit knowledge is tapped with a standard well-formedness task and explicit knowledge by a recognition test for the studied materials. The amnesics showed impaired recognition for the same stimulus items that, under the implicit task instructions, produced behavior indistinguishable from normal controls. That amnesics show consistent and severe impairments on tests of recall or recognition for items presented just a few minutes before testing strongly suggests that they have not, in any meaningful way, explicitly learned the material in the first place. It also supports the notion of a neuroanatomical dissociation between the two modes of cognitive functioning.

Moreover, the differences in implicit and explicit learning in amnesics have been demonstrated experimentally in nonhuman mammals. Briefly, animals whose experimentally induced brain lesions are designed to replicate those typically found in human amnesia demonstrate analogous patterns of impaired and preserved functioning (Mahut & Moss 1984; Mishkin 1982; Zola-Morgan & Squire 1986). The degree of correspondence in the behavioral findings from the human and animal literatures has only recently become appreciated (Squire 1992); the evidence appears to weigh in favor of separate implicit and explicit cognitive processes that are based in distinct neurological systems. We fully expect that research on the physiological and neuropsychological domains will continue to provide major advances in the understanding of the brain structures and systems that differentially support explicit versus implicit cognitive functioning.

New evidence for unconscious sequence learning

Jonathan Reed and Peder Johnson

Department of Psychology, University of New Mexico, Albuquerque, NM 87131; pjohanson@bootes.unm.edu

Our comments on the Shanks & St. John (S&S) target article will concentrate primarily on the topic of implicit sequence learning. S&S have provided an excellent review of this literature; their evaluation of the relevant studies in regard to the Information and Sensitivity Criteria point out two of the most fundamental problems one encounters in attempting to interpret the results of these studies. S&S conclude that none of the reviewed studies meet both of these criteria, and, therefore, that no convincing evidence exists for implicit sequence learning. In addition, they describe a study (Shanks et al. 1994) showing that when these criteria were not met, results suggested that implicit sequence learning occurred but that when the criteria were met there was no evidence of implicit sequence learning. We have, however, recently reported evidence of implicit sequence learning (Reed & Johnson 1994), using procedures that we believe meet the Information and Sensitivity Criteria pro-

posed by S&S. For this reason, our commentary will provide a rather detailed description of our methods and findings, so that the reader can decide whether we have adequately addressed these issues.

Whenever a specific sequence is repeated, it introduces constraints on a variety of simple event frequencies. Learning any or all of these simple event frequencies may be sufficient to produce an indirect effect without the subjects actually learning the precise structure of the repeated sequence. This potential confound can only be eliminated if performance with a repeated sequence is compared to reaction times (RTs) associated with trials that maintain the same simple event frequencies as the repeated sequence.

In the Reed and Johnson (1994) study, subjects performed a serial reaction-time (SRT) task during which they were exposed to a repeating second order conditional (SOC) sequence (121342314324, where each digit refers to one of four horizontally displayed locations). The SOC structure refers to the fact that any pair of consecutive target locations predicts the exact location that follows, whereas a single target location does not. After training, subjects performed the SRT task for one block of trials with a new repeating SOC sequence (123413214243). This transfer sequence was selected because it matches exactly the training sequence in relation to simple event frequencies, but it is completely different from the training sequence in relation to SOC structure (e.g., 12 is followed by 1 in training and by 3 during transfer trials). With the introduction of the transfer trials, RTs slowed significantly. Because the training and transfer sequences were identical in relation to SOC structure, we felt that the indirect measure of learning met the Information Criterion, and we were hence justified in concluding that subjects had learned the SOC sequence structure (i.e., the index of learning was not confounded with the learning of other types of sequence structure).

Assuming that this observation constitutes indirect evidence of SOC learning, the question turns to whether the learning occurred without awareness. The Sensitivity Criterion proposed by S&S requires that the test for awareness be at least as sensitive as the indirect test in terms of its ability to detect relevant conscious information. This, they assert, can be accomplished by making the direct and indirect test as similar as possible in terms of retrieval cues and differing only in task instructions.

To determine whether learning of the SOC structure occurred without conscious awareness, Reed and Johnson (1994) used a recognition test of awareness. During the recognition test, subjects responded to a series of three target locations in the same manner as during training. After performing the SRT task for the three trials, subjects indicated whether or not the sequence of targets was part of the training sequence. Subjects performed this task 24 times; 12 times with training-consistent target triads, and 12 times with training-inconsistent triads. They performed at a chance level of accuracy, and we concluded that learning of the SOC structure had occurred without conscious awareness.

Because the results of Reed and Johnson (1994) and those of Shanks et al. (1994) are in opposition, a comparison between the two studies is warranted. Based on the brief description S&S provided of the Shanks et al. study, we are obviously limited in the analysis of the possible factors underlying the differences in findings. Three differences of potential relevance concern the type of indirect test and learning conditions used. Shanks et al. compared learning slopes of subjects trained with a repeating sequence to those of subjects who encountered a series of nonrepeating trials that were matched with repeating trials in relation to simple event frequencies. In contrast, Reed and Johnson used disruptions in RTs with the introduction of the new sequence as the indirect measure of learning. Given that we found that the magnitude for the disruption effects were not correlated with slope ($r = .0263$), it is

not clear that the two indices of learning are sensitive to the same factors.

A second difference between the two studies is that in the Reed and Johnson study subjects performed a secondary, tone-counting task (see Cohen et al. 1990) simultaneously with the SRT task, whereas subjects in the Shanks et al. study did not. Because subjects' incidental explicit learning of sequence structure has been found to be greater under single-task conditions than under dual-task conditions (Cohen et al. 1990), this difference in experimental procedure could account for the finding that subjects demonstrated explicit knowledge of the training sequence in the Shanks et al. study.

A third difference between the studies is the amount of sequence training that subjects received. In the Reed and Johnson study, subjects were exposed to the repeated sequence 136 times, whereas in the Shanks et al. study subjects were only exposed to 40 repetitions of the sequence during SRT training. Thus, subjects in the Shanks et al. experiment may not have been given enough opportunity to learn implicitly the sequence structure.

To summarize, we generally agree with the evaluation of the research on implicit sequence learning that S&S review, and, in particular, we concur with their notion that studies attempting to demonstrate unconscious learning need to meet the Information and Sensitivity Criteria they propose. However, because the study that we have described (Reed & Johnson 1994) appears to meet both of these criteria, we suggest that positive evidence for unconscious sequence learning does exist.

Learning strategies and situated knowledge

Antonio Rizzo^a and Oronzo Parlange^b

^aConsiglio Nazionale Ricerche-Istituto di Psicologia-Università di Siena, 53100 Siena; ^bDipartimento di Psicologia, Università di Padova, 35139 Padua, Italy; rizzo@kant.irmkant.rm.cnr.it and psico04@ipdunivx.unipd.it

Shanks & St. John (S&S) argue that there is little evidence for the existence of separate learning systems along the conscious/unconscious dimension. Instead, they claim, there is substantial evidence for more than one conscious learning strategy. Their argument first shows that there is no convincing evidence from experimental paradigms adopted in the implicit/explicit literature; then follows a review of results, mainly from concept-learning tasks, corroborating the thesis of two separate learning strategies: genuine rule learning and instance learning.

Here we would like to provide (1) further evidence from one of the same experimental paradigms adopted for claiming a dissociation between implicit and explicit learning (i.e., complex instrumental control tasks); (2) some implications of such evidence for clarifying the factors that are conducive to rule learning and instance learning.

(1) Studies on complex instrumental learning have often adopted different scenarios (e.g., city transportation, person interaction, sugar production, etc.) functioning according to an underlying rule. However, the kind of knowledge that subjects use under these circumstances has not received the attention it deserves.

Our approach (Cambiganu et al. 1993; Rizzo et al. 1993;) to the issue of learning to control the dynamic systems devised by Berry and Broadbent (1984) focused on the role of situated knowledge, namely, how knowledge transformation is constrained by content. According to the situated knowledge view (e.g., Light & Butterworth 1992), it is possible to predict that the cognitive difficulty in learning to control a system is not directly related to the underlying formal structure of the system but to the recruitment of schemata that can be properly adopted and modified. The dynamic systems used in the implicit/explicit

literature were not analyzed in this sense, though there have been several pieces of evidence for content effect (e.g., Berry & Broadbent 1984; Stanley et al. 1989). We have adopted four scenarios, the two scenarios originally devised by Berry and Broadbent (sugar production; person interaction) and two new ones (slope: positioning and keeping a trolley in a given state along a slope; tank: filling and keeping a given level in a tank). The rule governing the functioning of these four scenarios was the same. The two additional scenarios were designed with the aim of varying the congruence between the presumably available schemata of knowledge related to the scenario and the systems' behavior. The "slope" scenario tried to maximize the match whereas the "tank" scenario had the opposite purpose. The results showed different performances as a function of different scenarios despite the formal equivalence of the underlying rule governing the scenarios. Moreover, the knowledge content influenced the learning processes. Two results are of particular interest here.

First, in the tank scenario, subjects did not exhibit any improvement in performance. This is strong evidence against the thesis of two independent learning systems along the conscious/unconscious dimension. The constraint imposed by knowledge content on learning could in itself hardly be compatible with the occurrence of implicit learning, because implicit learning should involve the acquisition of abstract rules from S-R pairs. However, the lack of performance improvement exhibited by subjects in the tank scenario is not compatible with the thesis that implicit learning is driving the subjects' behavior, because the actual information available to the subjects was the same for three scenarios (slope, tank, sugar) and only superficially different for "person." If the existence of an implicit learning system is maintained, it would be difficult to explain why the tank scenario alone blocked not only an implicit induction of the underlying rule but even an implicit abstraction of any useful input-output regularity. On the contrary, these results are compatible with a situated knowledge approach, which has shown, in reasoning tasks, that if the content triggers knowledge that overrides the stated premises, the application of the inference rule may be blocked (subjects first need to decontextualize information). We accordingly found a performance improvement for the tank scenario when the subjects were provided with verbal instructions (Cambiganu et al. 1993).

Second, we obtained a high correlation between performance and relative verbalization scores in the four scenarios. Also, in the free text reports, the explanations advanced by the subjects were in keeping with the characteristics of the scenarios, even if not with the underlying rule. They could provide explanations such as, "Since the trolley was very high and in a steep position I had to apply a stronger force" in the slope scenario; or, "Clegg is spoiled and touchy so to keep him friendly I have to follow his moods" in the person scenario. Our conclusion was that situated knowledge explains the otherwise bizarre behavior of the system under control. The subjects' behavior (performance and verbalization) seems to vary according to the suggested mental model (see also Sanderson 1989).

(2) Considering the situated knowledge view and our results, we suggest that, at least for complex instrumental control tasks, rule learning is predominant with respect to instance learning. Instance learning would mainly take place when no clear content addressable schemata are available. Most of the evidence obtained in category learning also supports this view (i.e., instance learning is induced by unusual content, speeded or secondary tasks). Finally, if the context does not activate well-fitted schemata, a combination of micro-rules and instances is the knowledge pattern more likely to be involved in learning.

Criteria for implicit learning: Deemphasize conscious access, emphasize amnesia

Carol Augart Seger

Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90024-1563; seger@cognet.ucla.edu

Shanks & St. John (S&S) point out that implicit learning experiments often cannot meet strict criteria for the learning to be considered unconscious. This problem may be inherent in the tasks that illustrate implicit learning; they may involve evaluation of knowledge that is on the fringe of consciousness (Mangun 1993), and the supposedly explicit tasks they are compared to may also be sensitive to this fringe knowledge. Even if we do accept that consciousness is a difficult criterion, it does not make sense to drop it and simply lump all the tasks together as explicit instance learning tasks, as S&S suggest, because the tasks share other commonalities. Most important, they are preserved in amnesia, and therefore depend on brain structures other than the hippocampal system on which explicit memory is dependent.

Should we follow S&S's suggestion and not use consciousness as a criterion for differentiating learning processes? Consciousness certainly is problematic, for two related reasons. First, it is difficult to devise a measure of implicit learning that taps only unconscious knowledge. The different tasks used in implicit learning research use different dependent measures of learning, which I call response modalities (Seger 1994); at least two of these response modalities require subjects to make conscious decisions. In the conceptual fluency response modality (e.g., grammaticality judgments in artificial grammar studies) subjects make decisions concerning the rule-governedness of stimuli by a process of evaluating their subjective feelings of knowing, fluency, and familiarity. In prediction and control (e.g., the dynamic systems research of Berry & Broadbent, 1984; 1988, classified by S&S as complex instrumental learning), subjects demonstrate learning by their increased ability to predict or control the value of certain variables by manipulating values of other variables. These measures appear to tap into what Mangun (1993), extending the idea of William James, called the fringe of consciousness: a field of relatively unarticulated, vague experience, neither fully accessible to consciousness, nor fully separate. It is not surprising that if tested in the right ways, subjects can be led to make responses indicating that they have more consciously available knowledge than they show in free report measures.

The third response modality, efficiency, in which subjects show increased speed or accuracy in processing stimuli, is less susceptible to arguments that it reflects conscious knowledge, because subjects are not aware that they are making responses that reflect learning and do not have to make judgments and decisions based on fluency or feeling of knowing. It is in regard to sequence-learning studies that use this response modality that S&S's claims that implicit learning is not unconscious are the weakest. They merely note that the measure of awareness in contingent response tasks (complex sequence-learning tasks of the sort performed by Lewicki and colleagues: Lewicki et al. 1987; Lewicki et al. 1988) does not meet the Sensitivity Criterion, and that the selection of the random stimuli on which performance was compared to that of the serially structured stimuli in the serial reaction-time task was inappropriate (S&S report that in Shanks et al.'s [1994] study using pseudorandom stimuli that followed the first order frequencies of the sequence, subjects did not show learning when their performance on sequence and pseudorandom blocks was compared; however, other recent studies have found learning in using comparison stimuli with first order frequencies identical to those in the stimulus pattern; Frensch et al. 1994; Reed & Johnson 1994.)

The second reason that consciousness as a criterion for im-

Implicit learning is problematic in that it is difficult to come up with a measure of explicit knowledge that is independent of implicit learning. S&S argue that prediction and recognition tests display subject's explicit knowledge of the pattern; however, in many other contexts these sorts of tasks are used as demonstrations of implicit learning. Within the conceptual fluency response modality there is evidence that subjects can make many different kinds of judgments based on implicitly learned knowledge (such as judgments of liking of a stimuli; Gordon & Holyoak 1983). Recognition judgments can be made on the basis of implicit conceptual fluency also, if subjects are willing to make these judgments on the basis of feelings of familiarity or fluency rather than explicit episodic recall (Hirst et al. 1988; Jacoby & Dallas 1981; Johnson et al. 1991; see Seger, 1994, for a more extensive treatment). Prediction tasks are used in contingent response tasks as a measure of implicit learning (Kushner et al. 1991) and are similar to controlling variables in dynamic systems tasks.

Given these problems, perhaps it is wisest not to make consciousness a primary criterion for implicit learning. Although it is true that the different implicit learning tasks on the whole lead to remarkably little verbalizable conscious knowledge, certainly much less than learning by explicit conscious hypothesis testing, it appears that (except for experiments that use the efficiency response modality) learning is at least somewhat penetrable by conscious processes. However, deemphasizing the definitional role of consciousness does not mean that the tasks falling in the field known as implicit learning do not form a type of learning that is meaningfully different from explicit learning. The empirical evidence for dissociations between implicit and explicit tasks is stronger than that for which S&S give it credit. It is true that there is a paucity of cases in which each type of learning is affected in opposite ways by an experimental manipulation, perhaps because implicit learning is a very robust phenomenon with which it is difficult to interfere. However, a number of situations have been identified in which variables affect explicit learning without affecting implicit learning (Abrams & Reber 1988; Howard & Howard 1989; Masters 1992; Reber et al. 1980; Reber et al. 1991).

S&S suggest that implicit learning be considered merely explicit instance learning. This proposal ignores research indicating that implicit learning can sometimes be remarkably abstract (as in artificial grammar and contingent response task studies showing transfer between different letter sets and different sensory modalities; Lewicki et al. 1992; Reber 1969; Manza & Reber 1992a). In addition, implicit learning tasks share other common features indicating that they form a meaningful cognitive category: low individual differences (Reber 1992a) and, most important, preservation in cases of amnesia. Artificial grammars, other visual patterns, sequences, covariations, and dynamic systems, despite the differences of response modality and types of stimuli used, are learned normally by amnesic patients (Knowlton et al. 1992; Knowlton & Squire 1992; Musen & Squire 1993; Nissen et al. 1989; Squire & Frambach 1990). Preservation in amnesia strongly suggests that implicit learning tasks are dependent on neural systems other than the hippocampal-diencephalic episodic memory system that is compromised in amnesia (Squire 1992). S&S argue that the results of studies with amnesics are tangential, because amnesics may well be aware of learning. However, this is missing an important point: preservation under amnesia itself is a good criterion for classifying learning as implicit.

Dissociable learning and memory systems of the brain

Larry R. Squire,^a Stephan Hamann,^b and Barbara Knowlton^c

Veterans Affairs Medical Center, Departments of Psychiatry and Neurosciences; ^{b,c}Department of Psychiatry, University of California at San Diego, San Diego, CA 92161; ^alsquire@ucsd.edu

In their target article, Shanks & St. John (S&S) undertake a heroic review of human learning, drawing from many domains of evidence based on normal subjects, in order to ask whether there exist independent and dissociable human learning systems. The review proceeds by considering two dimensions on which separable learning systems have been purported to differ: awareness versus unawareness of what is being learned, and learning by acquiring rules versus learning by memorizing instances. At the end of their review, the authors conclude that the first distinction (awareness vs. unawareness) has not been established as a useful way to classify learning systems, whereas the second distinction (rule learning vs. instance memorization) provides a useful basis for classification.

There are two difficulties. First, in separating the topics of learning and memory, and focusing on learning, the target article has a peculiar emphasis and largely neglects the interesting distinction between explicit (declarative) memory and implicit (nondeclarative) memory. Second, despite the admirable scope of the review, had the authors cast a still wider net for relevant data across the whole discipline of cognitive neuroscience, including research with amnesic patients, studies of experimental animals, and consideration of the anatomy and function of brain systems, they might have reached different conclusions. The perspective presented here is that implicit learning leads to implicit memory, and evidence for dissociations within memory are relevant to the nature of learning as well.

To begin, we suggest that the distinction between explicit and implicit (declarative and nondeclarative) forms of learning and memory rests on several kinds of evidence. The distinction does not stand or fall according to the status of aware versus unaware learning. For example, explicit and implicit forms of learning and memory depend on the integrity of specific brain systems. Note also that implicit learning is not the single thing depicted in the target article, but a collection of separable abilities (skill and habit learning, priming, emotional conditioning, skeletal muscle conditioning, and nonassociative forms of learning such as habituation and sensitization), each dependent on particular brain systems (the neostriatum, the neocortex, the amygdaloid complex, the cerebellum, and specific reflex pathways). Many of the details of how these systems support, and otherwise contribute to, explicit and implicit memory remain unknown, but the biological facts are important for how the relationship between explicit and implicit memory should be understood.

Another differentiating feature of these forms of memory is that explicit memory is flexible and available to multiple response systems, whereas implicit memory tends to be less flexible and limited by the context in which the original learning occurred. These conclusions come from important animal studies, which show that animals with lesions of the hippocampal system who succeed at learning a task learn qualitatively different information from normal animals (Eichenbaum et al. 1988; Saunders & Weiskrantz 1989). Finally, a growing body of work with experimental animals has invoked the same or a similar distinction (e.g., declarative vs. procedural) to understand the neural and functional organization of memory. Such a distinction was needed to understand the data, without any recourse to issues of awareness and unawareness (Eichenbaum et al. 1992; Mishkin et al. 1984; Packard et al. 1989; Sutherland & Rudy 1989; Zola-Morgan & Squire 1984).

Quite apart from the work with animals, awareness and

unawareness are important properties of explicit and implicit learning in humans. Consider the domain of perceptual priming, namely, the improved ability to detect or identify recently presented stimuli (Schacter et al. 1993; Tulving & Schacter 1990). Priming is intact in amnesia. Subjects are conscious and aware of the stimulus presentations, to be sure, but they are not aware that their perceptual apparatus is being altered so that later they will better perceive the stimuli. Indeed, the concept of priming is absent from their mental lexicons. S&S state in the target article, "to infer unconscious learning from implicit retrieval, the subject must be unaware of the relevant relationship that occurred in the study episode, in addition to being unaware of the episode itself." In the case of priming, normal subjects are unaware in the first sense, and amnesic patients are unaware in both senses. How can subjects be aware of the "relevant relationship" if subjects have no concept that a relationship exists between stimulus exposure and later perceptual speed? Priming is manifest through performance (at which point some subjects may notice that some items are easier to see), but the memory content (increased fluency) is not otherwise accessible to conscious recollection. During initial presentation of stimuli, subjects are aware only of perceiving the items, not of learning to perceive them more efficiently.

Similarly, it is difficult to imagine what normal or amnesic subjects are aware of when they improve their ability to read mirror-reversed text (Cohen & Squire 1980). Perhaps subjects become aware that they are improving a little, but what is it that they know about how this is happening? Such facilitation is probably specific to the instances that have been practiced, not a reflection of rule learning (Musen et al. 1990); but the distinction between instance memorization and rule learning is orthogonal to the distinction between explicit and implicit learning.

Early exposure to speech leads to stored representations that reflect the properties of the native language and that guide speech perception and production. By the age of 12 months, changes have occurred that are unique to each language environment. For example, experience during the first year determines which phonetic boundaries can be heard and which cannot. These changes are longlasting, resistant to modification, and nonconscious. They are expressed through performance because they reflect changes in the perceptual apparatus. They are not available as recollections and do not have the quality of "pastness" that is an attribute of explicit forms of memory.

Consider also the case of the classically conditioned eyeblink response (Thompson 1990). Regions of cerebellum and related circuitry in the brainstem are essential for the acquisition, storage, and expression of the conditioned response. In rabbits and cats, transection of the brainstem just rostral to the pons spares both the acquisition and the retention of the conditioned response (Mauk & Thompson 1987; Norman et al. 1977). This transection eliminates input from the forebrain to the cerebellum and brainstem. If the entire forebrain (including the neocortex, striatum, basal forebrain, hippocampus, amygdala, and thalamus) is not needed for conditioning, then it is difficult to imagine how conscious knowledge of CS-US associations plays a role in the learning. It seems more likely that the development of awareness about the contingencies in human conditioning experiments is epiphenomenal and not necessary for the conditioning to occur.

Finally, consider the case of habituation, sensitization, and classical conditioning in the invertebrate *Aplysia*, where the neurobiology and behavior have been studied in considerable detail (Hawkins et al. 1993). These examples of learning surely proceed without awareness of any cumulating memory content. In the course of vertebrate evolution, we suppose that phylogenetically early kinds of learning and memory remained, and a new kind of learning system developed, one with new properties that included the capacity to acquire information that is available to conscious recollection (see also, Sherry & Schacter 1987).

Learning and memory systems are thought to operate in parallel, just as the sensory systems do. In the case of vision, one processing stream proceeds ventrally, culminates in area TE in the inferotemporal region, and is concerned with object perception. Another processing stream proceeds dorsally, culminates in area PG in the parietal lobe, and is concerned with visuospatial information (Ungerleider & Mishkin 1982; Goodale 1993). Perception of an object in space involves concurrent activity in areas TE and PG, as well as in other areas. Consider what happens to this neural account of perception when one applies the approach taken in the target article, namely, that the term "implicit" properly refers to a particular kind of retrieval from a common knowledge base. This view is akin to supposing, in the case of perception, that some area (or system) other than areas TE and PG is actually more fundamental to visual perception, and areas TE and PG are simply retrieval routes by which different aspects of a percept (object identity or visuospatial information) are expressed. The idea that there is only one knowledge base and that explicit and implicit memory phenomena reflect dissociations of retrieval ignores biology. The learning and memory systems of the brain are separable and parallel. One system supports the acquisition, storage, and retrieval of conscious knowledge. Other systems support nonconscious learning and its various products: the dispositions, preferences, skills, and habits that are expressible through performance, but accessible to awareness only to the extent that subjects can (epiphenomenally) observe and monitor what they do.

Whither learning, whither memory?

Michael A. Stadler and Peter A. Frensch

Department of Psychology, University of Missouri-Columbia, Columbia, MO 65211; psymike@mizzoul.missouri.edu; and psypeter@mizzoul.missouri.edu

Often, the devil is in the details, but on the details we generally agree with Shanks & St. John (S&S). Instead, the devil may be in one of S&S's main points, namely, the assumption that "in order for us to infer unconscious learning from implicit retrieval, the subject must be unaware of the relevant relationship that occurred in the study episode, in addition to being unaware of the episode itself" (sect. 2.3.1) and that "such inferences will have to be accepted if unconscious learning is to be established" (sect. 2.3, para. 5). S&S's definition of learning relies on the nature of learning's result rather than on the learning process itself. We believe it is necessary to distinguish between the learning process, which may be implicit or explicit, and the result of that process, memory, which may be accessed implicitly or explicitly.

Learning is implicit when it is unaffected by intention (Reber 1989a). Likewise, memory is implicit when retrieval is unintentional (Graf & Komatsu, in press).¹ In our view, the implicit or explicit nature of learning and memory are functionally independent; therefore, any combination of learning and memory orientations is logically possible. Both learning and memory retrieval may be implicit, or both may be explicit. If explicit memory is poor, the former condition meets S&S's criteria for implicit learning; the latter is typically called remembering (Tulving 1985) and is exemplified by students studying for and taking exams.

There are also circumstances, however, where learning is explicit and memory is implicit. Here, the relevant stimuli or stimulus relations are explicitly encoded but they later affect behavior implicitly, with no intentional memory retrieval. During the development of automaticity for example, performance is ponderous, as subjects intentionally begin learning a task, but it becomes efficient with practice. Performance may reach the point where previous experiences have their effects without

explicit retrieval of those memories (cf., Logan 1988). When a subject's explicit memory for those experiences is poor, this case would be called implicit learning under S&S's definition, even though learning was intentional.

It is also possible for learning to be implicit with an explicit, intentionally retrievable memory as a result. For example, in the serial reaction-time studies of implicit learning, learning may proceed implicitly, yet subjects may over time become more and more aware of, and able to remember explicitly, some of what was learned. Even though the critical information was initially encoded implicitly, S&S's criteria would indicate that learning was explicit.

From the above examples, it should be clear that the learning process and the subsequent retrieval of learned information must be distinguished. Although learning and memory are very closely related, they are logically independent constructs and must be treated as such, else we risk calling explicit learning implicit, and implicit learning explicit.

In line with our definition of implicit learning, we prefer a research strategy based on comparisons of learning under implicit and explicit orientations (i.e., instructions). Ultimately, the best evidence for the distinction between implicit and explicit learning would be to find qualitatively different results under the implicit and explicit orientations. Whether subjects are or are not aware that they have learned something is not the most compelling theoretical issue. More important is the question of whether learning proceeds differently under implicit and explicit orientations and, if so, how (cf., Cheesman & Merikle 1986).

As a postscript, we note that S&S have suggested (sect. 2.7.2, para. 17) that Stadler's (1989) test of subjects' awareness, on which the subjects performed at chance, might have been flawed. The subjects were trained extensively on 24, seven-trial patterns in a visual search task where the patterns determined target location. In the test, they were asked to predict the location of the target on the seventh trial after responding to the first six. Patterns were ordered randomly and each appeared twice during the test. S&S argue that during the test, interference might have caused the subjects to forget what had been learned. They suggest testing this hypothesis by seeing whether subjects performed better early in the test than late. As it turns out, each subject's correct responses were distributed evenly over the course of the test, with no evidence of a preponderance of correct responses in the early trials. No subject was correct more than twice in the first eight trials (chance was two correct), and none was correct more than five times in the first 16 (chance was four). There is, therefore, no evidence to support Shanks & St. John's forgetting hypothesis.

ACKNOWLEDGMENT

Order of authorship was determined by a coin toss.

NOTE

1. Graf and Komatsu point out that awareness need not be preceded by intention. That is, people might be aware of unintentionally retrieved memories. We will not consider the issue in this limited space, because it does not alter our main point. For the sake of argument, we shall treat unintentionally retrieved memories as if people are unaware of them.

Is awareness necessary for operant conditioning?

Frode Svartdal

Department of Psychology, University of Tromsø, 9037 Tromsø, Norway;
frodes@psyk.uit.no

This commentary addresses the main question raised by the authors ("Can learning occur without awareness?") by focusing on the awareness issue in human operant or instrumental

conditioning. Shanks & St. John (S&S) examine only a very limited set of studies on instrumental conditioning, and some findings taken as evidence against unconscious learning (sect. 2.6.1) are only tangential to the awareness issue. For example, the general finding that performance and contingency verbalizations are often associated (e.g., Catania et al. 1989) provides no answer to the question of whether human operant conditioning can occur without contingency awareness; neither do demonstrations that performance and judgment are related in simple operant tasks (Shanks & Dickinson 1991).

Proper assessment of the awareness issue requires experiments capable of testing the possibility of nonconscious contingency processing. As indicated in the target article, such experiments are problematic, for several reasons. One not discussed by S&S is the strong effect explicit rules have on performance and the interaction of such rules with implicit contingencies. Many experimental designs have functioned much like guessing games in that ambiguities of the experimental situation encourage active hypothesis generation, thus promoting verbal control of performance. For example, Hefferline et al. (1959) arranged operant contingencies requiring very small muscular movements outside voluntary control, but artifacts from gross voluntary responses associated with subject hypotheses excluded a convincing conclusion about nonconscious contingency control. Another approach attempted to mimic contingency-shaping (i.e., reduce verbal control) by minimizing experimental instructions (Matthews et al. 1977). However, placing a subject in an experimental situation without explicit instructions may enhance rather than reduce verbal control through the development of self-instructions (e.g., Lowe 1979). The verbal conditioning literature (see review by Brewer 1974) indicates similar problems: exposing a subject to an ambiguous situation may encourage the formulation of hypotheses about the experiment, thus promoting explicit verbal control that may interact, often favorably (Dulany 1961), with the hidden contingency.

A more refined solution would be to provide explicit instructions for one task, but expose attributes of responding under that task to a concurrent, implicit operant contingency. Instructions would then control one dimension of action, and the implicit contingency would, if effective, control another dimension. Through making the instructed task difficult, one might be reasonably certain that subjects will allocate extensive attentional resources to that task; and arranging the implicit contingency for a different (but not incompatible) dimension of behavior, renders any shift in attention from the instructed dimension to the noninstructed one unlikely. Note that this strategy is somewhat different from comparable studies of implicit learning in process control (e.g., Hayes & Broadbent 1988) and sequence learning (e.g., Lewicki et al. 1987), since the implicit and explicit (instructed) contingencies in such tasks are for the same response dimension; in the design outlined, instructions and reinforcement contingencies operate on different dimensions.

In line with this reasoning, Svartdal (1992) exposed subjects to rather difficult tasks and provided feedback for "correct" answers. Instructions required one left or right button press per trial in response to a discrimination task, and feedback was delivered following correct answers. Unknown to subjects, at a predefined point in the experimental session the feedback criterion changed from position (left, right) to force: different groups of subjects were now reinforced for pressing with increased or decreased force compared to individually calculated baseline forces. Results from several experiments using this dual-task design demonstrate behavioral sensitivity to the implicit operant contingencies in that pressing force increased or decreased as required (Svartdal 1992; Svartdal & Mortensen 1993). No indication of verbal control of the relevant force dimension was detected in subsequent questionnaires and interviews. One can be reasonably sure, therefore, that the observed force changes were generated by mechanisms outside conscious control.

Interestingly, making the primary instructed discrimination task more difficult did not seem to have a negative effect or adaptation to the implicit contingency; on the contrary, adaptation tended to be reliable in the high-difficulty condition but unreliable in the easy condition (Svartdal 1992). Thus, because only marginal attentional resources could be allocated to the hidden contingency under the difficult instructed task, this result lends further support to the hypothesis that the force changes were generated by nonconscious mechanisms.

As noted by S&S (sect. 2.6.1), the Information Criterion may be difficult to satisfy in operant tasks, one reason being that free operant experiments allow little control over the subject's behavior. Note, however, that in these experiments response rate was under instructional control: all subjects emitted a fixed number of responses, one per trial. A second threat to the Information Criterion is the possibility of correlated hypotheses (Dulany 1961). In their discussion of another set of experiments (Svartdal 1989; 1991), S&S suggest that systematic tempo changes could be induced by specific response strategies (e.g., subjects resting their hand in certain positions). This is a valid threat to the Information Criterion only insofar as these particular hand positions were under conscious control. There are no indications in our data of the use of such conscious strategies. For example, following the experimental session, we asked subjects to formulate advice to another hypothetical person on how to master the task. Answers were typically very general (e.g., "stay attentive") or irrelevant to the contingency.

The Sensitivity Criterion, if taken literally, may be harder to satisfy. Tests of awareness during conditioning are likely to interact with task performance, sometimes to make the implicit contingency more salient to subjects. For this reason, postexperimental interviews and questionnaires may be preferable in situations where the salience of the contingency must remain low. As argued by S&S, this strategy may render awareness tests less sensitive. However, the sensitivity of an awareness test cannot be evaluated out of context; in general, it should be evaluated relative to performance. If behavioral sensitivity to the contingency is high, contingency awareness can be suspected; in contrast, if behavioral sensitivity is low, awareness of the contingency should be low. In our studies we generally find relatively low, though significant, effects on the performance measure. Low behavioral sensitivity to an implicit contingency may occur for two reasons, however; because that contingency is difficult to process (as, for example, in some serial-learning tasks), or because awareness is not involved in contingency processing. Force and tempo contingencies of the sort we have arranged are easily processed if detected. Relatively low behavioral sensitivity to such contingencies may therefore be taken as evidence that they are not consciously detected, that is, that learning occurs with no or low contingency awareness.

The idea of nonconscious processing of operant and other learning contingencies may be controversial, but the view, favored by Shanks & St. John, that assumes conscious processing of all action-relevant information, is also controversial and fails to capture important and intriguing aspects of human behavioral regulation. Their view is falsified by one good demonstration of nonconscious learning. The task then becomes that of assessing the relative contributions of conscious and nonconscious processes in action control (cf., Shanks & Dickinson 1991, p. 360).

Are infants human?

H. S. Terrace

Department of Psychology, Columbia University, New York, NY 10027;
terrace@columbia.edu

Does learning occur without awareness? Shanks & St. John (S&S) claim it does not. In their selective review of the litera-

ture, they find "little actual support for unconscious rule induction . . . or for the unconscious learning of any other type of information." S&S qualify their claim by limiting their focus to human learning. That dubious restriction does not, however, salvage a futile exercise in armchair psychology, in which introspection is given special status as a source of hypotheses about learning.

S&S's view that the contents of consciousness are crucial to our understanding of learning or memory is shared by many cognitive psychologists. That makes it all the more important to show why the contents of consciousness cannot serve as an adequate explanation of any psychological phenomenon. That is not to deny that consciousness is a valid subject of inquiry. It is; and because it is such a difficult topic, it will continue to attract a wide variety of investigators for the foreseeable future. At issue is whether anything is gained by regarding consciousness as an independent, as opposed to a dependent, variable.

From the outset S&S take the position that learning with consciousness (or awareness, a term they use interchangeably with consciousness) is the norm. They deny the existence of "a quite separate system that operates independently of awareness," on the grounds that tests for consciousness are not "sufficiently sensitive to detect exhaustively *all* conscious information" (italics added).

S&S insist that two criteria must be satisfied to show that learning occurs without awareness. An Information Criterion assesses what the subject can report about the learning task. Before concluding that subjects are unaware of the information relevant to the learning task (*I*), the experimenter must show that the information about which the subject is queried "is indeed the information responsible for performance changes." The subject may, for example, have learned *I**, a derivative of *I*, instead of *I*. If the subject is aware of *I**, but not *I*, an experimenter who only tests for knowledge of *I* would conclude erroneously that subjects did not acquire any information about the nature of the learning task (*I*). A Sensitivity Criterion must also be satisfied to show that the test for information is sensitive to all relevant sources of conscious knowledge.

Much of S&S's article is devoted to showing that studies claiming learning to have occurred without awareness are unable to satisfy at least one of their criteria. S&S's criticisms are unconvincing, because there is no principled way to constrain the number of ways in which *I** or some undetected aspect of consciousness can manifest itself. Space limitations prevent me from commenting on other aspects of S&S's cavalier dismissal of such studies. Instead, I will address what I regard as a more basic issue that follows from their preoccupation with adult human subjects who can introspect. How do human infants learn?

Consider an experiment of the type performed by Rovee-Collier and her colleagues on infants as young as 2-3 months (Rovee-Collier 1990; Rovee-Collier et al. 1985). The infant lies comfortably in her crib and can gaze at a mobile suspended above her head. With little effort, she learns to make the mobile turn by pulling a string with her foot. By altering either the order of the items hanging from the mobile or the context in which the infant is tested, Rovee-Collier provides extensive evidence that infants develop expectancies regarding the appearance of particular items and, in general, that they learn and remember many specific features of the mobile and its surround. In this and other studies of infant learning there is no basis for applying either of S&S's criteria for conscious learning. Whether infants are aware of their expectations is entirely a matter of speculation. Since consciousness in infants cannot be subjected to empirical study, there is no basis for assuming that infants are conscious of what they learn.

The literature on animal learning, of which S&S make no mention, raises problems similar to those encountered in studies of learning in infants. S&S go to considerable lengths to argue that awareness is necessary for the occurrence of Pavlovian

and operant conditioning in human subjects. Curiously, they say nothing about the implications of that view for theories of animal conditioning. Are we to assume that different sets of principles are needed to account for conditioning in animals and in human subjects, and that the difference stems from the fact that, in some instances, adult humans can describe the contingencies used by the experimenter?

Instead of framing hypotheses of learning around the introspective reports of adult human subjects, I recommend the generally accepted scientific strategy of working from the simple to the complex. A good place to start is to recognize the compelling evidence of learning without consciousness in infants and animals. Another simplification follows from the fact that animals and infants lack language. During the normal course of events, infants acquire language. Animals do not. In both instances, however, there is abundant evidence of thinking without language. It should be possible, at least in principle, to study simple and complex learning in subjects who lack consciousness and language and thereby to evaluate the contributions of both factors in a developmental perspective (Terrace 1985).

Little is known about the processes that enable a child to attend to the contents of consciousness and to relate them to what the child learns: the covariation of environmental events. There are theories available, however, formulated by behavioral and cognitive psychologists, about the development of consciousness (Nisbett & Wilson 1977; Skinner 1945). These theories, which suggest the kind of research needed to evaluate properly the role of consciousness in learning and other cognitive processes, assume that infants develop sustained consciousness of their experiences ever so slowly under the tutelage of their parents, siblings, friends, teachers, and so on.

Because learning in infants is initially unconscious, it is much more difficult to study in infants than in adult human subjects, who can introspect and report the contents of their consciousness. That is a defensible rationale for studying learning in adult human subjects. It is unlikely, however, that such research will provide a basis for formulating basic principles of learning; at least not until it is recognized that learning with consciousness presupposes the much more pervasive and fundamental ability to learn without consciousness.

ACKNOWLEDGMENT

The research described in this commentary was supported by two grants from NIMH (MH40402).

On the creation of classification systems of memory

Daniel B. Willingham

Department of Psychology, University of Virginia, Charlottesville, VA 22903; dbw8m@virginia.edu

Shanks & St. John (S&S) set out to convince the reader that their classification system (CS) of memory (fragment based/rule based) is superior to another CS (implicit/explicit). I see two difficulties with their argument. Both are best understood in the context of the creation of CSs.

A CS is a method of classifying objects – or, in this case, memories – as similar or dissimilar to one another. Creating a CS requires selecting attributes of memories that are used to divide memories into classes. How are the attributes selected? One can construct a phenetic CS, which entails listing many attributes of memories and conducting a cluster analysis (e.g., Sokal & Sneath 1963) to both derive the classes and identify which attributes are important. Phenetic taxonomy is rarely used today, because selecting the attributes to include, the way to measure them, and the clustering metric are inherently subjective processes (Fleishman & Quaintance 1984).

The preferred method entails three steps. First, a theoretical principle is selected. The theoretical principle describes some property of nature that the CS is to describe and it also dictates which attributes of an object will be used to distinguish classes (Ridley 1986). For example, the theoretical principle of most CSs of species is phylogeny; they describe the branching of the evolutionary tree. Within memory research, two theoretical principles have been used: neurobiology, and information processing.

Second, attributes of objects are selected that will distinguish classes. The attributes selected depend on the theoretical principle. If the theoretical principle is neurobiology, a relevant attribute might be the contribution of a neurotransmitter to encoding the memory (e.g., level of acetylcholine (ACh) in the basal forebrain). If the theoretical principle is information processing, a relevant attribute might be the dependence on a hypothetical representation or process (e.g., use of an induced rule).

Third, the values of the attributes for class inclusion are selected. For example, if the theoretical principle is neurobiology and the attribute is the effect of the level of ACh in the basal forebrain, one determines the levels of ACh that describe memory classes. For example, there might be two classes: memories impaired by low levels of ACh, and memories not impaired.

S&S have overlooked the role of selecting a theoretical principle, and they have, in my view, erred in their interpretation of how attribute values may be selected.

S&S's failure to appreciate the role of a theoretical principle is revealed in their treatment of the amnesic data. They rightly point out that normal learning by amnesic patients does not bear on whether patients were aware of the contingency at the time of learning. However, S&S fail to recognize and articulate the two theoretical principles (neurobiological and information processing) embedded in their criticism, and thus overlook the real reason the amnesic data are important; namely, because these data yield a neurobiologically based CS of memory that parallels an information-processing-based CS of memory. Paradigms that S&S cite as potential examples of unconscious learning (which therefore fall into the same class in an information-processing-based CS) elicit normal performance in amnesia (and therefore fall into the same class in a neurobiological CS): classical conditioning (Weiskrantz & Warrington 1979), artificial grammar learning (Knowlton et al. 1992), instrumental control tasks (Squire & Frambach 1990), and serial response-time tasks (Nissen & Bullemer 1987). In addition, a variety of semantic and repetition priming effects (see Shimamura, 1986, for a review) fall into the same class by an information-processing principle (they are dissociable from awareness in control subjects) and by a neurobiological principle (they are preserved in amnesia).

This concordance is important, as most researchers would agree that there is a correspondence between neuroanatomic structure and cognitive representation; these two theoretical principles should lead to the same CS, and they do.

The fragment-based/rule-based CS does not account for data from amnesic patients. S&S point out that intact learning in amnesia is almost always fragment based, but amnesic patients are impaired in learning situations that would be classified as fragment based (e.g., a list-learning task).

The third step in CS development bears on S&S's chief argument against the implicit/explicit distinction. Given that an attribute for classification is to be consciousness of the contingency at learning, one must select a value of this attribute to distinguish classes of memories. S&S argue that the attribute is irrelevant because there is insufficient reason to believe that any learning is unconscious. They argue, in effect, that the attribute has no variance and so cannot be used to differentiate memories.

I disagree with the contention that unconscious learning has not been demonstrated. One must bear in mind that any putative demonstration of unconscious learning entails accep-

tance of the null hypothesis. One can always argue that the test of awareness was insensitive, did not match the conditions present at encoding, or was unreliable. Caution is appropriate in calling subjects unaware, but one must ask how an experimenter can be any more certain that subjects do not have explicit knowledge when subjects repeatedly deny noticing a sequence, make errors when forced to guess about its characteristics, and fail to recognize it (Willingham et al. 1993).

Even if unconscious learning has not been satisfactorily demonstrated, consciousness may still be an important attribute in CSs of memory. S&S take the position that consciousness can only take on two values (conscious or unconscious) as an attribute in a CS. A complete lack of awareness may be one end of a spectrum of states of awareness of the contingency at learning; if unconscious learning cannot take place, that does not mean that the rest of the spectrum is irrelevant.

In sum, the implicit/explicit CS appears preferable to the fragment-based/rule-based CS because: (1) the latter is discordant with the amnesic data, whereas the former fits those data and shows that two different theoretical principles – information processing and neurobiology – lead to the same CS; and (2) the chief argument against the implicit/explicit CS (that learning is never unconscious) is difficult to falsify, because one who believes the testing of awareness was not adequate can raise objections that cannot be discounted. Furthermore, if unconscious learning did not exist, it would not sound a death knell for the implicit/explicit CS. For this CS to have validity, there need only be variability in the attribute of consciousness, and predictive power in that variability.

ACKNOWLEDGMENT

The author thanks Maggie Keane for very helpful comments on an earlier draft of this commentary.

Authors' Response

How should implicit learning be characterized?

David R. Shanks^a and Mark F. St. John^b

^aDepartment of Psychology, University College London, London WC1E 6BT, England; ^bDepartment of Cognitive Science, University of California at San Diego, La Jolla, CA 92093

Electronic mail: ^adavid.shanks@psychol.ucl.ac.uk;

^bmstjohn@cogsci.ucsd.edu;

The commentators on our target article have raised a very broad range of issues, and in our response we will attempt to answer as many of the key points as possible. We thank the commentators for the careful consideration they have given to our arguments. We should say at the outset, however, that although our arguments have been challenged at a number of places, we have not been led to abandon our view of the implicit learning literature. We continue to believe that although the instance/rule distinction provides a meaningful characterization of human learning, it remains highly questionable whether unconscious learning has been adequately demonstrated.

The organization of our response is shown in Table R1. Our plan is first to say something about the appeal of implicit learning, and then to consider the question of unconscious learning: whether our Information and Sen-

sitivity Criteria are appropriate, how consciousness and learning should be defined, what the role of consciousness in learning might be, and whether the further evidence for unconscious learning that commentators discuss warrants a revision of our conclusions. Following these sections, we will consider other dimensions for characterizing distinct learning systems. Is the fragment/instance memorization versus rule-induction distinction a useful and valid one? And finally, can we refine and further specify what is learned and later used by the fragment/instance system?

R1. The appeal of implicit learning

Carlson points out that the idea of implicit learning has an appeal that defies dismissal. Part of this appeal might simply be the cachet of a mysterious process. But more seriously, there clearly is something to implicit learning. Most, if not all, of the commentators agree that there is a real phenomenon, one characterized by Holyoak & Gattis as "(a) knowledge about covariations in the environment, (b) learned by exposure to stimuli exhibiting the covariations . . . (e) [knowledge that] is not fully verbalizable and (f) [knowledge that] is not manipulable in the sense that it cannot be re-represented explicitly to serve as input to other procedures." We agree; the task at hand is to further analyze this phenomenon. One question is whether the implicit learning process and the resulting information are unconscious. The other question concerns what this knowledge is like and what sort of learning process acquires it.

Several commentators (Berry, Lindsay & Gorayska, Reber & Winter) have challenged an important aspect of the overall logic of the target article. We have assumed that the sensible approach is to be skeptical about unconscious learning until it is demonstrated, but these commentators argue that exactly the opposite assumption should be adopted, namely, that unconscious processes have some sort of priority over conscious ones. For example, Berry says that we "adopt the position that unless researchers can prove conclusively that learning is unconscious then we should assume that it is conscious," but we "never justify why conscious processing should be the default assumption." Our argument is simply that explicit learning is the agreed starting point, since no one doubts that learning is at least some of the time accompanied by awareness of what is being learned. What remains to be shown is whether unconscious learning is also feasible.

The position advocated by Berry, Lindsay & Gorayska, and Reber & Winter is tantamount to holding that unconscious learning does not need to be demonstrated in the laboratory. But no amount of a priori argumentation is going to demonstrate its existence: what is needed is clear experimental evidence. If we do not know in a certain task whether knowledge is conscious or unconscious, it is surely foolhardy to *assume* the latter.

R2. Criteria for establishing unconscious learning

Before discussing potential problems with our conceptualization of implicit and explicit learning, we would like to mention briefly a very interesting article published after

Table R1. Organization of the authors' response to specific commentaries

Topics	Commentators
R1. The appeal of implicit learning	Berry, Carlson, Holyoak & Gattis, Lindsay & Gorayska, Reber & Winter
R2. Criteria for establishing unconscious learning	
R2.1. The Sensitivity and Information Criteria	Berry, Brody & Crowley, Carlson, Catania, Dienes & Perner, Holyoak & Gattis, Lindsay & Gorayska, Merikle, Overskeid, Perruchet & Gallego, Rakover, Reed & Johnson, Stadler & Frensch, Svartdal
R2.2. Defining consciousness and learning	Baeyens et al., Catania, Howe & Rabinowitz, Overskeid, Packard
R2.3. Measuring consciousness	Andrade, Baeyens et al., Catania, Dienes & Perner, Ennen, Nagata, Overskeid, Reber & Winter, Stadler & Frensch, Squire et al., Terrace
R2.4. Defining explicit learning and consciousness by verbal report	Catania, Cleeremans, Dienes & Perner, Lindsay & Gorayska, Merikle, Overskeid, Rakover, Terrace, Willingham
R2.5. Do prediction and recognition tasks measure conscious or unconscious knowledge?	Cleeremans, Dienes & Perner, Rakover, Seger
R2.6. The possible causal role of awareness	Catania, Poldrack & Cohen, Squire et al., Svartdal
R3. Further evidence for implicit learning	Overskeid
R3.1. The mere exposure effect	Bornstein, Brody & Crowley
R3.2. Conditioning	Baeyens et al., Kimmel
R3.3. Instrumental learning	Andrade, Berry, Overskeid, Rizzo & Parlangei, Svartdal
R3.4. Sequence learning	Cleeremans, Reed & Johnson, Stadler & Frensch
R3.5. Amnesia	Cleeremans, Holyoak & Gattis, Kourtzi et al., Packard, Poldrack & Cohen, Reber & Winter, Seger, Squire et al., Willingham
R3.6. Language learning	Goldstone & Kruschke, Lachter, Nagata, Reber & Winter
R4. How should implicit learning be characterized?	Holyoak & Gattis, Kourtzi et al., Merikle, Nagata
R4.1. Consciousness and models of implicit learning	Cleeremans
R4.2. The fragment/instance memorizer	Cleeremans, Kimmel, Marsolek, Perruchet & Gallego, Seger
R4.3. The relationship between memorization and rule induction	Goldstone & Kruschke, Howe & Rabinowitz, Nagata, Packard, Poldrack & Cohen, Reber & Winter

the target article was completed and which, we believe, provides overwhelming justification for our strict criteria. Farah et al. (1993) were interested in the relationship between implicit (or "covert") and explicit ("overt") processing in prosopagnosics, patients who, as a result of brain damage, have difficulty recognizing faces. The typical pattern of findings is that whereas prosopagnosics may not be able to name a familiar, visually presented face, they will nevertheless process that face faster than an unfamiliar face during the course of a task, thus suggesting that some face knowledge is preserved. For example, less time is taken to judge a pair of familiar faces the same than to judge a pair of unfamiliar ones the same. Just as with implicit learning, the suggestion is that the person may have knowledge that is not available to consciousness.

Farah et al., however, have provided compelling evidence that the dissociation between implicit and explicit processing in prosopagnosia is attributable to differential test sensitivity. They constructed a parallel distributed processing model which could be presented with coded representations of faces. Two aspects of the network's performance were monitored: overt recognition and speed of processing. When Farah et al. lesioned their network, they were able to reproduce all of the major

dissociations of implicit and explicit performance seen in prosopagnosics, but these dissociations came from a system that has a single (albeit distributed) knowledge source. Thus it follows that the dissociations must have arisen because the implicit tests were better able than the explicit ones to detect small amounts of residual knowledge in the system. [See also Farah: "Neuropsychological Inference with an Interactive Brain: A Critique of the 'Locality' Assumption" *BBS* 17(1) 1994.]

This conclusion is, of course, exactly parallel to our argument that many examples of apparent implicit learning can be explained by the fact that the explicit test does not meet our Sensitivity Criterion. Below, we describe some analogous simulations we have conducted of implicit learning in amnesia.

R2.1. The Sensitivity and Information Criteria. Several commentators endorsed the Sensitivity and Information Criteria (Berry, Brody & Crowley, Carlson, Perruchet & Gallego, Reed & Johnson, Svartdal), whereas others raised concerns over them. Dienes & Perner, Holyoak & Gattis, Lindsay & Gorayska, and Overskeid all argue that our criteria are too stringent and make it very difficult, in principle, to dissociate learning and awareness. Dienes and Perner, for example, propose a thought-experiment

in which a subject has instance-based, but unconscious, knowledge of a grammar; they argue that if the performance and awareness tests are made sufficiently similar to meet our criteria then it is very difficult to imagine a mechanism that would apply the knowledge in one case but not in the other.

We have two responses to this. The first is that there are many people who believe that there *are* studies that meet these conditions, in which one test does reveal knowledge whereas another very similar one does not (see commentaries by Reed & Johnson, Stadler & Frensch). The second is that if the criteria are relaxed, then although the likelihood of observing dissociations between our two tasks increases, so does the danger that the dissociation is due to differential sensitivity or a mismatch in the type of information examined by the two tests. We believe that if it is difficult to demonstrate unconscious learning, that is all the more reason to be cautious about its existence.

Catania, Holyoak & Gattis, and Lindsay & Gorayska go further and argue that we have effectively equated learning and awareness, with the consequence that we have (in Catania's words) "designed a set of conditions such that learning cannot occur without awareness." But this represents a gross misunderstanding of our position. We have suggested that there are two types of measures. Some, (conditioned responses, sequential reaction times, grammaticality judgments) provide indices of learning. Others (recognition, prediction, verbal reports) may provide useful evidence concerning the content of consciousness. Although it would be very controversial to claim that the learning tests provide pure (or even partial) assessments of awareness, we have shown that when the results of these two sorts of tests are compared, it turns out as a plain empirical fact that learning and awareness are typically associated.

Merikle elaborates some of the issues raised in the target article concerning the logic of single dissociations, arguing strongly that this logic will always founder on the impossibility of demonstrating beyond doubt that a given test of awareness is exhaustive. We are probably not quite as skeptical as Merikle about the possibility of establishing unconscious learning in experiments using the single dissociation logic, but we agree wholeheartedly that his concerns should be addressed more adequately by defenders of implicit learning. Merikle, Perruchet & Gallego, and Svartdal all advocate the exploration of alternative strategies, an idea we fully support.

Perruchet & Gallego and Brody & Crowley raise the very interesting but thorny issue of learning during the test. Perruchet & Gallego refer, for example, to evidence that above-chance performance in a grammaticality judgment test may be due to learning taking place in the test rather than to the retrieval of information from an earlier study phase. As evidence of this, Perruchet (1994) reports performance significantly above chance in subjects not given a study phase. A related effect, noted in the target article, is that subjects may observe their performance during testing and thereby become aware of their implicit knowledge. These possibilities are worrisome, because they challenge the idea that tests (either implicit or explicit) may reveal whether prior learning was conscious or unconscious. Rather than providing evidence about earlier processing, performance on the test is under the control of events happening during the test itself. We fully

agree with Perruchet & Gallego and Brody & Crowley that this is a real possibility that requires much further investigation.

Rakover offers an entirely new method of determining whether a task can be learned without awareness, but we believe his procedure is flawed. He proposes that a task is "not susceptible to introspective awareness" if the subject, having been told the rule governing the task but not the specific content of the rule, is unable to discover even a part of the content of the rule. As an example, he cites data reported in the Lewicki et al. (1987) study, which was extensively discussed in the target article. In a pilot study, Lewicki et al. informed subjects that there were rules governing the position of the target on complex trials and offered them \$100 for finding any parts of the rules. None was able to do so, from which Rakover concludes that the task is not susceptible to introspective awareness. He then suggests that if such an unsusceptible task is learnable, it follows that the learning must be unconscious. Because Lewicki et al. obtained reaction time (RT) speed-up when subjects were exposed to the RT version of the task, Rakover's analysis suggests that learning was unconscious. The structure of the task could not be determined by conscious analysis but could be learned implicitly.

The problem with such an account, however, is that it is left up to experimenters to specify in the susceptibility test the "rule" they think is relevant. But if this rule is not what is responsible for learning in the later performance test, the experimenter will erroneously conclude that learning was unconscious when it may well have been conscious. As we argued, the RT speed-up in Lewicki et al.'s experiments was almost certainly not due to knowledge of the rules that were required in the susceptibility test. In sum, Rakover's account falls foul of the consequences of not adopting the Information Criterion.

None of the commentators seems to question the status of the Information Criterion. We all appear to be in agreement that the test of awareness must focus on the same information that underlies successful task performance. One point made by Carlson, which we readily accept, is that it is not merely stimulus-specific information that is relevant but also self-specific information about such things as intentions. *All* of the information that plays a role in task performance – including such things as goals and intentions – must be known before we can judge whether an awareness test meets the criterion.

R2.2. Defining consciousness and learning. Several commentators (Baeyens et al., Catania, Overskeid, Packard) wondered why we had failed to define learning and awareness, but we regard doing so as a futile exercise. Surely the past thirty years of research on human concept learning should have taught us that some concepts are simply ill defined and fuzzy? And surely everybody realizes that any definitions of these terms we offer would immediately be challenged? Readers of Dennett (1991) will search the book in vain for a "definition" of consciousness, but just because a concept cannot be given an analytic definition, does not mean it is outside the scope of scientific investigation. [See also Dennett & Kinsbourne: "Time and the Observer: The Where and When of Consciousness in the Brain" *BBS* 15(2)1992.]

With respect to learning, Roediger (1993) offers a clear illustration of the difficulty in finding a suitable definition.

He reports that the average duration of labor in giving birth to first-born babies is 9.5 hours, whereas with later-born babies it is 6.6 hours. Clearly, with second and third children the amount of time the mother spends in labor is much less than with first children. Does this mean that something has been learned which explains the speed up? On the one hand, it could be argued that if repetition priming is an example of learning (as indexed, say, by latency of tachistoscopic identification), then so should this "priming" effect in childbirth. On the other hand, it seems strange to say that the female reproductive system is capable of "learning" and "remembering." We suggest that borderline cases such as this and the ones cited by Overskeid (e.g., the immune system "learning" to recognise viruses) illustrate the futility of trying to define learning.

Baeyens et al. claim we are arguing that "any verbal or nonverbal behavior which might reasonably be thought of as reflecting a judgment based on a belief-that-*x* should be treated as an index of awareness-of-*x*," but we certainly never argue for this in the target article. Apart from anything else, such a definition would probably make unconscious learning a logical impossibility. Our approach is wholly pragmatic – we offer no hard and fast criteria as to what counts as a test of awareness. However, we do believe that if a subject has an expectancy of a future event and is able to project that expectancy into a truly instrumental action (as opposed to a reflex), then the expectancy is almost certainly conscious. At present, it is fair to say that we do not know for certain whether responses in a prediction task are instrumental acts or reflexes (see Dickinson 1988 and Jacoby et al. 1993, for possible ways of distinguishing between these). We discuss the prediction task further in section R2.5. In fact, we are largely in agreement with Baeyens et al.'s criteria for what counts as consciousness. Our only dispute would be that it is a truism that "awareness-of-*x* is the presence . . . of a phenomenal state of having-the-subjective-impression-to-consciously-know-that-*x*."

Howe & Rabinowitz take us to task for being unspecific as to what it is that the subject needs to be aware (or unaware) of. They point out that there are three ways in which consciousness may be related to learning: "the organism as conscious while learning, the organism as being conscious of learning, and the organism being conscious of what is being learned." We thought we had been clear about this, but to reiterate, the question is whether or not the subject is aware of the information that is responsible for improved performance on the implicit test (i.e., the third case above). Whether an organism is aware in the general sense, or whether it is aware that it is learning, is simply irrelevant to the issue of implicit learning.

R2.3. Measuring consciousness. In their thoughtful commentary, Dienes & Perner suggest that performance on, for example, a prediction test measures the objective rather than the subjective threshold of awareness. Subjects may perform above chance on such a test despite believing that they are guessing, in which case their knowledge should be described as unconscious, and by inference, their learning should be described as unconscious as well.

We have two points to make about this. First, in the sorts of studies reviewed in the target article, we do not know whether this is what subjects believe – they may well believe that their prediction performance is above chance (though the unpublished study Dienes & Perner cite, in which subjects reported that they were guessing, sounds very promising). Second, and more important, the problem with the subjective/objective threshold distinction is the way in which the subjective threshold is assessed. Subjects are notoriously poor at knowing what randomness is, so when asked whether or not they believe they are guessing, there is no reason to suppose that a veridical answer concerning their state of awareness is given. Moreover, the subjective threshold is often assessed *post hoc*: at the end of a block of trials, subjects are asked to evaluate their performance across that block. It would hardly be surprising in the circumstances that very ephemeral states of consciousness during the prediction trials might be forgotten or discounted by the time of the evaluation. The subjective threshold technique may yield results different from the objective one only because the testing conditions are different and the subjective test is less sensitive. (This is much less of a problem in blindsight, of course, because the subjective judgment about performance is made in the presence of the stimulus to which the subject is responding. We certainly would not want to dispute the importance of the subjective/objective difference in such cases as blindsight).

In citing a test which they construe as measuring unconscious learning but which we view as measuring conscious knowledge, Baeyens et al. attempt to point out the unacceptability of our methods of assessing awareness. In the evaluative conditioning study they cite, subjects were exposed in the learning phase to a relationship between line thickness in a geometrical figure and the valence of a subsequently presented picture (positive or negative). In the critical test, subjects saw a geometrical figure and were told that a positive or negative picture would be presented subliminally immediately after it. Subjects had to say whether they thought they had seen a positive or negative picture, even though none was actually presented. Subjects performed above chance on this test: they reported seeing a positive slide, for example, if the geometrical figure had been paired in the study phase with a positive slide.

Baeyens et al. want to offer this as a test of unconscious knowledge, but we see no strong evidence supporting that position. It is equally possible that the subjects, as a result of the study phase, have a (very fragile) conscious expectancy of a picture of a certain valence given a certain geometrical figure. Their expectancy may then bias the response they give, such that this forced-choice test can be interpreted as assessing awareness. We believe the burden of proof is on those who support the idea of unconscious influences to demonstrate that their interpretation is the only viable one. In this case, we suggest it is not.

In support of unconscious processes in general, Overskeid claims there is much evidence that our decisions may often be taken by our brains prior to any conscious intention to act, and that we "jump on a running train believing we started it." As evidence, he cites the well-known experiments of Libet (1985), but these have been

subjected to so much criticism (see Dennett 1991) that it is highly premature to treat them as conclusive.

A number of commentators (Overskeid, Reber & Winter, Squire et al., Terrace) draw strong inferences from the fact that learning can occur in preverbal children and in lowly organisms such as the invertebrate *Aplysia*. The argument is that we are on the horns of a dilemma: either our thesis forces us to conclude that such organisms are conscious, or, if it is assumed that such organisms cannot possibly be conscious, then the forms of learning they are capable of must be unconscious. We believe that neither of these arguments is valid. First, just because a regularity occurs in humans does not mean it holds for other species. No amount of argumentation from animal experiments is going to alter the fact that a correlation between awareness and learning exists in humans.

Second, it cannot simply be asserted that nonhuman organisms are unconscious (see Dennett 1991; Griffin 1992, Nagel 1974). In fact, we believe that little productive purpose is to be gained in discussing animal consciousness at all, since there is no agreed-upon means of deciding what counts as consciousness in animals, and our intuitions clearly differ from those of some commentators. Reber & Winter are confident that animals are unconscious, but they provide no reasons for their view. Similarly, after describing evidence of learning in two- to three-month-old infants, Terrace states (correctly, in our view) that "whether infants are aware of their expectations is entirely a matter of speculation," but then goes on to conclude that "since consciousness in infants cannot be subjected to empirical study, there is no basis for assuming that infants are conscious of what they learn." But equally, there is no basis for assuming that they are *not* conscious. Terrace gives no rationale for his belief that there is "compelling evidence of learning without consciousness in infants and animals." To repeat, we doubt that the matter is decidable, and in any case it has little bearing on our conclusions. The target article, and the experiments it reviewed, are about adult human learning.

Catania proposes a number of animal experiments analogous to human implicit learning ones (see McLaren et al., in press, for similar examples) to illustrate his belief that it is only verbal reports, and not other forms of behavior such as recognition or prediction, that are of relevance to the concept of consciousness. [See also Lubinski & Thompson: "Species and Individual Differences in Communication Based on Private States" *BBS* 16(4)1993.] For example, in a pigeon analogue of a sequential RT experiment, it would be possible to set up both an RT measure and a prediction measure. Catania suggests that if the pigeon's RT and prediction responses correlated, such that RT improvements never occurred in the absence of above-chance prediction responses, our position would force us to conclude that the pigeon is aware.

We went to some lengths to acknowledge that prediction responses might be susceptible to unconscious influences (see below), so the conclusion that the pigeon is conscious would not necessarily follow. But our main response to Catania is to turn his question around: What would he conclude from his pigeon experiment if the RT and prediction measures dissociated? Would one not have to conclude that there is a source of information that is

available to one response but not the other, and (if the tasks are very similar) that what characterizes this information is that it is unconscious? Surely, on his reasoning, this would be even stronger evidence for conscious processes (in the control of prediction responses) in nonverbal organisms. Of course, we doubt that such dissociations occur in humans so there is little reason to anticipate them in other organisms, but that is no reason to believe that consciousness is a logical impossibility in animals.

Ennen recalls the Rylean distinction drawn between "learning-how" and "learning-that." She argues that learning-how does not involve acquiring information so it cannot be reported. Therefore, it must be learned implicitly. Unfortunately, although her argument seems sound, (that is, although unconscious learning appears to be a sound logical possibility), there does not appear to be any clear evidence for it. Take conditioning, for example. This is a perfect example of knowing-how, a form of learning that can be characterized as "the incremental modification of perceptual-motor connections over time such that the agent acquires a neurophysiological disposition for improved performance." Yet in humans, conditioning is always accompanied by awareness. Thus Ennen's analysis does nothing to challenge the claim that learning is always accompanied by consciousness.

Stadler & Frensch and Nagata suggest that learning is implicit when it is unaffected by intention. It is quite reasonable to define theoretical terms as one wants, but this creates an entity different from the "implicit" learning that is the focus of the target article, because intention is synonymous neither with awareness nor with abstraction. These commentators are surely right that learning can be either intentional or unintentional, but to focus on the learning *process* rather than the learned information is simply to change the subject. Remember that the thesis we are assessing proposes that information can exist in such a form as to be able to project itself onto behavior without being represented in consciousness. This is a thesis about information, not process.

Andrade points to a motivational issue in measuring awareness among patients under general anesthesia. She discusses the isolated forearm technique which reveals high levels of awareness in anesthetized people, but, as she notes, even this test is fallible, because of motivational changes in subjects. Not only is it the case, as we argued in the target article, that evidence for learning and memory during anesthesia is contradictory (see Merikle & Rondi, 1993, for an even more skeptical review), it is also doubtful that anesthetized subjects are always unconscious of events happening around them.

R2.4. Defining explicit learning and consciousness by verbal report. Some commentators (Catania, Dienes & Perner, Lindsay & Gorayska, Overskeid, Rakover, Terrace) suggest that verbal report is the best place to draw the line between explicit and implicit learning. Lindsay & Gorayska argue that it may be more difficult than we assumed to prove that another test is more sensitive or exhaustive than is verbal report, and therefore only the latter should be adopted as an index of consciousness. But not only does this immediately rule out animals and nonlinguistic humans as being conscious, it also makes unconscious learning a logical necessity. Whereas we are

accused of making unconscious learning a logical impossibility, Lindsay & Gorayska appear to be doing the exact opposite! We agree with Merikle that establishing a test as exhaustive is likely to prove very difficult, and we doubt that Lindsay & Gorayska have any case for establishing this with respect to verbal report.

The view that there is something special about verbal reports appears frequently in the commentaries, but we see no basis for adopting this strong position. There is no clear boundary between reports and other tests of awareness; instead, we take verbal reports to be merely one of several means of assessing awareness. For example, Marcel (1993) has shown that subjects can be much better at using responses other than verbal reports as indicators of awareness. Marcel presented subjects with a brief light close to the threshold of detection and instructed them to report on each trial whether they saw the light by blinking, pressing a button, and saying "yes." Subjects on the same trial reported with their finger that they saw the light while reporting orally that they did not, and vice versa. Most interestingly, verbal reports did not turn out to be the most sensitive index on detection: blinking, under some conditions, was far more sensitive. In the face of such findings, it is very difficult to assign a special role to verbal reports.

Other commentators point out that there is a qualitative difference between learning in subjects who can retrospectively report and justify their performance and subjects who deny having any knowledge of their performance. However, both the Sensitivity and Information Criteria suggest that verbal recall is a poor place to draw a line between learning systems. Reportable knowledge varies between subjects, their vocabularies, the retrieval cues available, prompting, set, certainty of the knowledge, and so on. We see little justification for a qualitative difference at the point of spontaneous or prompted verbal recall. We do see the level of verbal report as an interesting and legitimate independent variable, but one with quantitative rather than qualitative levels. And we agree with Willingham that quantitative differences make for poor taxonomies.

Two commentators (Cleeremans and Willingham) make the related proposal that consciousness varies on a continuum. As training continues, initially unconscious knowledge becomes increasingly conscious. Changes in the degree of consciousness would account for observed changes in the amount of information reported verbally as well as for changes in prediction accuracy and even reaction time. However, we prefer the interpretation that the information represented in consciousness increases in amount and in strength (i.e., certainty). Our argument, via the Sensitivity Criterion, is that verbal report places a high threshold of certainty and retrievability, whereas other tests, such as the prediction task, place a lower threshold. The lower threshold allows less certain information to be demonstrated.

Besides quantitative differences, of course, the content of verbal reports may vary qualitatively. As we discuss in the target article, learning (and testing) protocols may demonstrate heavy reliance on memorization and instance learning or they may demonstrate heavy use of reasoning and problem-solving skills and hypotheses. Note that although we believe that verbal reports should be treated with care, it is not the case (as Catania

suggests) that we have ruled out verbal reports as relevant to judging a learner's awareness. They are certainly relevant; they may just not be exhaustive.

R2.5. Do prediction and recognition tasks measure conscious or unconscious knowledge? Several commentators (Cleeremans, Dienes & Perner, Rakover, Seger) raise the issue, addressed at some length in the target article, that prediction and recognition tests may not be "process-pure," that is, they may not provide pure indices of conscious knowledge but may be contaminated by unconscious knowledge. We reiterate that we believe this is a perfectly valid objection which will only be resolved by further investigation. At present we probably do not have sufficient information to resolve the issue.

Dienes & Perner criticise prediction and recognition tests, because such tests do not require subjects to describe the underlying basis of their responses: a subject may make a correct prediction without being able to justify that prediction. But this would only be relevant if we thought that justification played a role in the performance measure (e.g., RT). That is, if subjects only responded quickly in the RT phase when they could describe the basis of their expectation about where the next signal would appear, then to meet the Information Criterion, the awareness test would have to pick out only those cases where justification was available. But in general we have no reason to believe that justifications play a role in performance tasks, in which case Dienes & Perner's position has little direct support.

R2.6. The possible causal role of awareness. We made no inferences concerning the possible causal role of awareness. Our aim was merely to show that, as an empirical fact, awareness is necessary for learning. Some commentators (Catania, Poldrack & Cohen) wonder about the specific role that consciousness plays in learning. The view that awareness is necessary for learning could entail a range of possibilities, with a causal role merely representing one extreme. At the other extreme, awareness may have no causal role with respect to learning but may be correlated with it via a third factor.

Far from being tangential to the question of implicit learning, (as Svartdal claims) the correlation between level of performance and verbal report demonstrates the coupling between performance and consciousness. We have no particular reason for favoring one interpretation of the role of consciousness over any other. We note, however, that subjects may consciously choose a learning strategy, such as hypothesis testing or memorization, and that this choice affects learning. In this sense, at least, consciousness affects what is learned.

Squire et al. claim that in cases where the two are correlated, consciousness is epiphenomenal to learning. Thus, despite the fact that conditioning cannot be shown to occur in the absence of awareness, the latter plays no causal role and is merely a nonfunctional "shadow" that accompanies learning. But, as Lovibond (1993) argues eloquently, the correlation between conscious knowledge and learning is incredibly close. The epiphenomenalist position therefore looks decidedly strange, because if consciousness is an epiphenomenon, it ought to be possible to dissociate it from learning. There is no evidence that this is possible, however. (We argue below that

amnesia does *not* represent a case in which awareness and learning are uncoupled.)

R3. Further evidence for implicit learning

Inevitably, commentators have cited studies that we failed to cover in the target article. We apologise if, by wittingly or unwittingly excluding certain pieces of evidence, we have reached erroneous conclusions.

Some commentators take us to task for considering such a limited and artificial range of learning tasks and for drawing more general conclusions about human learning than is warranted by these studies. To an extent, we cannot help but agree, but we also believe that complex real-world learning is largely reducible to simpler learning. By considering such basic learning paradigms as Pavlovian conditioning, we therefore hope to encompass more complicated and realistic learning situations. Besides, we know of no careful studies of the contents of people's consciousness when they are learning to play tennis (see Overskeid).

R3.1. The mere exposure effect. Bornstein maintains that data from experiments on the mere exposure effect provide compelling evidence for unconscious learning. In such experiments (e.g., Kunst-Wilson & Zajonc 1980), subjects show an increased preference for stimuli which have been repeatedly exposed for brief durations, but they cannot recognise those stimuli. We have no dispute over the robustness of such effects, but we do have strong reservations about whether they demonstrate implicit learning.

It is odd that Bornstein persistently refers to these experiments as involving stimuli perceived without awareness, because the evidence that the stimuli in most of these experiments are subliminal is very weak. For example, they are often presented unmasked, and a forced-choice present/absent test is rarely included (see Bornstein et al., 1987, for an exception). This is important, because if the original stimuli were genuinely subliminal, then any preference for them at test would have to represent a role for unconscious information.

Be that as it may, dissociations clearly do emerge between subsequent preference and recognition judgments, whether the stimuli were consciously perceived at study or not, and this finding may also provide evidence (albeit less direct) of unconscious processes. In this case, the idea would simply be that there is a source of information available for making the preference but not the recognition judgment, and this information may therefore be the end-product of an unconscious learning system operating in tandem with the conscious system. However, we pointed out that hypermnesia seems to occur for recognition responses in these experiments (Merikle & Reingold 1991), and we cited this as evidence that in making recognition judgments, subjects may initially discount a source of *conscious* information such as perceptual fluency (Brody & Crowley propose a similar account in their commentary). Bornstein does not provide an alternative explanation of the hypermnesia finding.

Moreover, it turns out that some very recent evidence supports our position and directly refutes the view that the preference and recognition judgments are based on

different sources of information. Whittlesea (1993) has shown that the dissociation of affect and recognition can occur *even for items that did not appear in the study phase*. In his experiment, he required subjects to rate words for pleasantness and also to judge whether they had been recently presented; in fact, the critical words had not been preexposed. Because of a manipulation of the context in which the words appeared, some words were processed more fluently than others, and this led them to be rated as pleasant, but these same words were not called "old." Thus the dissociation of affect and recognition occurs even for novel items.

This result directly contradicts the idea that above-chance affect judgments are based on unconscious information encoded during the study phase. Instead, as Whittlesea argues, the result must be attributable to the different ways in which information (such as perceptual fluency) concerning a test item may be interpreted in affect and recognition tests.

R3.2. Conditioning. Our conclusion that awareness is necessary for conditioning is disputed by Kimmel, who cites as positive evidence a study by Grings (1965) in which subjects acquired autonomic CRs but "when questioned after three days of conditioning . . . revealed no knowledge about stimulus relations." Kimmel also cites a study (Kimmel et al. 1983) in which subjects could be shown to have learned about one of two coloured lights that signalled shocks but were unable to report whether it was the blue or yellow light, and another study (Lachnit & Kimmel 1993) in which subjects apparently learned to respond appropriately to CSs even if they were unaware of the relevant reinforcement contingencies.

Kimmel seems to think that such results are persuasive since "the subjects . . . were questioned very closely, and the questioners were motivated to detect any sign of awareness. At the very least, Shanks & St. John should describe what they consider inadequate in interview procedures from studies claiming conditioning without awareness, instead of rejecting them peremptorily or acting as if they do not even exist." We did not dismiss these studies peremptorily; rather, we consider them to have exactly the same flaws as other studies that we did cite. These flaws include using an awareness index that fails to provide the subject with the same quality of retrieval cues as the conditioning measure.

It does not matter how closely the subjects are questioned or how motivated the questioners were. If the subjects' knowledge is fairly ephemeral, then any decrement in the conditions of testing between the implicit and explicit tests will lead to less information being revealed. We leave it as an exercise for the reader to decide whether results such as those Kimmel cites provide conclusive evidence of unconscious learning. Incidentally, we have been unable to evaluate the Lachnit and Kimmel (1993) study that Kimmel cites, because that report does not in fact include any discussion or data concerning aware versus unaware subjects.

We do not believe we have given an unfair description of Baeyens et al.'s data. When subjects are given a concurrent test of awareness (rather than just a postexperimental questionnaire), more of them appear to be aware of the conditioning contingencies. Is this because the concurrent measure is more sensitive, or because it

actually creates awareness? Baeyens et al.'s belief in unconscious learning requires them to exclude the former possibility, but we see no reason (and they provide no evidence) for such a conclusion. There are any number of reasons why a concurrent test might be more sensitive than a *post hoc* one. For further criticism of the data on evaluative conditioning, see Davey (1994).

R3.3. Instrumental learning. We applaud Svartdal's use of an interesting alternative to the standard logic of dissociation to demonstrate unconscious instrumental learning. Svartdal recommends providing explicit instructions that focus the subject's attention on one aspect of responding (e.g., position), but arranging for reinforcement to be under the control of variations in another dimension of responding (e.g., force). With such a procedure, Svartdal (1992) showed that variations in the dimension of responding that was in fact reinforced became larger and larger, apparently without the subjects being aware of this change.

As Svartdal notes, though, such an experiment is still likely to be dogged by the problem of correlated hypotheses. He dismisses the claim that correlated hypotheses may have been the basis of his subjects' responding, because when asked to formulate advice to another hypothetical person on how to master the task, subjects typically gave very general answers and failed to verbalise any such correlated hypotheses. But until we know what the putative correlated hypothesis might be, it is very difficult to know what we should be looking for in these protocols. If the hypothesis is only partially used, or if several different correlated hypotheses are used concurrently or on different trials, or if the hypothesis is very complex, subjects' inability to give useful verbal reports would hardly be surprising. Because there can be no doubt that correlated hypotheses play a crucial role in many instrumental tasks (Dulany 1961), we reiterate our conclusion that the existence of such hypotheses makes it very unlikely that clear evidence of unconscious learning will emerge from instrumental tasks.

Berry rightly points out that we should have discussed the study by Stanley et al. (1989), which used the yoking procedure that we commended in the target article. However, we believe that this study provides far from convincing support for implicit learning. In their first three experiments, Stanley et al. observed that verbal protocols from subjects who have had extensive practice at a control task (sugar production or person interaction) could significantly improve the performance of naive subjects, indicating that a good deal of knowledge was available to verbal report.

In Experiment 4, Stanley et al. (1989) selected a sample (12 out of 53 subjects) of people learning the control task who had a clear transition point where their performance showed a marked and relatively sudden improvement. Stanley et al. found that yoked subjects given the pretransition verbal protocols from these subjects showed just as much benefit as yoked subjects given the posttransition protocols. From this Stanley et al. concluded that a sharp performance increase is not immediately accompanied by an increase in the quality of verbal report, because otherwise the yoked subjects given posttransition transcripts should have shown a much greater benefit than those given the pretransition transcripts.

Although this result is certainly interesting, we believe that there are at least two problems with Stanley et al.'s interpretation. First, a statistical artifact may have been introduced by the procedure of selecting subjects who showed a sudden transition in performance. Naturally, random variation accounts for some of the fluctuations in performance in these tasks, but if it is merely chance that is responsible for the selected subjects' transitions, then there is no reason to expect their accompanying verbal protocols to show any marked improvement after the transition.

Second, Stanley et al. found in Experiment 3 that naive subjects given a simple rule did not perform at the maximum level achievable with that rule. Subjects were told to "select the response level half-way between the current production level and the target level," which if followed could lead to 83% responses on target. However, subjects achieved only 50% correct responses. The obvious conclusion is that even when subjects have access to a useful rule, they will not always use it. Perhaps they are not confident that a rule is correct until they have tested it by comparing what happens when they do and do not follow it. But if that is the case, then there is no justification for supposing that the yoked subjects in Experiment 4 should have shown an immediate improvement in performance when given the posttransition protocols. In sum, the yoked design Stanley et al. used is very promising (and without question shows that subjects may have access to far more knowledge than would be predicted by other forms of questioning), but it has not established that any significant amount of knowledge is unconscious.

Rizzo & Parlangei argue that in complex control tasks where some real-world schema is available to the subject, subjects will use that schema as a basis for rule learning. They report that in some circumstances a subject will be unable to learn to control a system formally identical to one they can learn to control if the former retrieves schema-based knowledge that is incompatible with the contingencies of the task. Rizzo & Parlangei suggest that this is inconsistent with the existence of an independent implicit learning system, because such a system would be unaffected by the specific content of the task. We are not sure that this line of reasoning is watertight, but it is certainly interesting that such differences can emerge. Rizzo & Parlangei also report obtaining strong correlations between performance and verbal report.

Overskeid claims that instrumental learning is possible during sleep, but we see no good reasons to assume that people are completely unaware when they are asleep. If it is difficult, as Andrade suggests, to prove beyond doubt that people are unconscious when they are anesthetized, then it will surely be a good deal harder to prove it when they are asleep. Moreover, numerous well-controlled studies have failed to reveal any signs of learning during sleep (British Psychological Society 1992; Wood et al. 1992)

R3.4. Sequence learning. Our interpretation of Stadler's (1989) study is very nicely refuted by Stadler & Frensch. We argued that the failure of the prediction test to reveal any savings (although the RT data indicated that learning had occurred) may have been due to interference and hence rapid forgetting during the test, which provided no feedback to maintain accurate performance. Our hypoth-

esis was that evidence of savings may have been present on the earlier test trials but abolished on later trials as interference built up. Stadler & Frensch report that prediction performance was at chance even from trial 1.

Because we have argued that this study goes a good way toward meeting the Sensitivity and Information Criteria, we are happy to acknowledge that the results appear to provide good evidence of unconscious learning. Our only outstanding concerns would be (1) that the assessment of awareness was based on a fraction of the number of observations that were used to demonstrate RT speed-up, introducing possible problems of statistical power, and (2) that we continue to believe that withholding feedback in the prediction test works against the possibility of finding significant savings. But because the awareness test provided no hint of savings, the former concern is probably unimportant, and the latter concern can easily be avoided by providing feedback and comparing prediction performance with that of a control group trained on random patterns. Stadler's (1989) study clearly merits replication and further exploration to reveal why this experiment obtained evidence of a dissociation between RT speed-up and prediction performance whereas so many others, reviewed in the target article, have failed to do so.

Reed & Johnson also report a new experiment that seems to meet our criteria while still obtaining a dissociation between performance and awareness. In that experiment, subjects trained on the repeating reaction time sequence 12134231432 were then tested on the sequence 123413214243, which is matched in terms of single-event and pairwise frequencies. The increase in response latencies that Reed & Johnson (1994) observed must therefore be attributable to genuine sequence learning, involving at least knowledge of second-order structure. That this knowledge was implicit is suggested by the fact that performance on generation and recognition tests was at chance. In the generation test, for example, subjects saw sequences of 2 target locations from the training phase and they then generated the 10 locations they thought had followed the target sequence in the training phase. Reed & Johnson found that subjects were no more likely to make a correct first response following the 2 targets than were control subjects trained on a random sequence. The recognition test also revealed no evidence of explicit sequence knowledge.

We are somewhat less convinced than with Stadler's (1989) experiment that this provides compelling evidence of unconscious learning. First, with a relatively small number of trials per subject, the generation and recognition tests may not have been powerful enough to detect small savings from the training phase. Second, it is not at all clear why Perruchet and Amorim's (1992) subjects were able to perform above chance in generation and recognition tests whereas Reed & Johnson's were not. It is true that there are several procedural differences between the tasks, but the discrepancy remains troublesome. Third, and most important, we have found in a replication of Reed & Johnson's experiment that subjects can perform well above chance in a generation test. We trained subjects on the same sequence and under the same conditions Reed & Johnson used, but we then required the subjects to predict target locations in a standard prediction test containing 96 trials, with feedback about correct target location. On the first cycle of 12

trials, subjects made reliably more correct predictions (45.2%) than would be expected if they only had knowledge of first-order structure (33.3%). Our subjects, therefore, did appear to have access on the generation test to second-order sequence knowledge. Of course, firm conclusions must await additional data.

Cleeremans suggests that "there is no evidence whatsoever that subjects have conscious access to . . . distributional information about the stimulus material." We disagree. Even if prediction performance is to some extent contaminated by unconscious processes, the results (including those of Cleeremans & McClelland 1991, but especially those of Perruchet & Amorim 1992) suggest that subjects have extensive conscious access to the relevant distributional information. More important, there is no evidence that they do *not*.

R3.5. Amnesia. Many of the commentators (Cleeremans, Holyoak & Gattis, Kourtzi et al., Packard, Poldrack & Cohen, Reber & Winter, Seger, Squire et al., Willingham) were concerned by the lack of weight we gave to data from memory-impaired patients. The argument for the significance of such data is clearly stated by Poldrack & Cohen. They claim that "it has been well documented that amnesic patients can exhibit normal learning of (motor, perceptual, and cognitive) skills, and can do so in the absence of the ability to recollect the learning experiences." If learned information can project itself onto performance without being available to consciousness, then surely we have *prima facie* evidence of implicit learning.

We acknowledge that such results provide, at first glance, an apparently strong case for an influence of unconscious information. However, we believe that the commentators have failed to give a fair or complete picture of the data from patients with anterograde amnesia. We believe there are many places where their argument needs a substantial amount of extra work, and in the present section we enumerate various reasons why amnesic data may be explained without appeal to implicit learning.

As a preliminary consideration, it is important to note that there are an enormous number of theories of amnesia, and how they bear on the issue of unconscious learning varies. The view that what is impaired in amnesia is the ability to have conscious declarative memory of prior learning episodes is only one of many theories, and it is by no means clear that this theory is correct (see Mayes 1988). If amnesia in fact turns out to be a selective deficit of contextual processing (Mayes 1988), or of configural learning (Kourtzi et al.; Rudy & Sutherland 1992), or of spatial learning (Hirst & Volpe 1984), or if it turns out to be attributable to any of the huge number of other factors that have been postulated, then its relevance to the issue of unconscious learning will be very limited.

Another reason why data from amnesics may not provide additional illumination concerning implicit learning is that the pattern of data seen in amnesics can readily be obtained with normal subjects anyway. As mentioned in the target article, memory-intact people can be induced to perform at above-chance levels on indirect tests of memory such as preference tests, while performing much more poorly (and perhaps even at chance) on recognition tests (e.g., Kunst-Wilson & Zajonc 1980). As explained in

the target article and in section R3.1, we see little reason to regard such results as evidence of unconscious learning. Poor recognition memory may occur simply because subjects discount a source of conscious information. But perhaps evidence from amnesia is more compelling?

It may be helpful to consider a specific example that is typical of the behavior of amnesics. Knowlton et al. (1992) trained normal and amnesic subjects in the typical manner of an artificial grammar study, and found that they performed at approximately equivalent levels in classifying test strings as grammatical or nongrammatical (normals = 66.9% correct, amnesics = 63.2% correct). However, on a recognition test in which old grammatical and new nongrammatical strings were presented, the amnesics (62.0% correct) performed significantly worse than the normal subjects (72.2% correct). Thus, on the test of conscious knowledge of which strings were study items, the amnesics were impaired, but this had no significant effect on their ability to use those items to guide classification.

Cleeremans, Holyoak & Gattis, Kourtzi et al., Packard, Poldrack & Cohen, Reber & Winter, Seger, Squire et al., and Willingham all believe that results such as this provide evidence of implicit learning. Yet the logic of the argument for this conclusion is rarely spelt out (though see Reingold & Merikle 1988 for an excellent discussion). Presumably, when fleshed out it would run in one of two ways, depending on the exact outcome of the recognition test:

1. If the amnesics are at chance on the recognition test but perform above chance on the implicit classification test, that means there is information that is available to the implicit but not to the explicit test. Subjects will be strongly motivated to use all available *conscious* information when performing the explicit recognition test, hence, the information that is projected onto the implicit but not the explicit test must be *unconscious*. And if that information is unconscious at the time of retrieval, then it is plausible to assume that it was registered unconsciously.

2. If the recognition test reveals above-chance performance, then the argument is slightly more complex and presumably runs as follows. In Knowlton et al.'s (1992) experiment, for example, it is no longer possible to say straightaway that the amnesics had access to a source of information on the implicit test that was not available to them on the explicit test, because recognition and classification are measured in quite different ways. What we can say is that the normal subjects have access to more information on the recognition test than do the amnesics. Because this information is assumed to be conscious, the only way the normals and amnesics could perform equally well on the *implicit* test is if the test is unaffected by conscious information (because normal subjects have more conscious information but they do not perform any better). Therefore above-chance performance on the implicit test must be attributable to *unconscious* information, and the argument proceeds as before.

With respect to the sorts of implicit learning experiments that have been conducted with amnesics, it is the second of the above arguments that is relevant rather than the first. In these experiments (e.g., Knowlton et al. 1992), explicit and implicit memory are assessed in quite different ways using different response scales, and recog-

nition is typically above chance. We know of no clear cases where identical measures are used (e.g., forced-choice recognition versus classification), and in any case our reservations about the mere exposure effect in normals would apply equally to such cases.

Thus we must evaluate the second of the above arguments. There are many places in which it could be challenged, but for present purposes it is vital to note that the line of reasoning only works if performance is truly equivalent in the amnesics and normals on the implicit test. If the normal subjects perform even only marginally better on that test, the argument will not go through, because it is possible that the extra conscious knowledge the normal subjects have (as indexed by superior performance on the explicit test) is what explains their superior performance on the implicit test.¹ We would then be left with no grounds for concluding that unconscious knowledge is playing *any* role at all.

We will argue that the data do not in fact provide strong support for the argument, because, contrary to what is usually claimed, implicit tests of knowledge typically *do* reveal impaired learning in amnesics (in situations where their explicit knowledge is poor). It is simply not true that amnesics perform normally on implicit learning tasks for which their explicit memory is impaired: at most, what can be said is that on some implicit tasks we do not know for certain whether they are impaired or not, whereas on others they are definitely impaired.

As the data cited above show, Knowlton et al.'s (1992) amnesic subjects performed worse (albeit nonsignificantly) than their control subjects, and this is typical of implicit learning studies with amnesics. For example, Kolodny (1994) found that amnesics were worse than normal subjects at learning to classify unfamiliar paintings. Squire and Frambach (1990) found that although amnesics were able to learn to control a complex task, their performance was worse than that of normals. Knowlton and Squire (1993) found that amnesics could correctly categorize the unseen prototype pattern of random dot stimuli, though at a numerically but nonsignificantly lower level of accuracy than could normals. Nissen and Bullemer (1987) observed that although amnesics could improve their performance in a sequential reaction time task, they learned less than normals, as indexed by the fact that their RTs increased by a smaller amount when they were transferred to a random sequence.

Similarly, Charness et al. (1988) found that amnesics were impaired in learning patterns in the Hebb digits task. Amnesics such as H. M. were impaired at learning puzzles such as the Tower of Hanoi (Gabrieli et al. 1987). Finally, despite repeated claims that conditioning occurs at normal rates in amnesics, there is clear evidence of impairments in conditional discrimination learning (Daum et al. 1991) and simple operant conditioning (Channon et al. 1993), if not in simple conditioning.

Note that we are not arguing that amnesics *always* show impaired learning. Rather, we are claiming that in those cases where their *explicit* memory for some information is impaired, their *implicit* memory for that information will be so as well. Thus the finding, cited by Squire et al., that amnesics show normal speed-up in learning to read mirror-reversed text is irrelevant unless it can be shown that they have poorer explicit memory for the knowledge responsible for their improved performance. That their

memory for the *content* of the text they have read may be poorer than that of normals is not sufficient to establish this fact. Reading mirror-reversed text may be based on something totally unrelated to text memory, such as knowledge of single mappings from reversed to normal letters.

Defenders of the view that amnesics show normal implicit, but not explicit, learning would almost certainly reply that, even granting the above cases, there are numerous other implicit tasks, such as repetition priming, in which the performance of amnesics is unquestionably normal (although they have impaired explicit memory), and that these cases provide incontrovertible evidence of implicit learning. For example, Squire et al. assert that priming is intact in amnesia. Again we have to disagree. The main problem is that amnesics almost invariably perform at lower baseline levels than normals, in which case it becomes very difficult to determine what counts as "normal" learning. Numerous studies have shown that, in absolute terms, the amount of repetition priming observed in amnesics and normals is equivalent, but Ostergaard (1994) pointed out that these amounts of priming occur against a background of different overall baselines. Because priming in normals is correlated with baseline latency, and because amnesics have a slower overall baseline, they need to show a larger absolute priming effect to achieve the same relative effect as normals. Amnesics show the same amount of priming in absolute terms, however, despite their slower baseline; hence, it follows that amnesics are *impaired*, not normal, on these tasks.

Part of the difficulty here is that to claim that amnesics learn normally in some tasks is to accept the null hypothesis that a difference does not exist (e.g., Knowlton et al.'s 1992 experiment), but often it is questionable whether the experiment had sufficient statistical power to reject the null hypothesis. Of course, one simple way around the problem would be to show that amnesics can perform *better* than normals on some implicit task (e.g., as a result of having been given more study trials or a shorter retention interval), coupled with *poorer* explicit memory. Our position seems to gain support from the fact that there appear to be no published cases of this sort. Granted, there are some studies in which amnesics have been tested after a shorter retention interval than normals such that performance on the implicit memory test is equated, and where the amnesics still appear to show a selective impairment of episodic memory (e.g., Schacter et al. 1984), but these results have proven very difficult to interpret, because the difference in performance on the implicit and explicit tests may be a function of retention interval (see Mayes 1992). A crossover interaction would be much more persuasive.

We believe that impaired learning in amnesia, as indexed by implicit tests, is the norm rather than the exception. That being the case, the argument that amnesic data demonstrate implicit learning comes completely unglued, because it is quite possible that the pattern of impairment seen in amnesia can be interpreted in terms of a single functional knowledge system: it is no longer necessary to assume two distinct forms of knowledge (conscious and unconscious) contributing to task performance.

By analogy to Farah et al.'s (1993) model, we assume that knowledge is represented in the weights of a parallel-

distributed memory system which can be probed both implicitly and explicitly. When lesioned, such a system would have sufficient knowledge to perform fairly well on an implicit test that can benefit from small and fragmentary amounts of knowledge, but the retained knowledge may not be sufficient to allow for the complete trace reintegration that is typically required in an explicit test. This is tantamount to saying that the explicit tests used to evaluate conscious knowledge in amnesics are generally less sensitive to small amounts of residual information in memory than the implicit tests used to measure performance.

Whatever the merits of our argument, it may well be objected that we have provided no direct evidence in favor of it. We have therefore carried out some computer simulations to try to reproduce Knowlton et al.'s (1992) pattern of data in a single system. The first author, in collaboration with N. Hursey, set up a standard back-propagation network to classify strings as grammatical or nongrammatical. The network was presented with grammatical and nongrammatical strings in the training phase and learned to produce the correct classification decisions across one set of output units. In addition to being taught to classify the strings, the network was simultaneously trained to reproduce the input string across a second set of output units, as in an "encoder" network. After "lesioning" (in fact involving the addition of random noise [mean = 0.0] to the weights) the network's classification and recognition performance was evaluated. For recognition, we measured the accuracy with which each input string was reproduced across the relevant output units.

The critical result was that classification accuracy was only mildly affected by the "lesion" but the network's ability to recognize the training patterns correctly was severely impaired, reproducing the exact pattern that Knowlton et al. (1992) observed in their human subjects. Such a result should not be surprising: it merely extends to amnesia the results reported by Farah et al. (1993; see sect. R2), and confirms the claim that recognition judgments are less sensitive to partial information than are grammaticality judgments. Our simulation also illustrates why it is so important that amnesics are impaired, even if only marginally, on implicit tests: it is because a connectionist network would have great difficulty modelling a case in which performance on only one type of test is impaired.

On the hypothesis that there is a single source of conscious information that is used on implicit and on explicit tests, amnesics would be predicted to perform worse on both; and this is exactly the pattern of results that is observed. In sum, we believe that much of the data from amnesic subjects performing implicit learning tasks of the sort reviewed in the target article can be interpreted by means of a single functional system, with some memory tests simply being more sensitive than others to small amounts of fragmentary information in the system. It is certainly possible that the behavior of amnesics is relevant to the question of unconscious learning, but we believe that advocates of that view will have to do a great deal of work to establish it.

R3.6. Language learning. Language acquisition is frequently cited as a prime example of unconscious rule learning (e.g., see commentaries by Lachter, Nagata,

Reber & Winter). Everything from spelling to grammar appears to be rule governed, and even though native speakers perform well on implicit tests of linguistic knowledge (e.g., grammaticality judgments), they are rarely able to report the relevant rules. Such a result is evidence for implicit learning provided the Information Criterion is met: as long as the rules identified by linguists are indeed the basis of peoples' implicit linguistic competence, their inability to report them suggests that the rules are represented in an unconscious form.

A number of researchers, however, have begun to challenge the idea that linguistic knowledge is rule based; they have shown instead that aspects of language may be learned by distributed instance memory. For example, Rumelhart and McClelland (1986) and Plunkett and Marchman (1991) describe connectionist networks that can learn to perform the mapping of English verbs from their root forms to their past tense forms. Sejnowski and Rosenberg (1987) and Seidenberg and McClelland (1989) constructed networks that learn to transform written words into spoken words, and Elman (1990) and St. John and McClelland (1990) showed that networks can learn simple aspects of English grammar and semantics (these are representative examples among a host of others). All of these network models learn by repeated exposure to instances. Over time, regularities in the data are incorporated into the weights. This instance-based learning procedure is the same procedure that was used by Dienes (1992) to model artificial grammar learning and by Cleeremans and McClelland (1991) to model serial reaction-time learning.

Lachter questions what children could be aware of that would allow them to make grammaticality judgments about intricately structured sentences. One might ask whether there are any distributed processing network models, such as those described in the target article, that can perform sophisticated grammaticality judgments. The current answer is no, but some network models are beginning to approach this level of sophisticated sentence processing. Elman (1993) described a network that accurately processes center-embedded sentences, and St. John (1994) found that a network can accurately process sentences with subject- or object-relative clauses. These networks could be modified to produce actual grammaticality judgments for artificial language strings.

Are these models focusing on the important structural organization of the languages on which they are trained? In analyses of the hidden unit representations that these networks develop, Elman (1993) has found evidence that they learn aspects of the phrase structure of the language. Given that these networks are developing a sensitivity to linguistic structure, the possibility that future distributed instance/memorization models will attain a sophisticated level of processing and be capable of making grammaticality judgments appears quite promising.

The research into network models of language learning is controversial and far from complete. For example, Pinker (1991) and Lachter and Bever (1988) argue that the detailed phenomena of the regular verb past tense require a rule system (see Goldstone & Kruschke's commentary). What we can conclude is that it has not been shown beyond doubt that linguistic knowledge is in the form of rules. That being the case, peoples' inability to articulate such rules is neither here nor there. Time, and

further research, will tell what informational processes underlie children's linguistic behavior and how far network models can be taken to model aspects of language learning without appealing to unconscious rule-learning systems.

R4. How should implicit learning be characterized?

Which dimensions make for sensible distinctions between learning systems? A common answer has been to use the dimension of consciousness, but we have not found a distinction between conscious and unconscious learning to be warranted. We also discussed in section R2.4 of this response the possibility that a distinction between verbally reportable and unreportable knowledge violates the Sensitivity and Information Criteria. Instead, we propose that the distinction between verbally reportable and unreportable knowledge is a useful, though fallible, indicator of a real distinction between rule learning and instance learning.

Several authors take issue with the rule/instance distinction or offer alternative ones. Nagata suggests that our view of what is learned in artificial grammar experiments does not differ markedly from Reber's. Whether or not this is the case, the distinction between rules and instances remains a valid one. Just because artificial grammar experiments may not distinguish between rule and instance learning, it does not follow that other experiments, such as that of Regehr and Brooks (1993), do not.

Holyoak & Gattis suggest the additional criterion of poor transferability and flexibility that, together with instance-based knowledge and limited reportability, seem to characterise what goes by the name "implicit learning." This matrix of characteristics seems to fit with and elaborate on the fragment/instance memorization system. Somewhat similarly, Kourtzi et al. propose a distinction between learning that depends on stimulus recoding and learning of stimulus-response associations. Direct S-R learning is assumed to be less flexible than stimulus recoding. We are not sure whether these two kinds of learning are meant to look different in on-line protocols or are meant to be subject to different independent variables, but in general we welcome the further exploration of these alternative/additional distinctions and their comparison to the instance memorization/hypothesis testing distinction we proposed.

Merikle says that the attempt to (dis)prove any given dissociation is doomed; simple dichotomies never seem to get settled (cf., Newell's 1973 "You can't play 20 questions with nature and win"). We feel that this position is too strong, because researchers have gathered a huge amount of data and learned a great deal about human learning systems from considering simple dichotomies.

R4.1. Consciousness and models of implicit learning.

Our speculations about how to model awareness in a PDP model are criticised by Cleeremans because such models are not aware of the abstraction process across items. Thus, assuming that the abstraction process in humans is analogous to the learning processes in a PDP network, this would appear to constitute a straightforward case of learning without awareness. To be absolutely clear, we are not proposing that networks are conscious. Rather, we are

assuming that they provide *models* of human learning and consciousness.

We agree wholeheartedly with Johnson-Laird's (1983, p. 449) point that "like the weather, [consciousness] is explicable by theories that can be simulated by computer programs, but it can no more be embodied within a computer than can an actual anti-cyclone: only organisms with brains can be conscious." For reasons that no one understands, some of the operations of the parallel network that is the brain happen to be accompanied by awareness of the content of what is being processed. We suggest that artificial networks are capable of modelling, albeit in a rudimentary way, the computational processes of the brain, but that these processes are only accompanied by consciousness in real brains. Nonetheless, we can quite happily hypothesise that it is a certain type of processing operation (in brains or artificial networks) that correlates with conscious experience.

We strongly disagree with Cleeremans's assertion that "the very fact that obviously nonconscious and elementary networks are able to account for human performance . . . has always struck me as evidence that awareness is in fact ancillary in these tasks." It would be bizarre to conclude that because a computer can simulate the flight of a spaceship, therefore rocket fuel is unnecessary for reaching the moon. For human learning, awareness may be as necessary as rocket fuel is for space travel.

R4.2. The fragment/instance memorizer. We agree with a number of commentators that rather than arguing about consciousness, we ought to examine and characterise the real differences between implicit and explicit learning. It was in this spirit that we took up the question of what is truly going on in implicit learning, namely, instance and fragment memorization.

We believe that the available data are best explained in terms of learning rules versus instances, and we are impressed by the data Marsolek cites in favour of distinct anatomical correlates of instance and rule learning. The next big question is, what does the instance/fragment memorization strategy look like, and how does it differ from the rule-learning strategy? Cleeremans says we need a process model, and we agree. But we feel that a step along the way toward that goal is to examine the sorts of information this system uses: what are instances and fragments, and is this information sufficient to model subjects' levels of performance? We reported several studies (Cleeremans & McClelland 1991; Dienes 1992; Perruchet & Pacteau 1990; Reber & Allen 1978; Servan-Schreiber & Anderson 1990) that address this question.

Still, many commentators wonder what is learned in implicit learning. What is this instance/fragment information like? Is it an explicitly represented list of instances, or is the representation less explicit, less individualised? And how is it acquired? We have suggested that knowledge of instances or fragments is represented across a large number of weights in a parallel processor. This distributed processor system is well characterized by Dienes's and Cleeremans & McClelland's connectionist models – as we point out in the target article. Still, it seems from Cleeremans's comments as well as those of others (Seger, Perruchet & Gallego) that we need further exploration of how these distributed processors work and how we can characterise their knowledge.

Whether or not subjects are aware of the abstraction *process*, they are plainly aware of the resulting knowledge of fragments and fragment frequencies (Perruchet & Pacteau 1990; Reber & Allen 1978), and likely continuations of strings and sequences (Dienes et al. 1991; Dulany et al. 1984). This information is available and even reportable at test time.

If we accept the equation of consciousness with activation, then subjects can be aware of the *level* of activation due to different frequencies and variabilities of fragments during training. More frequent fragments will be encoded more often, their weights will grow stronger and produce higher activation levels. As a result, these fragments will seem familiar, both during training and at test time. When new letters violate a common fragment or a light violates a common sequence, the subject may "feel" the violation as a mismatch between past predictions and current encodings via a general reduction in the activation of the network. Either way, the activations of the network provide information that the subject may report as knowledge of the items. With learning, the weights increase and therefore the activations increase. Should we call this change an increase in awareness or an increase in certainty of knowledge? We think the overwhelming evidence falls on the side of interpreting this increase as a change in certainty and quantity of knowledge.

We agree with Cleeremans that it is hard to imagine how knowledge encoded in these weights could be inspected directly. But, as we tried to point out above, we do not think direct inspection is necessary for the information in the weights to be consciously available.

Such knowledge does not require an explicit and individuated encoding of each stimulus. Rather, we point to Cleeremans's commentary as a good description of how a distributed processor acquires the statistical regularities of a corpus of stimuli and encodes that information in a distributed format. Contrary to Kimmel's and Seger's comments, distributed processors can perform a good deal of abstraction from a corpus of instances. Maybe the term "instance memorization" is not really adequate to convey this automatic form of statistical abstraction. Perhaps "distributed instance/fragment encoder" would better designate the process we wish to consider.

R4.3. The relationship between memorization and rule induction. A claim is made by Reber & Winter that our account fails to come to grips with the requirement that the memorial representation acquired in an artificial grammar learning experiment must be of a form that permits generalization and transfer to novel stimulus items with different physical instantiations. We do not disagree about the facts; even though we argued in the target article that much of the data from such studies can be understood simply in terms of instance or fragment memory, we did conclude in sections R2.5.3 and R3.2 that transfer to novel items with different physical instantiations is indeed possible (e.g., Altmann et al., in press).

However, contrary to what Reber & Winter argue, we maintain that this presents no great difficulty for our conception of implicit learning. In addition to memorizing fragments and instances of the study strings, subjects also engage in a certain amount of hypothesis testing, from which they induce abstract rules such as: "the first two symbols of a string cannot be the same." Such rules

allow them to perform slightly better than chance on a grammaticality test with entirely different items such as melodies. However, when tested on items in the same format as the training items, both the induced rules and the memorized fragments can contribute to the grammaticality judgement process, and hence performance is better than in the changed-format condition.

Packard suggests that the notion of a "rule" is not useful for characterizing types of learning, because a rule is merely a guide or principle and "almost any task can be construed to involve rule learning." Of course, as with the terms "learning" and "consciousness," a great deal of ambiguity is possible concerning what one means by the term "rule," but we had a specific concept in mind in the target article. A subject is responding on the basis of a rule if performance to test items is uninfluenced by similarity (e.g., to particular training items). Studies such as that of Nosofsky et al. (1989) illustrate how formal rule models can be constructed and how predictions based on instance memorization and rule use can differ at the quantitative level.

The relationship between distributed instance models that can perform some abstraction on the one hand and rule-learning systems on the other is a complex one. Goldstone & Kruschke and Howe & Rabinowitz question whether it is possible, given our present state of knowledge, to conclude that memorization and rule induction reflect the operations of distinct *systems*. We agree that conclusions should be tentative. We also agree with Goldstone & Kruschke, Nagata, and Poldrack & Cohen that there is enormous debate over what is an instance, what is a rule, the extent to which they interact, and what constitutes a model of either. But we doubt that memorization and rule induction can be seen as the endpoints of a continuum. Similarity either has a detectable effect or it does not: this is not a continuum.

Our concern is to explain the differences in learning and test behavior in experiments such as that of Reber et al. (1980) where explicitly instructed subjects performed quite differently from observation instructed subjects. Though we are not aware of any studies in which on-line protocols were taken from subjects in these conditions, our prediction is that the two groups of subjects would look radically different. A hypothesis-testing strategy is quite different from the distributed encoding of instances. It looks different in on-line protocols, as judged from studies of hypothesis testing (Bruner et al. 1956; Klahr et al. 1990), and a different set of variables affect behaviour (Mathews et al. 1989; Reber et al. 1980).

For example, Reber et al. (1980) found that the order of presentation of the stimuli only affected test performance for the verbally instructed subjects. Of course, as Goldstone & Kruschke might agree, it is likely that even these verbally instructed subjects, who were presumably following a hypothesis-testing strategy, would show some influences of particular instances in their behavior, but we would argue that these result from the influence of a separate instance-encoding system.

Some commentators (Goldstone & Kruschke, Poldrack & Cohen) cited Barsalou's (1990) claim to the effect that abstraction and exemplar theories are logically indistinguishable as evidence of the difficulty of maintaining a contrast between inducing rules and memorizing examples. But although Barsalou's argument is interesting, his

conception of an abstraction is very different from our notion of a rule, so we believe the relevance of his argument is limited. For Barsalou, an abstraction is a representation in which property information is centralised. For us, a rule does not encode information about studied exemplars at all, but rather is a generalisation about regularities in the exemplars.

To conclude, then, we agree that distributed instance models produce a significant amount of abstraction and therefore blur into what we think of as rules and rule-formation models. However, there is strong evidence for a very different strategy of hypothesis testing that looks quite different from memorization.

We fully agree with the commentators who argue that future research should take the direction of finding the differences (and mechanisms) that distinguish implicit from explicit learning. We think one profitable avenue will be to investigate the sorts of factors that will differentially influence a strategy of learning instances from one of testing hypotheses.

NOTES

1. Nagata and Reber & Winter seem to believe that if only a "small" amount of explicit knowledge is available, whereas a "large" amount of implicit knowledge can be shown, then implicit learning has been satisfactorily demonstrated. With respect to amnesia, these commentators would presumably argue that the small benefit the normal subjects show over amnesics in the implicit test cannot explain the large discrepancy on the explicit test, in which case argument 2 is still viable. Because the normal subjects have far more explicit knowledge than the amnesics, it follows that if the two tests relied on a common source of knowledge, the normals would have to perform considerably better than the amnesics on the implicit test. But their excess explicit knowledge is unable to boost performance very much on the implicit test. Hence the latter is principally sensitive to unconscious information. We believe this view may be quite widely held, but it is very important to emphasize that it is entirely fallacious. The problem is that a difference of $x\%$ correct on an implicit measure is as comparable with a difference of $y\%$ correct on an explicit measure as are apples with oranges. Just because we may use a similar measure (e.g., percent correct) in each case does *not* mean that they can be directly compared, because the tests will differ in all sorts of other ways. For example, the chance level of performance is often different in the two tests. It is in general not possible to translate a difference on a given response measure into information, and hence inferences cannot be drawn about whether a subject has "more" or "less" explicit than implicit knowledge. For the same reason, Reber & Winter's claim that implicit learning is less affected than explicit memory by such factors as individual differences and IQ is completely unproven. In the Reber et al. (1991) study, for example, IQ correlated less well with performance on an implicit grammaticality judgment task than with performance on a quite different series continuation task assumed to require explicit processing. But there is no reason to believe that the response scales for these two tasks are comparable or that the difference between them is attributable to the implicit-explicit dimension.

References

- Letters *a* and *r* appearing before authors' initials refer to target article and response respectively.
- Aaronson, D. & Scarborough, H. S. (1977) Performance theories for sentence coding: Some quantitative models. *Journal of Verbal Learning and Verbal Behavior* 16:277-303. [ASR]

- Abrams, M. & Reber, A. S. (1988) Implicit learning: Robustness in the face of psychiatric disorders. *Journal of Psycholinguistic Research* 17:425–39. [ASR, CAS]
- Adams, J. K. (1957) Laboratory studies of behavior without awareness. *Psychological Bulletin* 54:383–405. [aDRS]
- Aha, D. W., Kibler, D. & Albert, M. K. (1991) Instance-based learning algorithms. *Machine Learning* 6:37–66. [RLG]
- Allen, S. W. & Brooks, L. R. (1991) Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General* 120:3–19. [aDRS, RLC]
- Altmann, G. T. M., Dienes, Z. & Goode, A. (in press) On the modality-independence of implicitly learned grammatical knowledge: Implications for the onset of language acquisition. *Cognition*. [aDRS, DCB, ASR]
- Andrade, J. (in press) Learning during anaesthesia. *British Journal of Psychology*. [aDRS]
- Andrade, J., Munglani, R., Jones, J. G. & Baddeley, A. D. (in press) Cognitive performance during anaesthesia. *Consciousness and Cognition*. [JA]
- Baer, P. E. & Fuhrer, M. J. (1982) Cognitive factors in the concurrent differential conditioning of eyelid and skin conductance responses. *Memory and Cognition* 10:135–40. [aDRS]
- Baeyens, F., Eelen, P., Crombez, G. & Van den Bergh, O. (1992) Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy* 30:133–42. [aDRS, FB]
- Baeyens, F., Eelen, P. & Van den Bergh, O. (1990a) Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition and Emotion* 4:3–18. [aDRS, FB]
- Baeyens, F., Eelen, P., Van den Bergh, O. & Crombez, G. (1989) The influence of CS-US perceptual similarity/dissimilarity on human evaluative learning and signal learning. *Learning and Motivation* 20:322–33. [FB]
- (1990b) Flavor-flavor and color-flavor conditioning in humans. *Learning and Motivation* 21:434–55. [FB]
- Baeyens, F., Heremans, R., Eelen, P. & Crombez, G. (1994) Hidden-covariation detection and imagery ability. *European Journal of Cognitive Psychology* 5:435–56. [FB]
- Barsalou, L. W. (1990) On the indistinguishability of exemplar memory and abstraction in category representation. In: *Advances in social cognition* (vol. 3): *Content and process specificity in the effects of prior experiences*, ed. T. K. Srull & R. S. Wyer. [RLG, RAP]
- Barton, L. E., Brulle, A. R. & Repp, A. C. (1986) Maintenance of therapeutic change by momentary DRO. *Journal of Applied Behavior Analysis* 19:277–82. [GO]
- Beeman, M. (1993) Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain and Language* 44:80–120. [CJM]
- Bentall, R. P., Lowe, C. F. & Beasty, A. (1985) The role of verbal behavior in human learning: II. Developmental differences. *Journal of the Experimental Analysis of Behavior* 43:165–81. [GO]
- Berry, D. C. (1994) Implicit learning: Twenty-five years on. A tutorial. In: *Attention and performance XV*, ed. C. Umiltà & M. Moscovitch. MIT Press. [DCB]
- Berry, D. C. & Broadbent, D. E. (1984) On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology* 36A:209–31. [aDRS, DCB, AR, CAS]
- (1987) Explanation and verbalization in a computer-assisted search task. *Quarterly Journal of Experimental Psychology* 39A:585–609. [aDRS]
- (1988) Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology* 79:251–72. [CAS]
- (1990) The role of instruction and verbalization in improving performance on complex search tasks. *Behavior and Information Technology* 9:175–90. [aDRS]
- Berry, D. C. & Dienes, Z. (1991) The relationship between implicit memory and implicit learning. *British Journal of Psychology* 82:359–73. [aDRS]
- (1993) *Implicit learning: Theoretical and empirical issues*. Erlbaum. [DCB, ZD]
- Biederman, I. & Shiffrar, M. M. (1987) Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:640–45. [aDRS]
- Birdwhistell, R. L. (1970) *Kinesics and context*. University of Pennsylvania Press. [ASR]
- Boakes, R. A. (1989) How one might find evidence for conditioning in adult humans. In: *Aversion, avoidance and anxiety: Perspectives on learning and memory*, ed. T. Archer & L.-G. Nilsson. Erlbaum. [aDRS]
- Bornstein, R. F. (1989a) Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin* 106:265–89. [RFB]
- (1989b) Subliminal techniques as propaganda tools: Review and critique. *Journal of Mind and Behavior* 10:231–62. [RFB]
- (1992) Subliminal mere exposure effects. In: *Perception without awareness: Cognitive, clinical and social perspectives*, ed. R. F. Bornstein & T. S. Pittman. Guilford Press. [aDRS, RFB]
- Bornstein, R. F., Leone, D. R. & Galley, D. J. (1987) The generalizability of subliminal mere exposure effects: Influence of stimuli perceived without awareness on social behavior. *Journal of Personality and Social Psychology* 53:1070–79. [RFB]
- Bowers, J. S. & Schacter, D. L. (1990) Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:404–16. [CJM]
- Bowers, K. S. (1984) On being unconsciously influenced and informed. In: *The unconscious reconsidered*, ed. K. S. Bowers & D. Meichenbaum. Wiley. [aDRS]
- Brewer, W. F. (1974) There is no convincing evidence for operant or classical conditioning in adult humans. In: *Cognition and the symbolic processes*, ed. W. B. Weimer & D. S. Palermo. Erlbaum. [aDRS, SSR, FS]
- British Psychological Society (1992) *Subliminal messages*. [aDRS]
- Broadbent, D. E. (1971) *Decision and stress*. Academic Press. [aDRS]
- Broadbent, D. E., Fitzgerald, P. & Broadbent, M. H. P. (1986) Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology* 77:33–50. [aDRS]
- Brody, N. (1989) Unconscious learning of rules: Comment on Reber's analysis of implicit learning. *Journal of Experimental Psychology: General* 118:236–38. [aDRS]
- Brooks, L. (1978) Nonanalytic concept formation and memory for instances. In: *Cognition and categorization*, ed. E. Rosch & B. B. Lloyd. Erlbaum. [aDRS]
- Brooks, L. R. & Vokey, J. R. (1991) Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General* 120:316–23. [aDRS, HN, ASR]
- Bruner, J. S., Goodnow, J. J. & Austin, G. A. (1956) *A study of thinking*. Wiley. [aDRS]
- Burgess, C. & Simpson, G. (1988) Hemispheric processing of ambiguous words. *Brain and Language* 33:86–104. [CJM]
- Cambiganu, C., Rizzo, A., Parlangei, O., Farfaneti, G. & Bagnara, S. (1993) Implicit and explicit learning in controlling dynamic systems. A situated knowledge view. In: *Proceedings of the fourth European conference on cognitive science approaches to process control*, Copenhagen. CRI. [AR]
- Carew, T. J., Hawkins, R. D. & Kandel, E. R. (1983) Differential classical conditioning of a defense withdrawal reflex in *Aplysia californica*. *Science* 219:397–400. [ASR]
- Carlson, R. A. (1992) Starting with consciousness. *American Journal of Psychology* 105:598–604. [RAC]
- Carlson, R. A. & Dulany, D. E. (1985) Conscious attention and abstraction in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11:45–58. [RAC]
- Catania, A. C. (1992) *Learning* (3rd ed.). Prentice-Hall International. [GO]
- Catania, A. C., Matthews, B. A. & Shimoff, E. (1982) Instructed versus shaped human verbal behavior: Interactions with nonverbal responding. *Journal of the Experimental Analysis of Behavior* 38:233–48. [ACC]
- Catania, A. C., Shimoff, E. & Matthews, B. A. (1989) An experimental analysis of rule-governed behavior. In: *Rule-governed behavior: Cognition, contingencies, and instructional control*, ed. S. C. Hayes. Plenum. [aDRS, FS]
- Chan, C. (1992) Implicit cognitive processes: Theoretical issues and applications in computer systems design. Ph.D. thesis, University of Oxford. [ZD]
- Channon, S., Daum, I. & Gray, J. A. (1993) Operant conditioning after temporal lobe lesions in man: Conditional and simple discrimination learning. *Cortex* 29:315–24. [rDRS]
- Charness, N., Milberg, W. & Alexander, M. P. (1988) Teaching an amnesic a complex cognitive skill. *Brain and Cognition* 8:253–72. [rDRS]
- Cheesman, J. & Merikle, P. M. (1984) Priming with and without awareness. *Perception and Psychophysics* 36:387–95. [ZD]
- (1986) Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology* 40:343–67. [MAS]
- Chiarello, C. (1988) Semantic priming in the intact brain: Separate roles for the right and left hemispheres? In: *Right hemisphere contributions to lexical semantics*, ed. C. Chiarello. Springer-Verlag. [CJM]
- Chomsky, N. (1981) *Lectures on government and binding*. Holland: Foris Publications. [HN]
- Cleeremans, A. (1993a) Attention and awareness in sequence learning. In: *Proceedings of the fifteenth annual conference of the cognitive science society*. Erlbaum. [AC]
- (1993b) *Mechanisms of implicit learning: Connectionist models of sequence processing*. MIT Press. [aDRS, AC, ZD, GO, RAP]

References/Shanks & St. John: Dissociable learning systems

- Cleeremans, A. & McClelland, J. L. (1991) Learning the structure of event sequences. *Journal of Experimental Psychology: General* 120:235–53. [aDRS, AC]
- Cohen, A., Ivry, R. I. & Keele, S. W. (1990) Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:17–30. [aDRS; JR]
- Cohen, N. J. & Eichenbaum, H. (1993) *Memory, amnesia, and the hippocampal system*. MIT Press. [RAP]
- Cohen, N. J. & Squire, L. R. (1980) Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science* 21:207–10. [aDRS, EE, ZK, RAP, LRS]
- Cork, R. C., Kihlstrom, J. F. & Schacter, D. L. (1992) Absence of explicit or implicit memory in patients anesthetized with sufentanil/nitrous oxide. *Anesthesiology* 76:892–98. [aDRS]
- Curran, T. & Keele, S. W. (1993) Attentional and non-attentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:189–202. [AC]
- Daum, I., Channon, S., Polkey, C. E. & Gray, J. A. (1991) Classical conditioning after temporal lobe lesions in man: Impairment in conditional discrimination. *Behavioral Neuroscience* 105:396–408. [rDRS]
- Davey, G. C. L. (1984) Is evaluative conditioning a qualitatively distinct form of classical conditioning? *Behavior Research and Therapy* 32:291–99. [rDRS]
- (1992) An expectancy model of laboratory preparedness effects. *Journal of Experimental Psychology: General* 121:24–40. [aDRS]
- Dawson, M. E. & Schell, A. M. (1985) Information processing and human autonomic classical conditioning. In: *Advances in psychophysiology*, ed. P. K. Ackles, J. R. Jennings & M. G. Coles. JAI Press. [aDRS]
- De Houwer, J., Baeyens, F. & Eelen, P. (in press) Verbal evaluative conditioning with undetected US presentations. *Behaviour Research and Therapy*. [FB]
- De Houwer, J., Hendrick, H., Baeyens, F. & van Avermaet, E. (1994) Hidden covariation detection might be very hidden indeed. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. [aDRS]
- Dennett, D. C. (1991) *Consciousness explained*. Little, Brown. [rDRS]
- Dickinson, A. (1988) Intentionality in animal conditioning. In: *Thought without language*, ed. L. Weiskrantz. Clarendon Press. [rDRS]
- Dienes, Z. (1992) Connectionist and memory-array models of artificial grammar learning. *Cognitive Science* 16:41–80. [aDRS, RAP]
- Dienes, Z., Altmann, G. & Kwan, L. (submitted) Unconscious knowledge of artificial grammars is applied strategically. [ZD]
- Dienes, Z., Broadbent, D. E. & Berry, D. (1991) Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17:875–87. [aDRS, HN]
- Dienes, Z. & Perner, J. (forthcoming) Implicit knowledge in people and connectionist networks. In: *Implicit cognition*, ed. G. Underwood. Oxford University Press. [ZD]
- Dreyfus, H. L. & Dreyfus, S. E. (1986) Why skills cannot be represented by rules. In: *Advances in cognitive science*, ed. N. E. Sharkey. Ellis Horwood. [GO]
- Dulany, D. E. (1961) Hypotheses and habits in verbal "operant conditioning." *Journal of Abnormal and Social Psychology* 63:251–63. [arDRS, RAC, FS]
- (1968) Awareness, rules, and propositional control: A confrontation with S-R behavior theory. In: *Verbal behavior and general behavior theory*, ed. T. Dixon & D. Horton. Prentice-Hall. [RAC, SSR]
- Dulany, D. E., Carlson, R. A. & Dewey, G. I. (1984) A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General* 113:541–55. [aDRS, RAC]
- (1985) On consciousness in syntactic learning and judgment: A reply to Reber, Allen, and Regan. *Journal of Experimental Psychology: General* 114:25–32. [aDRS, RAC]
- Dulany, D. E. & Poldrack, R. A. (1991) Learned covariation: Conscious or unconscious representation? Paper presented at the Annual Meeting of the Psychonomic Society, San Francisco. [aDRS]
- Eich, E. (1984) Memory for unattended events: Remembering with and without awareness. *Memory and Cognition* 12:105–11. [aDRS]
- Eichenbaum, H., Fagan, A., Mathews, P. & Cohen, N. (1988) Hippocampal system dysfunction and odor discrimination learning in rats: Impairment or facilitation depending on representational demands. *Behavioral Neuroscience* 102:331–39. [LRS]
- Eichenbaum, H., Otto, T. & Cohen, N. J. (1992) The hippocampus – what does it do? *Behavior, Neurology, Biology* 57:2–36. [LRS]
- Elman, J. L. (1990) Finding structure in time. *Cognitive Science* 14:179–212. [rDRS, AC]
- (1993) Learning and development in neural networks: The importance of starting small. *Cognition* 48:71–99. [rDRS]
- Erdelyi, M. H. (1985) *Psychoanalysis: Freud's cognitive psychology*. Freeman. [RFB]
- Eriksen, K. A. & Simon, H. A. (1980) Verbal reports as data. *Psychological Review* 87:215–51. [aDRS]
- (1984) *Protocol analysis: Verbal reports as data*. MIT Press. [aDRS]
- Eriksen, C. W. (1960) Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review* 67:279–300. [aDRS]
- Estes, W. K. (1986) Array models for category learning. *Cognitive Psychology* 18:500–549. [aDRS]
- Farah, M. J., O'Reilly, R. C. & Vecera, S. P. (1993) Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review* 100:571–88. [rDRS]
- Fleishman, E. A. & Quaintance, M. K. (1984) *Taxonomies of human performance*. Academic Press. [DBW]
- Flexner, A. J. & Tulving, E. (1978) Retrieval independence in recognition and recall. *Psychological Review* 85:153–71. [SSR]
- Frensch, P. A., Buchner, A. & Lin, J. (1994) Implicit learning of unique and ambiguous serial transitions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:567–84. [CAS]
- Gabrieli, J. D. E., Keane, M. M. & Corkin, S. (1987) Acquisition of problem-solving skills in global amnesia. *Society for Neuroscience Abstracts* 13:1455. [rDRS]
- Ghoneim, M. M. & Block, R. I. (1992) Learning and consciousness during general anesthesia. *Anesthesiology* 76:279–305. [aDRS]
- Ghoneim, M. M., Block, R. I. & Fowles, D. C. (1992) No evidence of classical conditioning of electrodermal responses during anesthesia. *Anesthesiology* 76:682–88. [aDRS]
- Gibson, J. J. (1971) The information available in pictures. *Leonardo* 4:27–35. [RAC]
- Gluck, M. A. & Bower, G. H. (1988) From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General* 117:227–47. [aDRS]
- Gluck, M. A. & Myers, C. E. (1993) Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus* 3:491–516. [ZK]
- Goodale, M. A. (1993) Visual pathways supporting perception and action in the primate cerebral cortex. *Current Opinion in Neurobiology* 3:578–85. [LRS]
- Gordon, P. C. & Holyoak, K. J. (1983) Implicit learning and generalization of the mere exposure effect. *Journal of Personality and Social Psychology* 45:492–500. [RFB, CAS]
- Graf, P. & Komatsu, S. (in press) Process dissociation procedure: Handle with caution! *European Journal of Cognitive Psychology*. [MAS]
- Graf, P., Squire, L. R. & Mandler, G. (1984) The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10:164–78. [aDRS, ZK]
- Granda, A. M. & Hammack, J. T. (1961) Operant behavior during sleep. *Science* 133:1485–86. [GO]
- Green, R. E. A. & Shanks, D. R. (1993) On the existence of independent explicit and implicit learning systems: An examination of some evidence. *Memory and Cognition* 21:304–17. [aDRS]
- Greenspoon, J. (1955) The reinforcing effect of two spoken sounds on the frequency of two responses. *American Journal of Psychology* 68:409–16. [aDRS]
- Greenwald, A. G. (1992) Unconscious cognition reclaimed. *American Psychologist* 47:766–79. [aDRS]
- Greenwald, A. G., Spangenberg, E. R., Pratkanis, A. R. & Eskenazi, J. (1991) Double blind tests of subliminal self-help audiotapes. *Psychological Science* 2:119–22. [aDRS]
- Griffin, D. R. (1992) *Animal minds*. University of Chicago Press. [rDRS]
- Grings, W. W. (1965) Verbal-perceptual factors in the conditioning of autonomic responses. In: *Classical conditioning: A symposium*, ed. W. F. Prokasy. Appleton-Century-Crofts. [rDRS, HDK]
- Grings, W. W., Schell, A. M. & Carey, C. A. (1973) Verbal control of an autonomic response in a cue reversal situation. *Journal of Experimental Psychology* 99:215–21. [aDRS, ASR]
- Haith, M. M. & McCarty, M. E. (1990) Stability of visual expectation at 3.0 months of age. *Developmental Psychology* 26:68–74. [ASR]
- Hartman, M., Knopman, D. S. & Nissen, M. J. (1989) Implicit learning of new verbal associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:1070–82. [aDRS]
- Hawkins, R. D., Kandel, E. R. & Siegelbaum, S. A. (1993) Learning to modulate transmitter release: Themes and variations in synaptic plasticity. *Annual Review of Neuroscience* 16:625–65. [LRS]
- Hayes, N. A. & Broadbent, D. E. (1988) Two modes of learning for interactive tasks. *Cognition* 28:249–76. [aDRS, FS]
- Hefferline, R. F., Keenan, B. & Harford, R. A. (1959) Escape and avoidance

- conditioning in human subjects without their observation of the response. *Science* 130:1338-39. [FS]
- Hintzman, D. L. (1986) "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93:411-28. [aDRS]
- Hirst, W., Johnson, M. K., Phelps, E. A. & Volpe, B. T. (1988) More on recognition and recall in amnesics. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14:758-62. [CAS]
- Hirst, W. & Volpe, B. T. (1984) Encoding of spatial relations with amnesia. *Neuropsychologia* 22:631-34. [rDRS]
- 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. [aDRS]
- Holyoak, K. J. & Spellman, B. A. (1993) Thinking. *Annual Review of Psychology* 44:265-315. [KJH]
- Homa, D., Sterling, S. & Trepel, L. (1981) Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory* 7:418-39. [aDRS]
- Howard, D. V. & Howard, J. H. (1989) Age differences in learning serial patterns: Direct versus indirect measures. *Psychology and Aging* 4:357-64. [aDRS, CAS]
- Howe, M. L., Rabinowitz, F. M. & Grant, M. J. (1993) On measuring (in)dependence of cognitive processes. *Psychological Review* 100:737-47. [NB, MLH]
- Hugdahl, K. & Ohman, A. (1977) Effects of instruction on acquisition and extinction of electrodermal responses to fear-relevant stimuli. *Journal of Experimental Psychology: Human Learning and Memory* 3:608-18. [aDRS]
- Humphreys, M. S., Bain, J. D. & Pike, R. (1989) Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review* 96:208-33. [aDRS]
- Jacoby, L. L. (1991) A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language* 30:513-41. [aDRS]
- Jacoby, L. L. & Dallas, M. (1981) On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General* 110:306-40. [CAS]
- Jacoby, L. L. & Kelley, C. (1991) Unconscious influences of memory: Dissociations and automaticity. In: *The neuropsychology of consciousness*, ed. D. Milner & M. Rugg. Academic Press. [aDRS]
- Jacoby, L. L., Lindsay, D. S. & Toth, J. P. (1992) Unconscious influences revealed: Attention, awareness, and control. *American Psychologist* 47:802-9. [PMM]
- Jacoby, L. L., Toth, J. P. & Yonelinas, A. P. (1993) Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General* 122:139-54. [rDRS, MLH]
- James, W. (1890) *The principles of psychology*. Holt. [JL]
- Jelicic, M., Bonke, B., Wolters, G. & Phaf, R. H. (1992) Implicit memory for words presented during anaesthesia. *European Journal of Cognitive Psychology* 4:71-80. [aDRS]
- Johnson, J. S. & Newport, E. L. (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21:60-99. [JL]
- Johnson, P. E. (1983) What kind of expert should a system be? *The Journal of Medicine and Philosophy* 8:77-97. [GO]
- Johnson, W. A., Hawley, K. J. & Elliott, J. M. G. (1991) Contribution of perceptual fluency to recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17:210-23. [CAS]
- Johnson-Laird, P. N. (1983) *Mental models*. Cambridge University Press. [rDRS]
- Kemler Nelson, D. G. (1984) The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior* 23:734-59. [aDRS]
- Kihlstrom, J. F., Schacter, D. L., Cork, R. C., Hurt, C. A. & Behr, S. E. (1990) Implicit and explicit memory following surgical anaesthesia. *Psychological Science* 1:303-6. [aDRS]
- Kimmel, E. (1967) Judgments of UCS intensity and diminution of the UCR in classical CSR conditioning. *Journal of Experimental Psychology* 73:532-43. [HDK]
- Kimmel, H. D. (1974) Instrumental conditioning of autonomically mediated behavior in human beings. *American Psychologist* 29:325-35. [HDK]
- Kimmel, H. D. & Bevill, M. (1985) Habituation and dishabituation of the human orienting reflex under instruction-induced stress. *Physiological Psychology* 13:92-94. [HDK]
- Kimmel, H. D., Birbaumer, N., Elbert, T., Lutzenberger, W. & Rockstroh, B. (1983) Conditioned tonic stimulus control of nonspecific arousal. *Pavlovian Journal of Biological Science* 18:136-43. [rDRS, HDK]
- Klahr, D., Dunbar, K. & Fay, A. L. (1990) Designing good experiments to test bad hypotheses. In: *Computational models of scientific discovery and theory formation*, ed. J. Shrager & P. Langley. Morgan Kaufmann. [aDRS]
- Klayman, J. & Ha, Y. (1989) Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:596-604. [aDRS]
- Knopman, D. (1991) Unaware learning versus preserved learning in pharmacologic amnesia: Similarities and differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17:1017-29. [aDRS]
- Knowlton, B. J., Gluck, M. A. & Squire, L. R. (1993) Probabilistic category learning in amnesia. Proceedings of the Society for Neuroscience, Washington, D.C. [ZK]
- Knowlton, B. J., Ramus, S. J. & Squire, L. R. (1992) Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for instances. *Psychological Science* 3:172-79. [aDRS, KJH, ASR, CAS, DBW]
- Knowlton, B. J. & Squire, L. R. (1992) Intact prototype learning by amnesic patients: Evidence for parallel learning of item-specific and general information. *Society for Neuroscience Abstracts* 18:168. [CAS]
- Knowlton, B. J. & Squire, L. R. (1993) The learning of categories: Parallel brain systems for item memory and category knowledge. *Science* 262:1747-49.
- Kolodny, J. (1994) Memory processes in classification learning. *Psychological Science*. [rDRS]
- Kotovsky, K. & Simon, H. A. (1979) Empirical tests of a theory of human acquisition of concepts for sequential patterns. In: *Models of thought*, ed. H. A. Simon. Yale University Press. [aDRS]
- Krasner, L. (1958) Studies of the conditioning of verbal behavior. *Psychological Bulletin* 55:148-70. [aDRS, SSR]
- Krosnick, J. A., Betz, A. I., Jussim, L. J. & Lynn, A. R. (1992) Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin* 18:152-62. [FB]
- Kruschke, J. K. (1992) ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review* 99:22-44. [aDRS]
- Kuhl, P. K. & Metzoff, A. N. (in press) Evolution: Nativism and learning in the development of language and speech. In: *The genetic basis of language*, ed. M. Gopnik. Oxford University Press. [LRS]
- Kunst-Wilson, W. R. & Zajonc, R. B. (1980) Affective discriminations of stimuli that cannot be recognized. *Science* 207:557-58. [arDRS, RFB]
- Kushner, M., Cleeremans, A. & Reber, A. (1991) Implicit detection of event interdependencies and a PDP model of the process. In: *Proceedings of the thirteenth annual meeting of the cognitive science society*. Erlbaum. [CAS]
- Lachnit, H. & Kimmel, H. D. (1993) Positive and negative patterning in human classical skin conductance response conditioning. *Animal Learning and Behavior* 21:314-26. [rDRS, HDK]
- Lachter, J. & Bever, T. G. (1988) The relation between linguistic structure and associative theories of language learning: A constructive critique of some connectionist learning models. *Cognition* 28:195-247. [rDRS, JL]
- Lea & Simon (1979) Problem solving and rule induction. In: *Models of thought*, ed. H. A. Simon. Yale University Press. [aDRS]
- Lewicki, P. (1986) Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12:135-46. [aDRS]
- Lewicki, P., Czyzewska, M. & Hoffinan, H. (1987) Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition* 13:523-30. [aDRS, SSR, CAS, FS]
- Lewicki, P., Hill, T. & Bizot, E. (1988) Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology* 20:24-37. [aDRS, CAS]
- Lewicki, P., Hill, T. & Czyzewska, M. (1992) Nonconscious acquisition of information. *American Psychologist* 47:796-801. [CAS]
- Lewicki, P., Hill, T. & Sasaki, I. (1989) Self-perpetuating development of encoding biases. *Journal of Experimental Psychology: General* 118:323-37. [aDRS]
- Libet, B. (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences* 8:529-66. [rDRS, GO]
- Light, P. & Butterworth, G. (1992) *Context and cognition*. Harvester Wheatsheaf. [AR]
- Loftus, E. F. & Klinger, M. R. (1992) Is the unconscious smart or dumb? *American Psychologist* 47:761-65. [aDRS]
- Logan, G. D. (1988) Toward an instance theory of automatization. *Psychological Review* 95:492-528. [MAS]
- Lovibond, P. F. (1992) Tonic and phasic electrodermal measures of human aversive conditioning with long duration stimuli. *Psychophysiology* 29:621-32. [aDRS]

- (1993) Conditioning and cognitive-behaviour therapy. *Behaviour Change* 10:119–30. [rDRS]
- Lowe, C. F. (1979) Determinants of human operant behavior. In: *Advances in analysis of behavior, vol. 1: Reinforcement and the organization of behavior*, ed. M. D. Zeiler & P. Harzem. Wiley. [FS]
- Lowe, C. F., Beasty, A. & Bentall, R. P. (1983) The role of verbal behavior in human learning: Infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior* 39:157–64. [GO]
- Mace, F. C. & Knight, D. (1986) Functional analysis and treatment of severe pica. *Journal of Applied Behavior Analysis* 19:411–16. [GO]
- MacKay, D. M. (1974) The mechanics of tacit knowing. *IEEE Transactions on Systems, Man, and Cybernetics*. [ROL]
- Mahut, H. & Moss, M. (1984) Consolidation of memory: The hippocampus revisited. In: *Neuropsychology of memory*, ed. L. R. Squire & N. Butters. Guilford Press. [ASR]
- Malt, B. (1989) An on-line investigation of prototype and exemplar strategies in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:539–55. [aDRS]
- Mandler, G., Nakamura, Y. & Van Zandt, B. J. S. (1987) Nonspecific effects of exposure to stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:646–48. [RFB]
- Mangun, B. (1993) Taking phenomenology seriously: The “fringe” and its implications for cognitive research. *Consciousness and Cognition* 2:89–108. [CAS]
- Manza, L. & Reber, A. S. (1992a) Implicit learning: Transfer across form and sensory modality. Paper presented at the meeting of the Psychonomic Society, St. Louis, MO. [CAS]
- (1992b) Inter- and intra-modal transfer of an implicitly acquired rule system. Unpublished manuscript. [ASR]
- Marcel, A. (1983) Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology* 15:197–237. [aDRS]
- (1993) Slippage in the unity of consciousness. In: *Experimental and theoretical studies of consciousness*, ed. G. R. Bock & J. Marsh. Wiley. [rDRS]
- Marinkovic, K., Schell, A. M. & Dawson, M. E. (1989) Awareness of the CS-UCS contingency and classical conditioning of skin conductance responses with olfactory CSs. *Biological Psychology* 29:39–60. [aDRS]
- Marsolek, C. J. (submitted) Abstract-visual-form representations in the left cerebral hemisphere. *Journal of Experimental Psychology: Human Perception and Performance*. [CJM]
- Marsolek, C. J., Kosslyn, S. M. & Squire L. R. (1992) Form-specific visual priming in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:492–508. [CJM]
- Marsolek, C. J. & Schacter, D. L. (submitted) Form-specific visual priming for new associations in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. [CJM]
- Marsolek, C. J., Squire, L. R., Kosslyn, S. M. & Lulenski, M. E. (submitted) Form-specific explicit and implicit memory in the right cerebral hemisphere. *Neuropsychology*. [CJM]
- Martin, I. & Levey, A. B. (1978) Evaluative conditioning. *Advances in Behaviour Research and Therapy* 1:57–102. [aDRS]
- (1987) Learning what will happen next: Conditioning, evaluation and cognitive processes. In: *Cognitive processes and Pavlovian conditioning in humans*, ed. G. Davey. Wiley. [aDRS]
- Masters, R. S. W. (1992) Knowledge, nerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology* 83:343–58. [CAS]
- Mathews, R. (1990) Abstractiveness of implicit grammar knowledge. Comments on Perruchet and Pacteau’s analysis of synthetic grammar learning. *Journal of Experimental Psychology: General* 119:412–16. [DCB]
- Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R. & Druhan, B. (1989) Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:1083–1100. [aDRS, DCB, ASR]
- Matthews, B. A., Shimoff, E., Catania, A. C. & Sagvolden, T. (1977) Uninstructed human responding: Sensitivity to ratio and interval contingencies. *Journal of the Experimental Analysis of Behavior* 27:453–67. [FS]
- Mauk, M. D. & Thompson, R. F. (1987) Retention of classically conditioned eyelid responses following acute decerebration. *Brain Research* 403:89–95. [LRS]
- Mayes, A. R. (1988) *Human organic memory disorders*. Cambridge University Press. [rDRS]
- (1992) What are the functional deficits that underlie amnesia? In: *Neuropsychology of memory*, ed. L. R. Squire & N. Butters. Guilford Press. [rDRS]
- McClelland, J. L. & Rumelhart, D. E. (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review* 88:375–407. [aDRS]
- (1985) Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General* 114:159–88. [aDRS]
- McLaren, I. P. L., Green, R. E. A. & Mackintosh, N. J. (in press) Animal learning and the explicit/implicit distinction. In: *Implicit and explicit learning of languages*, ed. N. Ellis. Academic Press. [rDRS]
- Medin, D. L. (1986) Comment on “Memory storage and retrieval processes in category learning.” *Journal of Experimental Psychology: General* 115:373–81. [RLG]
- Medin, D. L. & Florian, J. E. (1992) Abstraction and selective coding in exemplar-based models of categorization. In: *From learning processes to cognitive processes: Essays in honor of William K. Estes, vol. 2*, ed. A. F. Healy, S. M. Kosslyn & R. M. Shiffrin. Erlbaum. [aDRS]
- Medin, D. L. & Schaffer, M. M. (1978) Context theory of classification learning. *Psychological Review* 85:207–38. [aDRS]
- Merkle, P. M. (1992) Perception without awareness: Critical issues. *American Psychologist* 47:792–95. [ZD, PMM]
- Merkle, P. M. & Cheesman, J. (1986) Consciousness is a “subjective” state. *Behavioral and Brain Sciences* 9:42. [FB]
- Merkle, P. M. & Reingold, E. M. (1990) Recognition and lexical decision without detection: Unconscious perception? *Journal of Experimental Psychology: Human Perception and Performance* 16:574–83. [aDRS]
- (1991) Comparing direct (explicit) and indirect (implicit) measures to study unconscious memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17:224–33. [arDRS, AC]
- (1992) Measuring unconscious perceptual processes. In: *Perception without awareness: Cognitive, clinical, and social perspectives*, ed. R. F. Bornstein & T. S. Pittman. Guilford Press. [aDRS, FB]
- Merkle, P. M. & Rondi, G. (1993) Memory for events during anesthesia has not been demonstrated: A psychologist’s viewpoint. In: *Memory and awareness in anesthesia*, ed. P. S. Sebel, B. Bonke & E. Winograd. Prentice-Hall. [rDRS]
- Mishkin, M. (1982) A memory system in the monkey. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 298:85–92. [ASR]
- Mishkin, M., Malamut, B. & Bachevalier, J. (1984) Memories and habits: Two neural systems. In: *Neurobiology of learning and memory*, ed. J. L. McCaugh, G. Lynch & H. M. Weinberger. Guilford Press. [aDRS, LRS]
- Moerman, N., Bonke, B. & Oosting, J. (1993) Awareness and recall during general anesthesia. *Anesthesiology* 79:454–64. [JA]
- Morris, R. G. M. (1984) Is the distinction between procedural and declarative memory useful with respect to animal models of amnesia? In: *Neurobiology of learning and memory*, ed. J. L. McCaugh, G. Lynch & H. M. Weinberger. Guilford Press. [aDRS]
- Munglani, R., Andrade, J., Sapsford, D. J., Baddeley, A. D. & Jones, J. G. (1993) A measure of consciousness and memory during isoflurane administration: The coherent frequency. *British Journal of Anaesthesia* 71(5):633–41. [JA]
- Musen, G., Shimamura, A. P. & Squire, L. R. (1990) Intact text-specific reading skill in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:1068–76. [aDRS, LRS]
- Musen, G. & Squire, L. R. (1993) Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:789–98. [CAS]
- Nagata, H. (1988) The relativity of linguistic intuition: The effect of repetition on grammaticality judgments. *Journal of Psycholinguistic Research* 17:1–17. [HN]
- (1989a) Effect of repetition on grammaticality judgments under objective and subjective self-awareness conditions. *Journal of Psycholinguistic Research* 18:255–69. [HN]
- (1989b) Judgments of sentence grammaticality with differentiation and enrichment strategies. *Perceptual and Motor Skills* 68:463–69. [HN]
- (1989c) Judgments of sentence grammaticality and field-dependency of subjects. *Perceptual and Motor Skills* 69:739–47. [HN]
- (1992) Anchoring effects in judging grammaticality of sentences. *Perceptual and Motor Skills* 75:159–64. [HN]
- (submitted) On-line and off-line reflexive resolution in Japanese logophoric sentences. [HN]
- Nagel, T. (1974) “What is it like to be a bat?” *Philosophical Review* 83:435–45. [rDRS]
- Natsoulas, T. (1992) Is consciousness what psychologists actually examine? *American Journal of Psychology* 105:363–84. [RAC]
- Neely, J. H. (1989) Experimental dissociations and the episodic/semantic memory distinction. In: *Varieties of memory and consciousness*, ed. H. L. Roediger & F. I. M. Craik. Erlbaum. [aDRS]

- Neisser, U. (1991) Two perceptually given aspects of the self and their development. *Developmental Review* 11:197-209. [RAC]
- Nelson, T. O. (1978) Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Human Learning and Memory* 4:453-68. [aDRS]
- Newell, A. (1973) You can't play twenty questions with nature and win: Projective comments on the papers of this symposium. In: *Visual information processing*, ed. W. Chase. Academic Press. [rDRS]
- Newport, E. L. (1990) Maturational constraints on language learning. *Cognitive Science* 14:11-28. [JL]
- Niedenthal, P. M. (1990) Implicit perception of affective information. *Journal of Experimental Social Psychology* 26:505-27. [FB]
- Nisbett, R. E. & Wilson, T. D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84:231-59. [HST]
- Nissen, M. J. & Bullemer, P. (1987) Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology* 19:1-32. [aDRS, KJH, DBW]
- Nissen, M. J., Knopman, D. S. & Schacter, D. L. (1987) Neurochemical dissociation of memory systems. *Neurology* 37:789-94. [aDRS]
- Nissen, M. J., Willingham, D. & Hartman, M. (1989) Explicit and implicit remembering: When is learning preserved in amnesia? *Neuropsychologia* 27:341-52. [CAS]
- Norman, R. J., Buchwald, J. S. & Villablanca, J. R. (1977) Classical conditioning with auditory discrimination of the eyeblink reflex in decerebrate cats. *Science* 196:551-53. [LRS]
- Nosofsky, R. M. (1986) Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115:39-57. [aDRS]
- Nosofsky, R. M., Clark, S. E. & Shin, H. J. (1989) Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:282-304. [aDRS, RLG]
- Nosofsky, R. M., Palmeri, T. J. & McKinley, S. C. (1994) Rule-plus-exception model of classification learning. *Psychological Review* 101:53-79. [RLG]
- O'Keefe, J. & Nadel, L. (1978) *The hippocampus as a cognitive map*. Oxford University Press. [aDRS]
- Ostergaard, A. L. (1994) Dissociations between word priming effects in normal subjects and patients with memory disorders: Multiple memory systems or retrieval? *Quarterly Journal of Experimental Psychology* 47A: 331-64. [arDRS]
- Overskeid, G. (1992) Kontingensforming og kunnskapsformidling: Et grunnleggende problem ved konstruksjon av ekspertsystemer [Contingency shaping and transfer of knowledge: A basic problem in the construction of expert systems]. *Tidsskrift for Norsk Psykologforening* 29:210-21. [GO]
- (submitted) Behaviorist or cognitivist - who can tell the difference? The case of implicit and explicit knowledge. [GO]
- Packard, M. G., Hirsh, R. & White, N. M. (1989) Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience* 9:1465-72. [MGP, LRS]
- Packard, M. G. & McGaugh, J. L. (1992) Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. *Behavioral Neuroscience* 106:1-7. [MGP]
- Packard, M. G. & White, N. M. (1991) Dissociation of hippocampus and caudate nucleus memory systems by post-training intracerebral injection of dopamine agonists. *Behavioral Neuroscience* 105:295-306. [MGP]
- Parkhouse, J. (1960) Awareness during surgery. *Postgraduate Medical Journal* 36:674-77. [JA]
- Perruchet, P. (1993) Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:223-28. [aDRS]
- (in press) Learning from complex rule-governed environments: On the proper functions of nonconscious and conscious processes. In: *Attention and Performance XV: Conscious and nonconscious information processing*, ed. C. Umiltà & M. Moscovitch. MIT Press. [aDRS, PP]
- Perruchet, P. & Amorim, M.-A. (1992) Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:785-800. [aDRS, DCB, AC]
- Perruchet, P. & Gallego, J. (1994) Association between conscious knowledge and performance in normal subjects: A reply to Cohen and Curran (1993) and to Willingham, Greeley, and Bardone (1993). *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:1438-44. [aDRS]
- Perruchet, P., Gallego, J. & Pacteau, C. (1992) A reinterpretation of some earlier evidence for abstractiveness of implicitly acquired knowledge. *Quarterly Journal of Experimental Psychology* 44A:193-210. [aDRS]
- Perruchet, P., Gallego, J. & Savy, I. (1990) A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology* 22:493-516. [aDRS]
- Perruchet, P. & Pacteau, C. (1990) Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General* 119:264-75. [aDRS, DCB]
- (1991) The implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. *Journal of Experimental Psychology: General* 120:112-16. [PP]
- Pinker, S. (1991) Rules of language. *Science* 253:530-35. [RLG, rDRS]
- Plunkett, K. & Marchman, V. (1991) U-shaped learning and frequency effects in a multi-layer perceptron: Implications for child language acquisition. *Cognition* 38:43-102. [rDRS]
- Polanyi, M. (1966) *The tacit dimension*. Doubleday. [ASR]
- Posner, M. I. & Keele, S. W. (1970) Retention of abstract ideas. *Journal of Experimental Psychology* 83:304-8. [aDRS]
- Rabinowitz, F. M. (1993) Unpublished data. [MLH]
- Rabinowitz, F. M. & Howe, M. L. (1994) Development of the middle concept. *Journal of Experimental Child Psychology* 57. [MLH]
- Rakover, S. (1993) Empirical criteria for task susceptibility to introspective awareness and awareness effects. *Philosophical Psychology* 6:451-67. [SSR]
- Rathus, J., Reber, A. S., Manza, L. & Kushner, M. (in press) Implicit and explicit learning: Differential effects of affective states. *Perceptual and Motor Skills*. [ASR]
- Reber, A. S. (1967) Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 5:855-63. [aDRS, ACC, AC, HN]
- (1969) Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology* 81:115-19. [aDRS, RFB, ACC, ASR, CAS]
- (1976) Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory* 2:88-94. [ACC]
- (1989a) Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* 118:219-35. [aDRS, RFB, HN, MAS]
- (1989b) More thoughts on the unconscious: Reply to Brody and to Lewicki and Hill. *Journal of Experimental Psychology: General* 118:242-44. [HN]
- (1990) On the primacy of the implicit: Comment on Perruchet and Pacteau. *Journal of Experimental Psychology: General* 119:340-42. [DCB]
- (1992a) The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition* 1:93-133. [CAS]
- (1992b) An evolutionary context for the cognitive unconscious. *Philosophical Psychology* 5:33-51. [KJH]
- (1993) *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford University Press. [ASR]
- Reber, A. S. & Allen, R. (1978) Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition* 6:189-221. [aDRS, HN, ASR]
- Reber, A. S., Allen, R. & Regan, S. (1985) Syntactical learning and judgment, still unconscious and still abstract: Comment on Dulany, Carlson, and Dewey. *Journal of Experimental Psychology: General* 114:17-24. [aDRS, RAC]
- Reber, A. S., Kassin, S. M., Lewis, S. & Cantor, G. (1980) On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory* 6:492-502. [aDRS, CAS]
- Reber, A. S. & Lewis, S. (1977) Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition* 5:333-61. [aDRS, ASR]
- Reber, A. S., Walkenfeld, F. F. & Hernstadt, R. (1991) Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17:888-96. [HN, ASR, CAS]
- Reed, J. & Johnson, P. (1994) Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:585-94. [JR, CAS]
- Regehr, G. & Brooks, L. R. (1993) Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General* 122:92-114. [aDRS]
- Reingold, E. M. & Merikle, P. M. (1988) Using direct and indirect measures to study perception without awareness. *Perception and Psychophysics* 44:563-75. [aDRS, RFB, PMM]
- Richardson-Klavehn, A. & Bjork, R. A. (1988) Measures of memory. *Annual Review of Psychology* 39:475-543. [aDRS]
- Ridley, M. (1986) *Evolution and classification*. Longman. [DBW]
- Ristau, C. A., ed. (1991) *Cognitive ethology: The minds of other animals*. Erlbaum. [ASR]

References/Shanks & St. John: Dissociable learning systems

- Rizzo, A., Parlangei, O., Cambiganu, C. & Bagnara S. (1993) Control of complex systems by situated knowledge: The role of implicit learning. In: *Human-computer interaction, 19b: Software and hardware interfaces*, ed. M. J. Smith & G. Salvendy. Elsevier. [AR]
- Roediger, H. L. (1990) Implicit memory: Retention without remembering. *American Psychologist* 45:1043-56. [aDRS]
- (1993) Learning and memory: Progress and challenge. In: *Attention & performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, ed. D. E. Meyer & S. Kornblum. MIT Press. [rDRS]
- Rosch, E. H. (1975) Cognitive representations of semantic categories. *Journal of Experimental Psychology* 104:192-233. [aDRS]
- Rosenfarb, I. S., Newland, M. C., Brannon, S. E. & Howey, D. S. (1992) Effects of self-generated rules on the development of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior* 58:107-21. [aDRS]
- Ross, B. H. (1987) This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:629-39. [RLG]
- Rovee-Collier, C. (1990) The "memory system" of prelinguistic infants. *Annals of the New York Academy of Sciences* 608:517-42. [ASR, HST]
- Rovee-Collier, C., Griesler, P. C. & Earley, L. A. (1985) Contextual determinants of retrieval in three-month-old infants. *Learning and Motivation* 16:139-57. [HST]
- Rudy, J. W. & Sutherland, R. J. (1992) Configural and elemental associations and the memory coherence problem. *Journal of Cognitive Neuroscience* 4:208-16. [rDRS]
- Rumelhart, D. E. & McClelland, J. L. (1986) On learning the past tense of English verbs. In: *Parallel distributed processing: Explorations in the microstructure of cognition, vol. 2: Psychological and biological models*, ed. J. L. McClelland & D. E. Rumelhart. MIT Press. [rDRS]
- Russell, I. F. (1986) Comparison of wakefulness with two anaesthetic regimens. *British Journal of Anaesthesia* 58:965-68. [JA]
- (1989) Conscious awareness during general anaesthesia: Relevance of autonomic signs and isolated arm movements to depth of anaesthesia. In: *Ballière's Clinical Anaesthesiology: Depth of Anaesthesia*, ed. J. G. Jones. Balliere. [JA]
- Ryle, G. (1949) *The concept of mind*. Viking Penguin. [EE]
- St. John, M. F. (1994) Learning language in the service of a task. *Technical report*, Department of Cognitive Science, University of California, San Diego. La Jolla, CA. [rDRS]
- St. John, M. F. & McClelland, J. L. (1990) Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence* 46:217-57. [rDRS]
- Sanderson, P. M. (1989) Verbalizable knowledge and skilled task performance: Association, dissociation, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:729-47. [aDRS, DCB, AR]
- Saunders, R. C. & Weiskrantz, L. (1989) The effects of fornix transection and combined fornix transection, mammillary body lesions and hippocampal ablations on object-pair association memory in the rhesus monkey. *Behavioral Brain Research* 35:85-94. [LRS]
- Schacter, D. L. (1987) Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:501-18. [aDRS]
- (1989) On the relation between memory and consciousness: Dissociable interactions and conscious experience. In: *Varieties of memory and consciousness*, ed. H. L. Roediger & F. I. M. Craik. Erlbaum. [aDRS]
- (1990) Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. *Annals of the New York Academy of Sciences* 608:543-71. [CJM]
- (1992) Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *Journal of Cognitive Neuroscience* 4:244-56. [CJM]
- Schacter, D. L., Chiu, C. Y. & Ochsner, K. N. (1993) Implicit memory: A selective review. *Annual Review of Neuroscience* 16:159-82. [LRS]
- Schacter, D. L., Harbluk, J. L. & McLachlan, D. R. (1984) Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior* 23:593-611. [rDRS]
- Schank, R. C. (1982) *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press. [RLG]
- Schmidt, H. (1969) Precognition of a quantum process. *Journal of Parapsychology* 33:99-108. [aDRS]
- Schneider, W. & Shiffrin, R. M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review* 84:1-66. [JL]
- Schwender, D., Kaiser, A., Klasing, S. & Peter, K. (1992) Explicit and implicit memory and mid-latency auditory evoked potentials during cardiac surgery. *Anesthesia and Analgesia* 74:S269. [JA]
- Seamon, J. G., Brody, N. & Kauff, D. (1983a) Affective discrimination of stimuli that are not recognized: Effects of shadowing, masking, and cerebral laterality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9:544-55. [NB]
- (1983b) Affective discrimination of stimuli that are not recognized: II. Effects of delay between study and test. *Bulletin of the Psychonomic Society* 21:187-89. [NB, RFB]
- Seger, C. A. (1994) Implicit learning. *Psychological Bulletin* 115:163-96. [CAS]
- Seidenberg, M. S. & McClelland, J. L. (1989) A distributed, developmental model of word recognition and naming. *Psychological Review* 96:523-68. [rDRS]
- Sejnowski, T. J. & Rosenberg, C. R. (1987) Parallel networks that learn to pronounce English text. *Complex Systems* 1:145-68. [rDRS]
- Sells, P. (1985) *Lectures on contemporary syntactic theories: An introduction to government-binding theory, generalized phrase structure grammar, and lexical function grammar*. Center for the Study of Language and Information. [JL]
- Servan-Schreiber, D., Cleeremans, A. & McClelland, J. L. (1991) Graded state machines: The representation of temporal contingencies in simple recurrent networks. *Machine Learning* 7:161-93. [AC]
- Servan-Schreiber, E. & Anderson, J. R. (1990) Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:592-608. [aDRS]
- Shanks, D. R. (1993) Human instrumental learning: A critical review of data and theory. *British Journal of Psychology* 84:319-54. [aDRS]
- Shanks, D. R. & Dickinson, A. (1990) Contingency awareness in evaluative conditioning: A comment on Baeyens, Eelen & van den Bergh. *Cognition and Emotion* 4:19-30. [aDRS]
- (1991) Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Memory and Cognition* 19:353-60. [aDRS, FS]
- Shanks, D. R. & Gluck, M. A. (1994) Tests of an adaptive network model for the identification and categorization of continuous-dimension stimuli. *Connection Science* 6:69-99. [aDRS]
- Shanks, D. R., Green, R. E. A. & Kolodny, J. (1994) A critical examination of the evidence for nonconscious (implicit) learning. In: *Attention and performance XV: Conscious and nonconscious information processing*, ed. C. Umiltà & M. Moscovitch. MIT Press. [aDRS, JR, CAS]
- Sherry, D. F. & Schacter, D. L. (1987) The evolution of multiple memory systems. *Psychological Review* 94:439-54. [LRS]
- Shettleworth, S. J. (1993) Varieties of learning and memory in animals. *Journal of Experimental Psychology: Animal Behavior Processes* 19:5-14. [MLH]
- Shiffrin, R. M. & Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84:127-90. [JL]
- Shimamura, A. P. (1986) Priming effects in amnesia: Evidence for a dissociable memory function. *Quarterly Journal of Experimental Psychology* 38A:619-44. [DBW]
- Shimamura, A. P. & Squire, L. R. (1989) Impaired priming of new associations in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:721-28. [aDRS]
- Shimp, C. P., Sabulsky, S. L. & Childers, L. J. (1989) Preference for starting and finishing behavior patterns. *Journal of the Experimental Analysis of Behavior* 52:341-52. [ACC]
- Shin, H. & Nosofsky, R. M. (1992) Similarity scaling studies of "dot pattern" classification and recognition. *Journal of Experimental Psychology: General* 121:278-304. [aDRS]
- Simon, J. R. & Craft, J. L. (1989) The effect of prediction accuracy on choice reaction time. *Memory and Cognition* 17:503-8. [aDRS]
- Skinner, B. F. (1945) The operational analysis of psychological terms. *Psychological Review* 52:270-77. [ACC, HST]
- (1957) *Verbal behavior*. Appleton-Century-Crofts. [ACC]
- (1984a) Behaviorism at fifty. *Behavioral and Brain Sciences* 7:615-67. [ACC]
- (1984b) An operant analysis of problem solving. *Behavioral and Brain Sciences* 7:583-613. [aDRS]
- (1984c) The operational analysis of psychological terms. *Behavioral and Brain Sciences* 7:547-81. [ACC]
- Skorstad, J., Gentner, D. & Medin, D. (1988) Abstraction processes during concept learning: A structural view. In: *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*. Montreal, Canada. [RLG]
- Slobin, D. (1978) A case study of early language awareness. In: *The child's conception of language*, ed. A. Sinclair, R. Jarvella & W. Levelt. Springer-Verlag. [JL]
- Smith, E. E. (1989) Concepts and induction. In: *Foundations of cognitive science*, ed. M. I. Posner. MIT Press. [PP]
- Smith, E. E., Langston, C. & Nisbett, R. (1992) The case for rules in reasoning. *Cognitive Science* 16:1-40. [aDRS]
- Smith, J. D. & Kemler Nelson, D. G. (1984) Overall similarity in adults'

- classification: The child in all of us. *Journal of Experimental Psychology: General* 113:137-59. [aDRS]
- Smith, J. D. & Shapiro, J. H. (1989) The occurrence of holistic categorization. *Journal of Memory and Language* 28:386-99. [aDRS, RLG]
- Sokal, R. R. & Sneath, P. H. A. (1963) *The principles of numerical taxonomy*. Freeman. [DBW]
- Spielberger, C. D. & DeNike, L. D. (1966) Descriptive behaviorism versus cognitive hypothesis in verbal operant conditioning. *Psychological Review* 73:306-26. [SSR]
- Squire, L. R. (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review* 99:195-231. [aDRS, ASR, CAS]
- Squire, L. R. & Frambach, M. (1990) Cognitive skill learning in amnesia. *Psychobiology* 18:109-17. [rDRS, CAS, DBW]
- Squire, L. R. & McKee, R. (1992) Influence of prior events on cognitive judgments in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:106-15. [aDRS, ZK]
- Stadler, M. A. (1989) On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:1061-69. [aDRS, DCB, SSR, MAS]
- Stanley, W. B., Mathews, R. C., Buss, R. R. & Kotler-Cope, S. (1989) Insight without awareness: On the interaction of verbalization, instruction and practice in a simulated process control task. *Quarterly Journal of Experimental Psychology* 41A:553-77. [arDRS, DCB, AR, ASR]
- Sutherland, R. W. & Rudy, J. W. (1989) Configural association theory: The role of the hippocampal formation in learning, memory and amnesia. *Psychobiology* 17:129-44. [LRS]
- Svartdal, F. (1989) Shaping of rule-governed behaviour. *Scandinavian Journal of Psychology* 30:304-14. [aDRS, FS]
- (1991) Operant modulation of low-level attributes of rule-governed behavior by nonverbal contingencies. *Learning and Motivation* 22:406-20. [aDRS, FS]
- (1992) Sensitivity to nonverbal operant contingencies: Do limited processing resources affect operant conditioning in humans? *Learning and Motivation* 23:383-405. [rDRS, ACC, FS]
- Svartdal, F. & Mortensen, T. (1993) Effect of reinforcer value on sensitivity to nonverbal operant contingencies in humans. *Quarterly Journal of Experimental Psychology* 46A:347-64. [FS]
- Terrace, H. S. (1985) Animal cognition: Thinking without language. *Philosophical Transactions of Royal Society (London)* B 308:113-28. [HST]
- Thompson, R. F. (1990) Neural mechanisms of classical conditioning in mammals. In: *Behavioral and neural aspects of learning and memory*, ed. J. R. Krebs & G. Horn. Clarendon Press. [LRS]
- Thorndike, E. L. & Rock, R. T. (1934) Learning without awareness of what is being learned or intent to learn it. *Journal of Experimental Psychology* 17:1-19. [aDRS]
- Tulving, E. (1983) *Elements of episodic memory*. Oxford University Press. [aDRS]
- (1985) Memory and consciousness. *Canadian Psychology* 25:1-12. [MAS]
- Tulving, E. & Schacter, D. L. (1990) Priming and human memory systems. *Science* 247:301-5. [CJM, LRS]
- Tunstall, M. E. (1977) Detecting wakefulness during general anaesthesia for caesarean section. *British Medical Journal* 1:1321. [JA]
- Ungerleider, W. & Mishkin, M. (1982) Two cortical visual systems. In: *The analysis of visual behavior*, ed. D. J. Ingle, M. A. Goodale & R. J. W. Mansfield. MIT Press. [LRS]
- Vallacher, R. R. & Wegner, D. M. (1987) What do people think they're doing: Action identification and human behavior. *Psychological Review* 94:3-15. [RAC]
- Vaughan, W., Jr. (1988) Formation of equivalence sets in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 14:36-42. [ACC]
- Velmans, M. (1991) Is human information processing conscious? *Behavioral and Brain Sciences* 14:651-726. [aDRS, JL]
- Verplanck, W. S. (1992) Verbal concept "mediators" as simple operants. *Analysis of Verbal Behavior* 10:45-68. [ACC]
- Vokey, J. R. & Brooks, L. R. (1992) Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:328-44. [aDRS]
- Vokey, J. R. & Read, J. D. (1985) Subliminal messages: Between the Devil and the media. *American Psychologist* 40:1231-39. [aDRS]
- Ward, T. B. & Scott, J. (1987) Analytic and holistic modes of learning family-resemblance concepts. *Memory and Cognition* 15:42-54. [aDRS]
- Wason, P. C. (1968) Reasoning about a rule. *Quarterly Journal of Experimental Psychology* 20:273-81. [aDRS]
- Weiskrantz, L., ed. (1985) *Animal intelligence*. Oxford University Press. [ASR]
- Weiskrantz, L. & Warrington, E. K. (1979) Conditioning in amnesic patients. *Neuropsychologia* 17:187-94. [aDRS, DBW]
- Whittlesea, B. W. A. (1993) Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:1235-53. [rDRS]
- Whittlesea, B. W. A. & Dorken, M. D. (1993) Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General* 122:227-48. [aDRS, AC, HN, ASR]
- Widrow, B. & Hoff, M. E. (1960) Adaptive switching circuits. *Institute of Radio Engineers, Western Electronics Show and Convention Record*, Part 4: 96-104. [aDRS]
- Willingham, D. B., Greeley, T. & Bardone, A. (1993) Dissociation in a serial response time task using a recognition measure: Reply to Perruchet and Amorim. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:1424-30. [aDRS, DBW]
- Willingham, D. B., Nissen, M. J. & Bullemer, P. (1989) On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:1047-60. [aDRS, DCB, KJH]
- Winograd, T. (1975) Frame representations and the declarative-procedural controversy. In: *Representation and understanding: Studies in cognitive science*, ed. D. Bobrow & A. Collins. Academic Press. [aDRS]
- Wolpaw, J. R. & Carp, J. S. (1990) Memory traces in spinal cord. *Trends in Neurosciences* 13:137-42. [GO]
- Wood, F., Ebert, V. & Kinsbourne, M. (1982) The episodic-semantic distinction in memory and amnesia: Clinical and experimental observations. In: *Human memory and amnesia*, ed. L. S. Cermak. Erlbaum. [aDRS]
- Wood, J. M., Bootzin, R. R., Kihlstrom, J. F. & Schacter, D. L. (1992) Implicit and explicit memory for verbal information presented during sleep. *Psychological Science* 3:236-39. [aDRS]
- Zajonc, R. B. (1968) Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph* 9:1-27. [RFB]
- Zola-Morgan, S. & Squire, L. R. (1984) Preserved learning in monkeys with medial temporal lesions: Sparing of motor and cognitive skills. *Journal of Neuroscience* 4:1072-85. [LRS]
- (1986) Memory impairment in monkeys following lesions of the hippocampus. *Behavioral Neuroscience* 100:155-60. [ASR]

ANNOUNCING A NEW JOURNAL

Personal Relationships

Journal of the International Society for the Study of Personal Relationships

PREMIERE ISSUE • MARCH 1994

Editor

Patricia Noller, *University of Queensland, Australia*

Associate Editors

Rodney Cate, *Iowa State University*

Keith Davis, *University of South Carolina*

Sandra Metts, *Illinois State University*

Susan Sprecher, *Illinois State University*

Personal Relationships will be the official quarterly journal of the *International Society for the Study of Personal Relationships (ISSPR)*.

Aims and Scope

The aim of this international, interdisciplinary journal is to promote scholarship in the field of personal relationships throughout a broad range of disciplines including:

- psychology • psychiatry • communication studies
- sociology • anthropology • family studies

The subject matter and approach will be of interest to academics, researchers and practitioners alike.

Personal Relationships will cover topics such as:

- equity • conflict • jealousy • attachment • love
- costs and benefits of relationships • sibling rivalry
- social support • loneliness • kinship • gender and sexuality

All personal relationships, romantic, marital, parent-child, and friendship will be explored including those between romantic partners, spouses, parents and children, siblings, and friends.

Manuscript submission

Empirical research, review articles and descriptions of new research methods, from scholars working in all aspects of personal relationships and relevant disciplines, may be submitted.

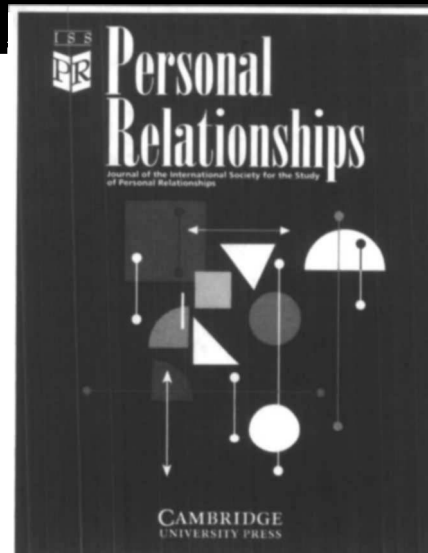
Please submit 4 copies of manuscripts in APA style. For expediency, a title page and abstract may be sent by E-mail or by fax to the editor who will advise the author of the assigned associate editor to whom the manuscript should be mailed.

Scholarly comment for selected lead articles may be sought by the editor and published to encourage debate over theoretical and empirical issues.

Please address material and inquiries to:

Patricia Noller, PhD., Editor, *Personal Relationships*, Dept. of Psychology, University of Queensland, St. Lucia, Queensland 4062, Australia.

Fax: (61-7) 365 4466 E-mail: pn@psych.psy.uq.oz.au



The **International Society for the Study of Personal Relationships** is the foremost scholarly society in this area of study. Members will automatically receive the newsletter, membership directory, and the journal, as well as an invitation to the biennial conference.

Through ISSPR you can meet others in the field, stay informed of important developments, keep up with current research and assessment measure, and discover useful literature.

For ISSPR inquiries

Please contact Dr. Catherine Surra, Child Development and Family Relations Dept., University of Texas at Austin, Gearing 117, Austin, TX 78712-1097. Phone: (512) 471-4120.

Personal Relationships subscription information

ISSN 1350-4126

Volume 1 • Numbers 1-4 • 1994

Institutional \$120.00

Individual \$60.00



CAMBRIDGE
UNIVERSITY PRESS

40 West 20th Street
New York, NY 10011

Call toll free 1-800-431-1580 x175:
or FAX 914-937-4712

MasterCard / VISA accepted.