DEBATES IN PSYCHOLOGY

Series editor Dianne C. Berry
Series advisor Lawrence Weiskrantz

Martin Conway (ed.) Recovered Memories and False Memories
Dianne C. Berry (ed.) How Implicit Is Implicit Learning?

How Implicit Is Implicit Learning?

Edited by

DIANNE C. BERRY

Department of Psychology
University of Reading

Oxford New York Tokyo
OXFORD UNIVERSITY PRESS
1997

SIX

A subjective unit formation account of implicit learning

PIERRE PERRUCHET and JORGE GALLEGO

6.1 INTRODUCTION

6.1.1 A preliminary definition of implicit learning

Given that the notion of implicit learning conveys partially different meanings for different authors, we must first specify what we consider to be the essence of the phenomenon we are considering. Hereafter, implicit learning designates an adaptive mode in which subjects' behaviour is sensitive to the structural features of a previously presented situation, without this adaptation being due to the intentional exploitation of subjects' explicit knowledge about these features.

The two components of this definition, namely (1) the behavioural sensitivity to the structure of a complex situation and (2) the lack of any explicit mediation for this sensitivity, have formed part of virtually all earlier definitions of implicit learning. Reber (1989) posited that the knowledge base acquired in implicit learning studies 'can be used implicitly to solve problems and make accurate decisions about novel stimulus circumstances'. Likewise, Cleeremans (1993, p. 199) noted that 'implicit learning tasks do not *require* explicit or conscious access to the information on which performance is based'. Defining implicit learning in such a way removes any doubt about the existence of the phenomenon in natural settings. Although it would be difficult to draw up a list of the relevant natural phenomena, it is arguable that the acquisition of language and the development of perceptual–motor skills or social abilities, are largely the result of implicit learning processes. The fact that the phenomenon is conveniently reproduced in experimental contexts is known primarily from introspective reports. The most common result of implicit learning studies is that normal subjects report that they have not controlled their behaviour on the basis of explicit representations of the situation. Evidence stemming from amnesic patients also supports the conclusion that performance may be causally independent of explicit representations. These patients exhibit a deep deficit in their ability to remember explicitly any component of the training session. Nevertheless, their performance has been found to improve with practice in implicit learning settings (e.g. Knowlton and Squire 1994; Nissen and Bullemer 1987).

Our contention is that the other features that most authors consider to

be indissociable from the notion of implicit learning are either optional or unwarranted. For example, implicit learning is typically observed in situations where subjects have been exposed to the structured situation in a relatively passive way. That is, subjects are not instructed to engage rule searching, hypothesis testing, and similar analytical activities, although their attention is usually directed to the material itself. However, it is arguable that implicit learning occurs in all situations, even when subjects have been instructed to search for rules. The phenomenon is more easily observed in conventional situations, because the effect of strategic control on performance is reduced. But the absence of active rule searching is a procedural convenience rather than a defining feature of implicit learning.

It is also worth stressing that our definition departs from that of some contributors, in that we do not posit that explicit knowledge about the training situation is lacking, or even limited. It is, of course, possible to include this property in the concept of implicit learning given that terminological issues are arbitrary, but we will see later that this property is not observed in the studies commonly categorized as investigating implicit learning, at least in the studies using conventional laboratory situations in normal subjects. Our claim is that access to consciousness is simply not needed for behavioural adaptation.

6.1.2 Overview of the chapter

In keeping with a general principle of parsimony, we started our research programme on implicit learning with the aim of examining to what extent the well-established principles put forward in the literature on learning and memory to account for elementary phenomena, such as classical conditioning and priming, were able to account for the data collected in this new domain. With regard to the well-investigated, earlier situations, implicit learning paradigms have the peculiarity that subjects are faced with complexly-structured material. The challenge was to account for the apparent sensitivity of subjects to this complex structure within the conceptual frameworks put forward for learning in far simpler settings. The model that is presented below matches these requirements insofar as it involves only the learning processes that are engaged and have observable consequences on performance in all situations, *whether they are governed by complex rules or not*. As it will (hopefully) be made clear later, sensitivity to complex rules emerges as a by-product of processes observable in very simple training situations.

The bulk of this chapter focuses on the most conventional implicit learning situations, namely the repeated sequence tasks and the artificial grammar learning settings. In both cases, subjects are exposed to material which is governed by sequential rules, that is to say by rules that define the nature of an event (from a small set of possible events) as a function of one or more preceding events. In artificial grammar learning, the events are generally consonant letters, and subjects' performance is assessed through their judgements of the grammaticality of new sequences. In repeated sequence tasks, the events are targets located at several places on a computer screen, and performance is assessed through choice reaction times to these targets.

We start with a task analysis (Section 6.2) in order to examine what structural information is really embedded in these situations. We show that most of the regularities embedded in the situations can be described as first and second order sequential rules. We then argue (Section 6.3) that subjects capture most of these regularities by building small, disjunctive units designed to optimize parsing of the material, in a sense that will be specified later. We devote the major part of the chapter to providing a theoretical account of these phenomena. We suggest that new processing units are built through associative processes, leading to the unitization of the set of primitive features simultaneously processed in the attentional focus. These new processing units shape subjects' phenomenal experience of the world (Section 6.4). This account is then compared with alternative accounts, primarily the Reber abstractionist theory and the memory-based, or episodic, accounts (Section 6.5).

Of course, implicit learning is not restricted to repeated sequences tasks and artificial grammar learning. There are other, generally more complex, experimental implicit learning situations and, moreover, all laboratory situations are intended to reproduce some aspects of (still more complex) real-world situations of learning. The relevance of our framework to these other situations will be evaluated in Section 6.6. To conclude, we summarize our account as a small number of key propositions (Section 6.7).

6.2 IN SEARCH OF THE INFORMATION CONTENT OF THE TASKS

What structural information is in fact embedded in the tasks used in implicit learning studies? In repeated sequence and artificial grammar learning tasks, the primitive components (e.g. letters in grammar, spatial locations in sequence tasks) are limited in number and quickly identified. The key structural constraints are the dependency rules, which determine both possible and impossible transitions between primitives. The crucial requirement is to identify the order of these rules, that is the number of prior events that should be considered in order to predict a target event. The minimum number of prior events to be considered may be defined by the fact that some possible successors are authorized while some others are not. To take an example: the consideration of one-order dependency rules is irrelevant (i.e. provide no information about the constraints of the situation) if the generative rules allow all the possible pairwise combinations between primitives. (We will see later that, in this latter case, information can be conveyed through the differential frequency of these units. However, we can ignore this factor for our present purposes). Likewise, the maximum number of prior events to be considered may be defined by

the fact that considering an antecedent event does not improve prediction of the target.

Let us first consider the case of repeated sequence tasks, that is tasks patterned after the Nissen and Bullemer (1987) paradigm. In these tasks, the knowledge of two-trial sequences is moderately relevant. For instance, in the sequence used in a large proportion of these studies (D–B–C–A–C–B–D–C–B–A–D . . .), 9 out of the 12 (4×3, given that repetitions are excluded) possible transitions are used. However, the knowledge of sequences comprising three locations accounts for many of the constraints of the situation, given that only 10 out of the 36 (4×3×3) possible three-trial sequences are valid. Thus, the uncertainty about the location of the next target is only moderately reduced when only the one-order dependency rules are considered, but considerably reduced when second-order dependency rules are taken into account. In fact, in the typical sequence, the only information conveyed by higher order dependency rules concerns the location following C–B (when only three-trial sequences are considered, C–B may be followed by D or A). This means that, in the conventional repeated sequence task (and similar analyses could be undertaken with other material used in published studies), almost all the constraints embedded in the repeated sequence may be captured when only second-order dependency rules are considered, whatever the other contextual elements. This is true even of the more complex sequences that have been used so far, such as the so-called 'ambiguous sequences' of Cohen *et al.* (1990), and the 'second-order conditional sequences' of Reed and Johnson (1994; see also Gomez, submitted).

The matter is somewhat more complex in the field of artificial grammar learning. Indeed, because the material is segmented into strings of limited length, the rules determine properties other than the possible transitions between letters. For instance, the nature of the letters permissible at the start and at the end of the strings also determines well-formedness. We will return to these questions later. For the moment, however, we shall consider only those rules that relate to sequential dependencies. It may be noted that in most, if not all, of the finite-state grammars used in the earlier published studies, the knowledge of the grammatical bigram units is relevant. In most studies, the grammars are composed of five letters, thus making 25 possible bigrams; roughly half of these possible bigrams are typically defined as admissible (e.g. 12 in Dulany *et al.* 1984; 14 in Gomez and Schvaneveldt 1994 and in Servan-Schreiber and Anderson 1990), hence maximizing information content at the bigram level as high as it can be. By way of illustration, let us assume that subjects encode only the 12 permissible bigrams from the material used initially by Reber (1976) and subsequently by many others, and that they classify test items as grammatical if they consist only of permissible bigrams (whatever their location in the strings). This would result in the production of exactly 90% correct responses, a success level that greatly exceeds observed performance! Of course, consideration of the information provided at the trigram level still improves simulated performance, although the effect is considerably limited because of obvious ceiling effects.

This analysis reveals that first-order and second-order dependency rules capture virtually all the structural constraints of the most common situations of implicit learning. Note that such a conclusion is not trivial. It is a widespread view that improved performance implies knowledge of more abstract properties of the situations. For instance, it is sometimes claimed that performance improvement in serial reaction time tasks implies the (implicit) knowledge that the very same long sequence of events is continuously cycled. It should be realized that this form of knowledge is useless. Of course, it can be argued that certain other implicit learning situations can not be described in the same terms, and that subjects also improve their performance in these situations. We shall postpone until the final section our discussion of how the present analysis can be broadened to take account of improved performance in such situations. In the next section, we will show how subjects capture the dependency rules in the situations where they are clearly relevant.

6.3 THE FORMATION OF SUBJECTS' UNITS

6.3.1 Subjects encode small units

Our proposal is that subjects do not directly process the dependency rules which structure the material, but rather encode the material as a succession of small units that partially embed the structuring rules. The data on which our arguments are grounded differs depending on whether we are considering artificial grammar or sequence learning tasks, partly because reaction times, unlike grammaticality judgments, provide a trial-by-trial measure of performance which allows a more direct assessment of the units which are really encoded by subjects. However, the different lines of investigation tend toward similar conclusions.

In the field of artificial grammar learning, Perruchet and Pacteau (1990) tested the hypothesis that subjects essentially learn the bigrams that compose the study strings. They reasoned that if subjects learn only bigram information when faced with the whole strings, the direct presentation of the bigrams should not change the final performance. They therefore compared the performance of subjects trained, as usual, using the complete grammatical strings, with the performance of subjects trained using the bigrams from which these strings are composed. This prediction was confirmed: when all the non-grammatical test items began with a valid first letter, the performance of subjects who had learned from the strings and those who were trained using the component bigrams were statistically indistinguishable. Additional support for the formation of small chunks of letters is provided by the fact that the verbal reports of subjects who were asked to give verbal instructions to a yoked partner during the study phase primarily refer to specific bigrams or trigrams (Mathews *et al.* 1989).

However, there is evidence that subjects also acquire other information. Using

the same general paradigm as Perruchet and Pacteau, Gomez and Schvaneveldt (1994) also observed that the performance of subjects trained with bigrams consistently exceeded a random response level in a grammaticality test, but, contrary to the results obtained by Perruchet and Pacteau, their performance was noticeably poorer than that of subjects trained in standard conditions. They concluded that subjects in standard conditions learn more about the language than permissible pairs of letters. The discrepancy between the results of Perruchet and Pacteau and those of Gomez and Schvaneveldt may be due to a difference in the way the study material was displayed. The items were displayed in a single list in the Perruchet and Pacteau study, while they were displayed in small subsets with instructions demanding item-by-item processing in Gomez and Schvaneveldt. It is quite possible that the subjects who viewed complete strings in Perruchet and Pacteau in fact allocated only a low level of attention to the individual items as a whole, and instead searched for regularities by continuously scanning the whole list. This would have artificially reduced their subsequent performance, making it indistinguishable from that of the bigram-trained subjects. Whatever the value of this *post-hoc* account, the available evidence (see also Dienes *et al.* 1991) supports the conclusion that subjects typically learn more than valid bigrams when learning from grammatical strings (note that Perruchet and Pacteau also showed that subjects learned about the valid initial letters of strings (Experiment 1), and, to a lesser extent, about the location of permissible bigrams (Experiment 2)).

These results do not rule out the view that subjects encode small units, but they suggest that the potential content of these units should be extended. The subjects' units may comprise surface features other than the constituent letters of the strings. For instance, the constituent letters may be associated with their location in the strings, thus resulting in the formation of units such as '– VM' (i.e 'VM in the beginning position').

In sequence learning, the nature of the units governing reaction time shortening has been directly documented by Cleeremans and McClelland (1991). Their data show that subjects form more than just pairwise associations. However, the size of the relevant units appears to be severely constrained. Cleeremans (1993, ch. 5) reports what seems to be the largest value: after considerable practice (60 000 trials), subjects appear to be sensitive to the information conveyed by the trial four steps away from the current trial. This limit was observed in a situation specially designed in such a way that a consideration of fourth order dependency rules improved predictions, and this effect occurred even when the three intervening trials provided no information about the outcome. Unsurprisingly, there is evidence that the task becomes more difficult when the number of predictive elements increases. The results obtained by Cohen *et al.* (1990) may be rephrased in this way. What they call 'unique associations' are in fact cases in which two-trial units are relevant: the first item perfectly predicts the second. In contrast, two-trial units are not (or are only partially) informative in what they call 'ambiguous associations'. In this case, it is necessary to consider three or more trial units in order to capture the structure of the situation. The work of these authors demonstrates that it is more difficult to learn 'ambiguous associations' than 'unique associations'.

6.3.2 *Subjects' units are disjunctive and tend toward an optimal parsing of the material*

We must now discuss an important, although systematically neglected, issue. This relates to the problem of knowing whether subjects form physically separate entities, or overlapping units. For instance, does a subject faced with ABCDE form the units, say, ABC and DE, or ABC, BCD, and CDE? To suppose that subjects form disjunctive units raises a second issue, which relates to the way in which they segment the training material. By way of illustration, consider Table 6.1a, which represents two modes of segmentating a block of three 10-trial cycled sequences. Although the *total* number of units (and hence the mean length of the units) does not differ in the two modes of parsing, it can be seen that the strings are described alternatively as being composed of three or seven *different* units. The matter is more complex in the artificial grammar field, because the strings differ from each other. However, the problem is similar. Table 6.1b represents two modes of segmentation of three strings of letters, in which the strings are described alternatively as composed of four or seven different units.

As we shall show later, these problems are crucial for a theory of implicit learning aimed at accounting for the mechanism responsible for the formation of the cognitive units. However, they are also surprisingly difficult to address empirically. This is because a model positing overlapping units predicts a mean performance which is indistinguishable from that predicted by the alternative model, provided that different subjects form different units. Our proposal is that subjects form disjunctive units, which are consistent across the study material.

There is evidence that subjects partition the material while encoding it in the study session. For example when subjects are asked to write down the study

Table 6.1 Two modes of parsing a 10-trial sequence repeated three times (a), and three grammatical strings of letters (b); although, within each domain, segmentations (1) and (2) parse the material into the same number of units, they differ in economy, because segmentation (1) leads to the formation of a larger number of *different* units than segmentation (2).

a	
1	D B C/A C B/D C B A/D B C/A C B/D C B A/D B C/A C B/D C B A
2	D B C/A C B/D C B A/D B C A/C B D/C B A/D B C/A C B D/C B A
b	
1	M/T M V M T V?R/X R M V/R X R
2	M T/M V M T/V/R X R M V/R X R

items of an artificial grammar experiment, they frequently reproduce the strings as separate groups of letters (Servan-Schreiber and Anderson 1990). The claim is that these perceptual units lie at the root of the units which underlie performance improvement. Experimental support for this view has been provided by studies in which the nature of the perceptual units has been constrained by experimental manipulation. In artificial grammar learning, Servan-Schreiber and Anderson (1990) introduced spaces between small groups of letters within each string, and, in a repeated sequence task, Stadler (1993, 1995) introduced a temporal pause at some locations within the sequence. In both sets of experiments, these (spatial or temporal) intervals were positioned in such a way as to induce a more or less consistent parsing of the material. Consistent parsing is induced by introducing the intervals in a consistent way, as illustrated in Table 6.1. Servan-Schreiber and Anderson described the product of their grammar (taken from Reber 1976) alternatively as the combination of 27 or 9 component units, depending on the way these units were chosen. The main result of these experiments is that subjects' performance is sensitive to the mode of presentation. Consistent grouping induces better performance, both in artificial grammar learning (Servan-Schreiber and Anderson 1990) and in sequence learning (Stadler 1993, 1995). These data provide striking evidence that the units formed during the coding stage are the relevant units for learning.

The fact that consistent grouping induces the best performance does not in itself indicate that subjects behave in such a way. During the first stage of exposure to the material, segmentation is presumably induced by salient surface features, such as those indicated by gestalt theorists, and/or subjects' background knowledge. For instance, in artificial grammar learning, the string TV??? will be more probably segmented as TV/??? than TV?/??. This does not necessarily lead to efficient coding. Table 6.1.b illustrates a case where the best coding requires the division of the TV fragment. Now, there is evidence to suggest that subjects perform segmentation efficiently. For instance, in the Servan-Schreiber and Anderson studies, the performance of subjects for whom no special structuring has been artificially induced approximates to the performance of subjects who viewed the well-structured strings. In a still unpublished experiment, Perruchet *et al.* (in preparation) provided more direct evidence of the fact that training leads to an improvement in the efficiency of initial segmentation. They asked their subjects to read each string generated by a finite state grammar and, immediately after reading, to mark with a slash bar the natural segmentation positions. They repeated the same task after a phase of familiarization with the material, which consisted, depending on the group involved, of learning items by rote, performing a short-term matching task, or searching for rules. Subjects formed the same number of total units before and after the training phase, thus indicating that they did not tend to form larger and larger units. However, the number of different units reliably decreased, whatever the task the subjects had performed during training. The conclusion is that exposure to structured material modifies the natural coding of the material and leads to more efficient segmentation.

Up to now, we have taken it for granted that consistent partitioning provides optimal coding. Indeed, describing the material with a smaller number of different units appears to be a more economical procedure; it is obviously easier to learn a small number of units, some of which are repeated, than a larger number of different units which are presented once. However, it must be realized that an economical, consistent coding results in the loss of part of the information conveyed through the strings. For example, the fact that the grammar allows the succession TM and TV is not represented in the economical coding shown in segmentation (2) in Table 6.1b, whereas the coding shown in segmentation (1) preserves this information. What is lost in the consistent partition of sequential material is the probability of transition between primitives belonging to contiguous units. A preliminary observation is that the establishment of an economical coding for material generated by a finite-state grammar logically implies that the material is segmented at the points where the transition probabilities are the weakest. Looking again at Table 6.1b, it appears, if we consider the first letter string, that the segmentation MT/MV is more economical than the segmentation M/TMV, because the bigram MT is present elsewhere in the material, whereas the bigram TM is unique to this string. The formation of an optimal parsing results in some possible transitions being ignored, but, as a matter of fact, these will necessarily be the least frequent transitions.

However, it can be argued that neglecting the boundaries is not only the least bad alternative, but is fully adaptive in itself. This point deserves being emphasized, because early versions of the present account are sometimes described as relying on 'fragmentary knowledge', with the tacit implication that, although knowledge of fragments may be sufficient, more extensive knowledge would be better. In fact, paradoxically, the reverse is true. Coding the material as 'fragments', far from being a degraded procedure, is a highly powerful mode of coding. Indeed, in the standard artificial grammar situations, dealing with small units facilitates transfer and generalization. This happens because, given the structure of finite-state grammars, new items are often formed by recombining old components. Componential coding is specially powerful if components fit with the main paths of the grammar, and this is precisely the logical outcome of partitioning performed in accordance with the principles described above. The greater explanatory power of componential knowledge (with regard to knowledge of complete strings) in accounting for transfer performance has been demonstrated in the Perruchet (1994*b*) reanalysis of the Vokey and Brooks (1992) data. Whereas Vokey and Brooks invoked two independent processes, one assessing the similarity of the test items with the most similar training exemplar, and the other computing similarity to the whole set of training items, Perruchet (1994*b*) showed that reliance on componential knowledge accounts for the overall variance in performance. We will see later that this may be even more advantageous when accounting for learning in real-life settings than in standard laboratory paradigms, given the hierarchical structure of some of the encountered situations.

6.3.3 Subjects' units are normally available in explicit memory

As noted in the Introduction, amnesic patients perform normally in implicit learning tasks, thus providing evidence that the ability to evoke the component units explicitly does not directly mediate performance in these tasks. But this result does not dispense with the need to examine whether normal subjects are ordinarily able to recollect explicitly their componential knowledge.

Researchers in the artificial grammar field have used a variety of methods. Some of these methods are akin to recall tests. For instance, Reber and Lewis (1977) asked subjects to generate grammatical strings from a subset of letters presented in scrambled order. Mathews *et al.* (1989) asked their subjects to provide their yoked partners instructions on how to classify. Other tasks are based on a recognition procedure. For instance, Dulany *et al.* (1984) instructed subjects to underline the part of the test strings that violated rules. Perruchet and Pacteau (1990) presented all the possible bigrams separately and asked subjects to rate on a six-point scale the extent to which they believed that each of the bigrams was displayed in the study phase. Likewise, subjects in Dienes *et al.* (1991) were asked which letters could occur after different stems. Whatever the choice of the test, the evidence that subjects are able to retrieve some units composed of a few atomistic elements of the situations is overwhelming. To quote Reber and Lewis' conclusion: 'Clearly, with only minimal exposure to grammatical constraints, subjects learned the bigram patterns of the language to an impressive degree' (1977, p.345).

Similar studies in repeated sequence tasks are less frequent, because researchers have persistently asked their subjects whether they have detected that a single long sequence was continuously repeated (despite the fact that, as has been shown above, such knowledge is of no use in the correct anticipation of the next event). However, the available evidence also suggests that subjects learn small units, although the size of the common units might be slightly larger than in artificial grammar learning. Perruchet and Amorim (1992) showed that subjects learned more than simple pairwise associations even after a very limited amount of training. For instance, when subjects were asked to reproduce the prior training sequence, their response for the trial n depended on the location of the target on trial n–1, but also in some cases, on trial n–2 Correlations over the serial positions between reaction time and explicit knowledge of three- and four-trial subsequences ranged from 0.63 to 0.98. Fine-grained analyses of the correspondence between explicit knowledge and reaction time provided an even more impressive idea of the goodness of fit, whatever the structure of the sequence (including or not unique pairwise associations) and the amount of attention devoted to the task (normal or dual-task conditions). Although using a partially different methodology, Shanks and Johnstone (in press) also report a close parallel between performance and explicit knowledge, and they reinterpret certain earlier results concluding in favour of independence (Reed and Johnson 1994). Of course, in consonance with the evolutive nature of subjects' units, obtaining substantial correlations is dependent on the synchrony of the measures of performance and explicit knowledge. As shown by Willingham *et al.* (1993), results in tests of explicit knowledge performed at the end of the training session do not correlate with performance measured in the early phase of practice.

The ability of the explicit knowledge base to account for performance improvement has been examined in a variety of studies. Generally, the procedure has consisted of examining whether the knowledge collected in explicit tests can be used in some performance model in order to classify the test items at the same level of accuracy as is achieved by actual subjects. Although relying on various performance models, Dienes *et al.* (1991), Dulany *et al.* (1984), and Perruchet and Pacteau (1990) have shown that simulations of performance based on the explicit recognition of small units conveniently accounted for observed performance. However, recall tests may reveal an insufficient knowledge base (Dienes *et al.* 1991; Jimenez *et al.* 1996; Mathews *et al.* 1989). To borrow the terms used by Berry (1994, p.771), the knowledge resulting from an implicit learning session has 'relative inaccessibility with free recall'.

To sum up, there is a close relationship between performance and explicit knowledge in normal subjects, although free recall tests may reveal an amount of knowledge more restricted than that revealed by performance. In our opinion, this close relationship is the fundamental result which has to be accounted for in any interpretive framework. This picture is at variance with a widely held view. Any reader who is even slightly acquainted with the literature on implicit learning knows that a majority of authors claim that implicit learning generates knowledge that is unavailable to explicit thought. This feature is considered as so important that it constitutes the first of the criteria put forward by several contributors to define implicit learning. To quote Reber (1989, p.219), 'Implicit learning produces a tacit knowledge base'. Seger (1994, p.164) echoes Reber's claim, asserting that subjects 'cannot provide a full (or, in many cases, any) verbal account of what they have learned'. Indeed, subjects are commonly surprised when they are asked for explicit remembering after an implicit learning session because they feel that they gained no articulable knowledge during the training episode.

It is likely that subjects who are asked to report the regularities they detected in the training situation search for some kind of abstract, general rules. Being unaware of such knowledge, they infer that they learned nothing at all. Our contention is that many researchers have long followed the same reasoning. The claim that a dissociation exists is largely due to a failure in identifying the nature of the knowledge underlying performance improvement or, in other words, a failure to fulfill Shanks and St. John's (1994) information criterion. There is agreement on the fact that subjects are typically unable to verbalize the

rules used by the experimenter to build his or her material, whether this relates to the existence of a finite state grammar in artificial grammar learning, or the fact that a single long sequence is continuously repeated in serial reaction time tasks. If one holds that performance improvement derives from knowledge of the abstract rules, then it follows that subjects are not consciously aware of this knowledge base. Irrelevant factors will clearly be thought to have an unconscious effect by those who invoke them. The literature is full of examples in which initial claims for unconscious learning have subsequently been discounted, because it later became apparent that performance was grounded on knowledge that was not considered in the original studies. An early example is the reanalysis by Dulany (1961) of the data on verbal operant conditioning. More recent illustrations include the reappraisal of the results reported by Lewicki *et al.* (1988) by Perruchet *et al.* (1990), and the reinterpretation of McGeorge and Burton (1990) result's by Wright and Burton (1995).

6.4 A SUBJECTIVE UNIT FORMATION ACCOUNT OF IMPLICIT LEARNING

We have first observed that the rules structuring the usual situations of implicit learning can be described as low-order sequential dependency rules. We have then shown that subjects capture most of these rules through the coding of the material into small and disjunctive units consisting of a few primitive elements. As training progresses, the original sequence is chunked into a minimum number of different units. In addition, experimental data show that subjects perform above a random response level in explicit tests asking for the recall or recognition of these units, although recall and recognition may be impaired in amnesic patients without detrimental consequences on performance.

We must now set these empirical data within a theoretical framework. Several questions have to be answered. Why do subjects encode the material in the form of small units? What are the mechanisms that account for the formation of an optimal coding? What is the exact status of subjects' units in the architecture of the mind, and how do they facilitate performance? And last, but not least, how are we to account for the – at least partial – integration of the units within the subjects' explicit knowledge base?

6.4.1. Accounting for small, disjunctive, and dynamic units

We propose to account for the formation of new units, and their evolution with task practice, within a theoretical framework borrowed from conditioning research. This does not mean that implicit learning can be equated with classical or operant conditioning. In fact, besides their greater simplicity, conventional conditioning situations differ from implicit learning paradigms in several features, notably the involvement of aversive or appetitive stimuli.

Moreover, conditioning phenomena are themselves multifaceted. Our concern here is the formation of new units. This aspect is undoubtedly one component of conditioning. However, many conditioning phenomena may be described more adequately as the elaboration of a relationship between two separate events, allowing, for instance, the first event to trigger preparatory behaviour for the second event (see Perruchet 1984 for a discussion of this distinction).

However, we also think that the theories put forward in the conditioning area may be of relevance for implicit learning. Of course, just as the conditioning phenomena are multifaceted, there are several theoretical views of conditioning. We subscribe here to a general model in which the attention devoted to the stimulus is construed as the major explanatory principle for association elaboration (e.g. MacKintosh 1975). More precisely, we posit that a new unit is formed as an automatic and mandatory consequence of the concurrent attentional processing of a few events. Note that such a position has also been adopted in other contexts. For instance, Ceraso (1985) has emphasized the role of the processes involved in the initial perceptual apprehension of the stimulus in the formation of psychological units, and Logan and Etherton (1994) have shed light on the role of attention in constructing an instance.

The formation of associations is known to result from the repeated presentation of primitive elements in spatial or temporal contiguity. If one considers that the primitive elements for the subjects are the letters in artificial grammar learning and the location of a target in sequence learning, then the formation of units composed of contiguous events, which is at the heart of implicit learning, may be viewed as a typical phenomenon of associative learning. The embedding of some contextual elements in the composition of units, such as the location of valid bigrams and trigrams within strings in artificial grammar, is itself consonant with associative learning phenomena. For instance, the dependency of conditioned reactions on contextual information has long been emphasized.

Associations stem from the concurrent processing of the primitives of the situations that can be simultaneously apprehended during initial coding. In the implicit learning context, the size of the units – a rough evaluation may be between two and four primitive components of the situation – may be directly related to the limited capacity of perceptual attentional processes. At the beginning of a training session, the segmentation of the material is presumably randomly determined, although natural saliency factors may contribute to the formation of these units. How can these units be modified to form optimal parsing of the material?

If one performs all the possible partitions of sequential material into a succession of units of a given range of size, the most economical parsing, as defined above, is that in which the number of different units is minimal. At first glance, selecting these units would seem to require sophisticated analytical tools. However, selection can be accounted for more simply within an associative framework. Indeed, a logical correlate of the minimum number

criterion is that the final units are the most frequent (for a fixed range of size). The detection by the subject of these most frequent units needs no special mechanism, but instead derives from the simple application of two fundamental principles. First, the most frequent associations are simply those that are privileged, in keeping with a general principle of associative learning, namely the beneficial effect of repetition. Second, the formation of stable, optimum units may be strengthened by another process, aimed at preventing the formation of concurrent, and unfrequent, associative links. One of the major developments in the study of conditioning in recent decades has been the discovery that associations between contiguous events tend to be selective and mutually exclusive. For instance, if two events A and B are paired with C, but, for some reason (e.g. saliency, position, relative number of pairing) A is a better predictor than B, then the association of B and C is impaired (e.g. Kamin 1969). This phenomenon is referred to as overshadowing in the conditioning literature. Consider now that in a string of five events (A, B, C, D, E) the more economical parsing, in the sense defined above, is AB/CDE. This means that in the whole series of training items, AB and CDE occur more frequently than, say, ABC and DE. Association principles account for the formation of AB and CDE associations, but also for the neglect of the less frequent associations, because the associations of B with A and of C with D would prevent the formation of the BC associations. This effect can again be interpreted as an attentional effect, the most frequent associations capturing all the available attention (e.g. MacKintosh, 1975).

6.4.2 Implicit learning shapes the phenomenal experience of the world

According to the dominant view, implicit learning generates implicit knowledge or implicit representations. Although there is not enough space here to analyse these aspects in detail (for interesting suggestions, see Berry and Dienes 1993, Chapter 8), it seems to be commonly accepted that the notions of knowledge and representation refer to some kind of cognitive database, that demands the mobilization of certain mechanisms or processes in order to be developed and used. This database is usually conceived as a pool of symbolic information, which may be progressively incremented, and whose use is optional. If one subscribes to this widespread view, the units issuing from the action of associative processes are implicit internal representations, which may be used by all-purpose, fixed mechanisms. Our proposal is, in some real sense, the exact opposite. We contend that these units are fully conscious, but that they do not share, at least initially, the properties of internal representations.

We claim that implicit learning modifies the way the data are *consciously* coded and perceived. This contention is a direct consequence of the mode of formation of the units, through the engagement of attentional processes. The units are in fact defined by the content of our conscious attentional percept. For instance, after exposure to a corpus generated by a finite state grammar including RXR as a frequent recursion (see Table 6.1b), subjects no longer perceive 'R' and 'X' as two familiar but separate entities, but perceive 'RXR' as an increasingly familiar unit. As a result, when faced with the string MTVRXR, subjects spontaneously encode MT/V/RXR, instead of, say, MTV/R/XR. The way the data are coded is changed, and this change directly affects our phenomenal experience. The process may be better understood intuitively in the context of natural situations, so it may be useful for the sceptical reader to reflect for a moment on his or her preferred prototypical situation of implicit learning, whether this situation concerns the acquisition of first or second language, natural categories, reading and writing abilities, or even sensitivity to musical structure. It is hardly defensible that our subjective experience of that part of the environment with which we interact in each of these cases remains unchanged while training progresses. On the contrary, what changes is precisely the way we consciously perceive and interact with the environment.

The changes in phenomenal experience and the observed improvement in performance can not be dissociated. They are two facets of the same phenomenon. The phenomenal experience is the experience of a subject performing a task, and the performance is a function of the conscious processing of the situation. More specific hypotheses depend on the behavioural register tapped by the implicit test. If we consider the prototypical grammaticality judgements, one possible explanation is that subjects interpret, more or less automatically, their level of perceptual fluency as an indicator of grammaticality. Strings which can be easily read because chunks of letters are directly perceived as familiar units, would tend to be judged as grammatical.

The perceptual units involved here must be distinguished from internal representations on at least two key points. Firstly, there is no sense in claiming that these units are 'used' whenever we deal with a task. The effects of processing an earlier episode are direct, mandatory, and not reversible. This means that it is quite impossible to deal with the incoming information in the same way before and after exposure to training items. The coding mode is changed in a way that cannot be short-circuited by any higher-order coding mechanisms, because such unchanging coding mechanisms do not exist. Secondly, these units are not internal representations because they are linked to the presence of the sensory input. Like any other conscious content, these units are accessible and may be incorporated into logical reasoning, decision-making, and other forms of controlled thinking, provided, however, that the external data are present.

6.4.3 The phenomenal experience of the world generates explicit knowledge

It is crucial to distinguish between the changes in our conscious experience of the outside world, which we have just discussed, and the elaboration of a

knowledge base about the world, which is the meaning attributed to the notion of explicit knowledge in the implicit learning area. Processing grammatical letter strings may lead to RXR becoming a familiar unit shaping subjects' conscious perception of the strings which contain this unit, but this effect must not be confused with the ability to recall that the letter strings displayed during a prior study phase comprised RXR as one component, or to recognize RXR in similar new items. Explicit knowledge is the database that is tapped by explicit postexperimental recall or recognition tests.

In order to become available in explicit memory, perceptual units must be transformed into internal representations. Perceptual units are linked to the presence of the sensory input. They exist only during the act of perceiving. With training, these units gain in autonomy, and can be evoked outside of their initial context. This decontextualization is presumably progressive, so that less and less context is needed for the evocation of a given unit. This progressive decontextualization fits well with the observation that unit knowledge is more readily available in recognition tests, in which a large part of the original context is preserved, than in recall test, in which representations must be evoked without perceptual support. Note that the scenario described here is in no way specific to the data presented in implicit learning studies: it simply corresponds to the ubiquitous principles of memory formation, in which the data-driven conscious percept becomes conceptually driven conscious memories.

This account predicts a close parallel between performance improvement and explicit knowledge about the training situation, provided that the Shanks and St John (1994) information criterion is fulfilled; that is, provided that the tests tap relevant componential knowledge. Indeed, performance is dependent on the change in conscious perception, which is also the source of internal, explicit representations. As noted above, this parallelism is indeed the most common experimental outcome and the data illustrating it provide the main direct experimental support for the present account. However, there are cases in which this parallelism breaks down, essentially in neurological disorders. How is this possible?

It must be noted that the observation of improved performance without explicit memory of the situation does not invalidate our account, since we do not maintain that performance is causally dependent on explicit knowledge. We have repeatedly claimed that such a view is unwarranted (Perruchet *et al.* 1990; Perruchet and Gallego 1993). Performance is linked to conscious perception, as shaped by unconscious associative processes, not to conscious mental representations. We contend that the attentional processing of the displayed material permanently modifies the subsequent conscious perception of related material as a mandatory consequence of the processing itself, *in amnesics as in normals*. In normal subjects, the attentional processing of an episode also triggers the ability to remember this episode explicitly. What is lost when neural lesions are present is the ability to form internal, autonomous representations from a conscious percept. However, this loss has no detrimental consequences on performance, because phenomenal experiences do not rely on these explicit representations.

6.4.4 *Implicit knowledge and consciousness*

In this scenario, what becomes of the notion of implicit knowledge or implicit representation, which is generally construed as the cornerstone of modern cognitive science? This construct simply becomes superflous. We contend that the implicit effects of earlier episodes are due to the direct modification of the processing mechanisms themselves, in accordance with associative principles. The involvement of the processes mobilized to cope with the demands of the task modifies these processes, as a 'mechanical' consequence of their involvement (Kolers and Roediger 1984). There is no place for implicit representations in the mind: there are only unconscious (associative) processes, which shape conscious contents and, in consequence, conscious representations.

Refuting the existence of unconscious representations, and hence of genuine unconscious cognition, leads to consciousness being given its full evolutionary significance. Consciousness becomes the only support of any process which manipulates knowledge. The scope of this chapter is limited to the standard experimental settings of implicit learning, in which the task demands make the explicit representations of the situation functionally irrelevant. However, our conception is fundamentally opposed to an epiphenomenalist stance. Let us suppose that the ultimate task implies some form of reasoning or abstraction based on the study material. Suppose, for instance, that a subject is faced with a set of strings generated by a finite state grammar, but that his or her task is to draw up a schema of the grammar. Now, the piecemeal information available to consciousness needs to be analysed. It is interesting to note again that the associative processes, by shaping the representation of the material, are the ideal preparation for a later analytical, problem-solving approach. Indeed, by isolating the most frequent units, the associative processes recode the material in a way that tends to make its deep structure directly available for controlled processing.

6.5 FACING THE UNIT FORMATION FRAMEWORK WITH ALTERNATIVE THEORIES

We must now to compare the present framework with competing models of implicit learning. This section is not intended to present an exhaustive discussion of the alternative theories, but only to stress some crucial points of divergence between our framework and the most developed alternative theories. We will see that the notion of abstraction is a recurrent issue. We will take this opportunity to examine in what sense forming new units through associative processes performs abstraction.

6.5.1 *The abstractionist theory*

According to the first historical interpretation of implicit learning (Reber 1967, 1989), subjects faced with a structured situation extract its deep regularities in order to store in memory a small number of ready-to-use abstract laws, and apply them whenever they may be useful in coping with a new situation. In other words, subjects encode the grammar, not the product of the grammar. Although this abstract coding demands that attention is allocated to the study material (Reber 1993, p. 95), it is performed by an autonomous, unconscious processor.

The abstract format of the representations generated during the study phase makes the performance independent of the specific material used in the experiment. Two forms of independence have been tested (Reber 1989, 1993). The first concerns the independence of the internal representations with regard to the specific sample of the products of the grammar that has been displayed to subjects. In support of an abstractionist standpoint, Reber and Lewis (1977) showed that subjects learning an artificial grammar seemed to be more sensitive to the frequency of the bigrams composing the full set of strings generated by the grammar than to the frequency of the bigrams composing the subset of strings displayed in the study phase. The second form of independence is related to the specific material in which the grammar is instantiated. The argument here is based on changed-letter transfer procedures. In a changed letter transfer procedure, the letters composing the study items are changed in a consistent way in the test (e.g. M is always replaced by Q, V by Z, and so on). The letters may also be replaced by tones of different frequencies (Altmann *et al.* 1995). Reber (1969), and several subsequent studies, show that subjects' performance exceeded a random response level in these conditions, although their performance is considerably worse than under standard test conditions. Reber and others claim that these results prove that subjects are able to abstract the 'syntax' of the displayed material, independently of the 'vocabulary'.

Our account stands in sharp contrast to an abstractionist perspective. Subjects are assumed to be sensitive only to the specific product of the generative grammar. There is no abstract, 'ready-to-use' knowledge base. Instead, the coding of the incoming information is modified as a function of the material previously processed, because processing any event facilitates the later processing of the same or similar event. Attention to the material is needed for learning, but, instead of being a mysterious condition required for the operation of a mysterious autonomous, unconscious processor, it shapes what is learned: subjects learn what they attend to.

How can our own framework encompass the experimental results put forward in support of the abstractionist theory? The first argument outlined above concerned Reber and Lewis' result that subjects are more sensitive to the properties of the whole set of strings that a grammar may generate than to the subset of strings actually presented to them. Perruchet, Gallego, and

Pacteau (1992) demonstrated that this result was entirely due to a set of biases inherent in the procedure.

The second argument is based on the changed-letter transfer procedure, and warrants close examination. The empirical reliability of the phenomenon has been questioned (e.g. Perruchet 1994*a*), because the effect is generally tenuous, and adequate methodological controls were lacking. However, recent studies (Altmann *et al.* 1995) have established more convincing empirical roots for the phenomenon. We acknowledge that changed-letter transfer is a reliable phenomenon, and is indicative of abstract processing. However, it is possible to encompass the phenomenon within a non-abstractionist framework.

We argue that no abstraction occurs in implicit conditions, that is in the conditions defined above, when the procedures that mediate the behavioural adaptation are not under conscious control. But, of course, we do not deny the existence of any form of abstract processing. For instance, Turner and Fischler (1993) and Whittlesea and Dorken (1993) showed that subjects can learn the abstract rules of the grammar underlying strings of letters when they are instructed to search for rules, or when they are given incidental instructions which guide them towards the structure of the material (see also Shanks and St John 1994, for additional evidence for abstraction in other areas of research). The point is not that abstract processing never occurs, but that this mode of processing requires explicit thinking.

There is some evidence that the processing involved during transfer is in fact intentional and explicit. In all published studies including the most recent ones (e.g. Manza and Reber this volume), subjects are informed of the rule-governed nature of the material before testing, and they know that they have to assess the well-formedness of the new items with regard to these rules. These instructions inevitably shift subjects to a rule-discovery mental set. The role of explicit reasoning in changed-letter transfer is suggested by the fact that transfer is performed better when the training session involved intentional (i.e. rule searching) rather than incidental instructions (Mathews *et al.* 1989). Whittlesea and Dorken (1993) failed to obtain changed-letter transfer in subjects whose attention was not focused on the structure of the situation. In the same vein, Gomez (submitted) showed that changed-letter transfer occurred only in subjects who have sufficient explicit knowledge of the rules. All these studies strongly suggest that transfer performance depends on the involvement of conscious and deliberate processes.

This conception raises a problem. Explicit abstraction at transfer time necessarily relies on the explicit memory of the material displayed during the training phase. If the available knowledge base essentially consists of bigrams and trigrams, it could be argued that this basis is too tenuous to infer rules which are of potential relevance for other material. In fact, there are both empirical and theoretical arguments that run counter to this objection. For instance, Mathews *et al.* (1989) show that transfer to a different letter set operates equally well in yoked subjects who learned from verbal instructions given by experimental subjects as in the experimental subjects themselves. The point is that most of these verbal instructions tell subjects to select or avoid specific bigrams or trigrams. At a theoretical level, Redington and Chater (1996) have shown how the abstract information which is relevant for transfer can be abstracted from componential knowledge. They showed for instance that a 'toy model' postulating that subjects learned only initial and final trigrams predicted a level of performance in a subsequent changed-letter transfer session that matched or exceeded that really observed.

Before concluding the discussion of the abstractionist position, we should mention a final issue. Reber (e.g. 1993) and others (e.g Mathews, this volume) have claimed that the abstractionist position accords with the principles of evolutionary biology. In our view, the abstractionist position runs counter to the evolutionary principles on two major points. Firstly, it is quite paradoxical to assume that (phylogenetically and ontogenetically) primitive mechanisms rely on sophisticated abstraction processes. Our reliance on associative mechanisms fits undoubtedly better with the common conception of evolution. Secondly, the abstractionist standpoint sees no role for consciousness in adaptation. Although we subscribe to the view that 'consciousness is a late arriver on the phylogenetic scene' (Reber 1993), it is also arguable that this arrival has been associated with deep modifications in the way living organisms adapt to their environment. If unconscious processes perform the same tasks as conscious processes and do so more efficiently (e.g. Lewicki *et al.* 1992), why should the appearance of consciousness seemingly coincide with improved adaptation?

6.5.2 The episodic account

The first historical alternative to the abstractionist position in the field of artificial grammar learning is the so-called instance-based or memory-based model proposed by Brooks (Brooks 1978, Vokey and Brooks 1992). In Brooks' model, subjects who are shown grammatical strings during the study phase store the strings in memory, without any form of condensation or summary representation. During the test phase, they judge for grammaticality of test strings as a function of their similarity to the stored strings. The instance-based model works because, if no special care is taken to generate the material, grammatical test items tend to look globally more similar to study items than ungrammatical test items.

In the first version of the model, similarity was assessed in terms of the common surface features of a test string and a single study string. However, this version has been shown to be unable to account for the subjects' overall performance when the material teases apart grammaticality and surface similarity. This limitation has been clearly shown in the Vokey and Brooks (1992) studies. As a part of their nicely designed experiments, Vokey and Brooks presented different test items that were equally similar to a specific study item (for instance, one study item was MXRTMVR while the tests items were

MXRTMXR and MXRTMTR), but differed in their grammaticality. The letter of the test items that differentiated these items from a specific study item either obeyed, or violated, the finite state grammar. If performance relies exclusively on surface similarity, the manipulation of grammaticality should have no effect. This was *not* the case in four independent experiments (see also McAndrews and Moscovitch 1985). In fact, manipulating grammaticality (while keeping the specific similarity constant) affected performance to approximately the same extent as manipulating specific similarity (while keeping grammaticality constant). In order to explain these results within a memory-based model, Vokey and Brooks (1992; see also Brooks and Vokey 1991) claimed that the similarity may be computed with the whole set of study items (Vokey and Brooks 1992), and may take account of certain abstract properties of the strings (Brooks and Vokey 1991). However, unlike the abstractionist model, there is no assumption that an abstract knowledge base is built during the study phase. Abstract processing is performed online during the test to cope with the specific demands of the situation.

The instance-based model has recently been developed into a so-called 'episodic account' of implicit learning (Neal and Hesketh 1997; Whittlesea and Dorken 1993). The basic idea is still that each instance is stored as a separate trace. However, the episodic account better integrates a number of ideas to emerge from the memory research of the last two decades. To cut a long story short, subjects no longer store an instance but process an episode (i.e. an instance in its context). Researchers now emphasise the fact that the memory trace is no more than the by-product of the processing operations engaged during study, and that retrieval depends on the overlap between the processing undertaken during the study and the test phase. Support for this view stems from the demonstration of the encoding specificity effects in research into artificial grammar. For instance, Whittlesea and Dorken (1993) show that performances in the test phase are better if the processing involved during the test (pronouncing or spelling the letter strings) matches the processing involved during the study phase.

Our framework has much in common with an episodic account. We also subscribe to a proceduralist view of memory, so that we have no need to reinterpret the majority of the arguments put forward in support of this view, such as those of Whittlesea and Dorken referred to above.

The first point of departure between our standpoint and an episodic account concerns the transformations of the external input during its coding by the subjects. In our view, the physical input is not integrated as such, but instead captured through the filters inherent in the properties of the attentional perceptual processes. We claim that subjects code sequential material as a succession of disjunctive chunks, as a mandatory consequence of the limitations of perceptual processes. In the field of artificial grammar learning, these divergencies are revealed by the fact that the episodic account assumes that subjects store a holistic representation of the displayed items, whereas

we posit that subjects parse the strings into successive fragments (note that the relevance of the distinction between whole item and fragments may be limited to the artificial grammar context. It is conceivable that in a different area, the displayed instances are smaller, and that a unit formation account considers units which are equal or larger in size than an episodic account. The point is that the size of the processing units in our account is determined by the constraints inherent in attentional processing).

Is there independent evidence that subjects process unitary representations of long strings of consonants? For instance, the episodic account postulates that a string such as 'MXRTMVR', presented among a long set of other strings composed of the same letters, can be stored in memory and subsequently judged as being similar to 'MXRTMXR' (these examples are taken from Vokey and Brooks 1992). The counter-intuitive nature of this assumption can not itself be thought of as a shortcoming, given that unconscious processes do not need, by definition, to be intuitively comprehensible. However, we should emphasize that there is no cross-domain support for this postulate. In implicit memory research, the items are generally words, for which there is evidence that literate subjects have formed unitary representations before the experiment starts. When non-words are used in this context, they generally include both vowels and consonants and look like real words. Consequently, it is likely that they make contact with pre-existing linguistic representations (see Dorfman 1994 for arguments). The few attempts to demonstrate priming with non-legal non-words have generally failed (but see Bowers 1994).

There is a second crucial difference between an episodic account and ours. An episodic account postulates that each episode is processed separately, while we assume that there is some process which accomplishes pooling across episodes. This departure is partially a consequence of the differences in the size of the processing units. If we consider full episodes including, for instance, a complete string of letters in its context, it is improbable that the very same episode will subsequently appear. At the same time, some small letter chunks typically occur many times during a training session. Our model assumes that the processing of these chunks is progressively modified. Simply applying basic associative processes leads to the selection of a few chunks frequently displayed, which structure the coding of new strings.

This discussion raises the question of knowing whether a unit formation account postulates abstraction. On the one hand, processing fragments of a global entity may be thought of as a form of abstraction (Mathews 1990). This is obviously formally correct. However, in our opinion, this claim would have psychological significance only if the processing of chunks results from analytical operations performed in order to extract componential information from the whole episode, which is conceived as a single primitive entity. If subjects necessarily encode the percept as fragments because of the limitations of their perceptual and attentional coding mechanisms, it seems psychologically irrelevant to refer to abstraction here. Units are not abstracted from large entities. Instead they are built through the associations of their primitive components.

On the other hand, it has been argued (Vokey and Brooks 1994) that a model such as ours postulates abstraction, because structural sensitivity is identified with the effects of the frequency of componential information on performance, and therefore relies on some kind of pooling across items. Vokey and Brooks are correct when they emphasize that our model relies on pooling mechanisms. However, they seem to confuse the idea that one aspect of abstraction implies pooling across items (a contention with which everybody would agree) with the reverse, namely that all forms of pooling necessarily imply abstraction. In fact, sensitivity to the frequency of stimuli is observed at all levels of the ontogenetic and phylogenetic scales. Sensitivity to the frequency of the stimulus can even be observed at the cellular level. The phenomenon appears to be an intrinsic characteristic of any learning and memory process, and claiming that it requires an abstractive mechanism deprives the term abstraction from any specific meaning.

We conclude that a unit formation account does not rely on abstractive mechanisms, whereas an episodic account necessarily relies on certain abstractive mechanisms in order to account for performance (see above). The parsimony of our account is illustrated by the reanalysis of the Vokey and Brooks (1992) data by Perruchet (1994*b*). It should be recalled that Vokey and Brooks dissociated the normally undifferentiated factors of specific similarity and grammaticality, and showed that both factors influenced performance. They concluded that both abstraction (viewed as a retrieval pooling process) and assessment of global similarity are involved in artificial grammar learning. Perruchet (1994*b*) showed that exclusive reliance on componential information – the demonstration itself involved the initial and terminal trigrams of letter strings – was able to account for the overall variance in performance elicited by the two factors manipulated by Vokey and Brooks. It should be remembered that this explanation relies exclusively on ubiquitous associative processes, and therefore in no way postulates any mysterious, powerful mechanism which supplies knowledge of which subjects have no introspective evidence, such as unconscious knowledge of abstract rules (e.g. Reber 1993) or the unconscious holistic representation of the external input (e.g. Higham and Vokey 1994).

6.5.3 A note on computational modelling

An associative theory of learning obviously accords better consonance with connectionist models than with symbolic models. An influential connectionist model of implicit learning was developed by Cleeremans (Cleeremans 1993; Cleeremans and McClelland 1991). At a general level, the ability of connectionist models to account for empirical data provides support for an associative model, insofar as both models are based on the progressive growth of associative links between elementary primitives. However, some remarks are in order. Firstly, beyond the fact that they agree with general associative principles, connectionist models may rely on postulates that contradict the basic postulates of the present account. In our opinion, this is the case with Cleeremans' simple recurrent network model in which, whatever the location in the sequence, predictions of the next event are derived from the few prior events, each event being successively target and predictor. This procedure is related to the notion of overlapping units, which contrasts with the ideas of disjunctive units postulated here. Secondly, non-connectionist models may rely on principles that accord better with certain aspects of the present account than a connectionist approach. The competitive chunking model proposed by Servan-Schreiber and Anderson (1990), in which chunks are disjunctive fragments of the items, is highly revealing on this point (however, the chunking model relies on a process of analytical fragmentation of a primitive entity that is, in some ways, the opposite of the present, associative account). Third, non-connectionist models starting from postulates which are starkly opposed to the present account have also been proved to be able to simulate human performance reasonably well (e.g. Ling and Marinov 1994).

These remarks are not intended to deny the interest of the computational approach, but rather to suggest that computational models ought to be tested in situations that are designed to contrast predictions of different models, in the same way that human subjects are tested. Simply showing that a model is able to learn from a structured situation does not tell us anything about the psychological mechanisms involved in the task, because widely divergent models all simulate performance improvement. An interesting illustration of a more fruitful approach was presented by Cleeremans (1994), who subjected his model to a transfer paradigm intended to compare predictions issuing from a rule-based theory on the one hand, and a memory based theory on the other (the paradigm was used on human subjects by Perruchet (1994*a*)).

6.6 A TENTATIVE GENERALIZATION OF A UNIT FORMATION ACCOUNT

So far our analysis has concerned situations of repeated sequence learning and artificial grammar learning. In this section, we will show first that our model is also relevant for the third situation commonly used in laboratory studies, namely process control tasks, provided that the notion of the subjects' unit of processing is enlarged to include certain components of the subjects' own actions.

A model whose area of application is limited to these experimental situations would have only limited heuristic value. The problem does *not* stem primarily from the specificity of the physical basis for the rules structuring the situations. For instance, it seems that the results obtained with finite-state grammars which use consonant letters as data, are generalizable to situations involving spatial positions (Cleeremans and McClelland 1991), familiar noises (Howard and Ballas 1982), geometrical figures, or even tones of various frequencies (Altmann *et al.* 1995). The problem is, in fact, a result of the nature of the rules themselves. In the three standard implicit learning paradigms, the structure can be described in terms of a set of first-, second- or third-order dependency rules. Of course, such rules may have some general applicability. However, it is worth examining whether our account can be generalized to cases where the structuring rules are no longer sequential.

6.6.1 When rules extend to subjects' own actions

In the process control tasks, first investigated by Berry and Broadbent (e.g. 1984), subjects are required to reach and maintain specified target values of an output variable by varying one or two input variables. Output and input variables are related by a linear equation including the current state of the system as one term. These tasks differ from the situations examined above in at least three key ways.

1. The system is regulated by a single, abstract rule (a linear equation), instead of being regulated by a large set of sequential dependency rules. However, this specificity turns out to be quite superficial. Although condensed into a linear equation, the structure embedded in the control tasks may be described alternatively as a set of simple pairwise associations between the current state of the system (or in some case the state of the system on trial n–1) and the action to be performed to reach the target value. In the context of the control tasks, this kind of information has been described as a look-up table. A look-up table comprises the set of the pairwise relations needed to control the system, and using a look-up table consists of performing the same action in response to new situations as in earlier successful trials. There is now compelling evidence that subjects in fact abstract no rules, and perform as described by a look-up table model (e.g. Marescaux *et al.* 1990; Dienes and Fahey 1995). The only exception occurs when the relations are determined by very simple and salient rules. In this case, subjects have been shown to be able to acquire and use abstract information (e.g. Dienes and Fahey 1995). However, this finding does not detract from our view of implicit learning, since, a priori, it is very likely that such performance is intentionally driven and consciously mediated.

2. Control tasks are interactive in the sense that the events presented to subjects depend on their own responses to prior events. Undoubtedly, this feature is present in many natural situations, and extending the validity of our model to these cases is a first step towards demonstrating its relevance in natural situations. To account for these cases, we need to adapt certain aspects of our earlier proposal. Indeed, we argue above that implicit learning shapes the way subjects *perceive* the world. In interactive tasks, subjects are not only perceiving a fixed input, they are also acting on the world. To integrate this fact into our account, we must accept that subjects' units may comprise subjects' actions as one component. In fact, this generalization raises no fundamental

problems. We ground unit formation on associative learning principles, and it is well known from behaviourist studies that pairwise associations may include subjects' behaviour. The distinction here exactly parallels the distinction between classical and operant forms of conditioning, and it is commonly taken for granted that the two forms of conditioning obey the same laws. In order to take account of these aspects, we may argue, more generally, that implicit learning shapes our phenomenal *experience* of the world. There the word 'experience' includes both the passive exposure to, and the active interaction with, the environment.

3. The third characteristic of the process control tasks is that subjects are engaged in goal-directed behaviour. In fact, this property raises no generalization problem at all. Our initial definition of implicit learning phenomenon did not require that subjects are diverted from the real objective of the task. On the contrary, it seems that asking for goal-directed behaviour ensures the attentional processing of the situation, a condition that we construe as necessary for the occurrence of implicit learning. An apparent contradiction to the necessity of attentional processing should be noted here. Hayes and Broadbent (1988) showed that performing the control task under dual task conditions had a positive effect on performance, with regard to normal, full attention training conditions. However, these results, although still advanced in support of a strong view of implicitness (e.g. Seger 1994), have not been replicated. The fact that divided attention conditions of training *impair* performance in this task (as in others) is now acknowledged (Green and Shanks 1993).

It is necessary to make a final comment about control tasks. Berry and Broadbent (1984) demonstrated an inverse relationship between the ability of subjects to describe the system adequately and their level of performance in the control task. Subjects who performed worse were more precise in their verbal reports, and vice versa. It comes as no surprise that these data are often put forward as compelling evidence for the dissociation of implicit and explicit knowledge (e.g. Seger 1994). In any case, they are not consistent with our account: if an explicit knowledge base results from the interpretation of a change in processing mode, then subjects whose performance improves the most should have the greatest opportunity to develop an explicit knowledge base. Although the dissociation observed by Berry and Broadbent (1984) is empirically reproducible, its interpretation has been reconsidered by Buchner *et al.* (1995). Buchner *et al.* showed that subjects tended *either* to explore all the possible transitions of the system and hence acquire extended explicit knowledge about it while performing poorly, *or* to keep the same response throughout the training session, hence ensuring good performance, but thereby depriving themselves of the opportunity to learn about other action/outcome relationships. This occurs because, in all the tasks used in this context, the best strategy to optimize performance is to respond in the same way throughout the session. Therefore, the initial evidence for an explicit/implicit dissociation turns out to be induced by the task itself, not by the subjects' processing system. The available evidence (e.g. Marescaux and Karnas 1991; Sanderson 1989) suggests that performance improvement is generally accompanied by the explicit knowledge of the relevant associations.

6.6.2 When rules are no longer sequential

Virtually all the knowledge needed to accomplish repeated sequence learning tasks, artificial grammar tasks, and control process tasks can be described, respectively, as a set of associations between a small number of spatial locations, associations between two or three letters, and between a response and its related outcome. The reason is that, in every case, the structuring rules can be (almost exhaustively) described as nth order dependency rules, where n varies between 1 and 3. It is no surprise that the formation of small units captures most of these structural regularities. This deep identity of structure has damaging consequences when assessing the relevance of the results obtained from implicit learning research to real-world situations. Fortunately, the rules employed in a small number of studies are not low-order sequential dependency rules. How well does our account work in these cases? The response is: surprisingly well! Let us consider two cases in turn.

The first case (Kushner *et al.* 1991; Reber 1993, p. 142) is especially interesting, because the rules that structure the situation are relational. Subjects viewed five stimuli displayed in rapid succession, with each stimulus appearing in one of three possible locations, arranged to form the vertices of an invisible triangle, ABC. The subjects were asked to predict where the sixth stimulus would appear, and were then informed of its correct location. This location was determined on the basis of the relation between the location of the second and fourth stimuli of the initial sequence. If these stimuli appeared at the same location, the sixth stimulus appeared in location A. If the relation between them was clockwise, the sixth stimulus appeared in location B, and if the relation was counterclockwise, the sixth stimulus appeared in location C. As shown by Clark and Thornton (in press), the problem posed by this kind of situation can not be solved through exploitation of observable regularities in the input data as originally coded. In order to make the problem tractable, the original input logically needs to be recoded, and this operation involves some abstractive process.

Perruchet (1994*a*) replicated the main result of Kushner *et al.* (1991), namely that subjects significantly improve the accuracy of their predictions across (somewhat extensive) training in the task. However, Perruchet also demonstrated, thanks to the use of a transfer design, that subjects learned without abstracting the relational rule in any way. The results were compatible with the hypothesis that subjects were able to form direct associations between the first set of stimuli and the correct outcome. The number of primitive components integrated in this associative network is a matter for further research. It is unclear, in particular, whether subjects extracted the position of the relevant items (that is, second and fourth) or processed the five-trial sequence as a whole (see also Cleeremans 1994). But what the study did show is that performance improvement can be accounted for by the formation of cognitive units through purely associative processes, in a situation where the rules are relational.

The second paradigm considered here was proposed by Lewicki *et al.* (1988). Subjects were submitted to a four-choice reaction time task, with the targets appearing in one of four quadrants. In each of the series of five-trial blocks, the first two locations of the target were randomly distributed, and the last three locations were determined by complex rules such as: 'When the third and the fourth locations in a five-trial logical block describes a horizontal movement, then the movement described by the fourth and the fifth locations is diagonal'. Although subjects were unable to verbalize the actual nature of the manipulation and, crucially, had no explicit knowledge about the partitioning of the material into logical blocks, performance on the last trials of each block improved at a faster rate and was better overall than performance on the first trials. The trouble with these results is that, although the rules are in fact second-order dependency rules and hence explicable in associative terms, they require the segmentation of the whole sequence in a succession of five-trial subsequences. As in the previous example, this operation involves a recoding of the input data that can *not* be accounted for by associative processes, because there is no property evident in the distribution of the conditional probabilities of individual events that would allow such units to emerge.

Perruchet *et al.* (1990) demonstrated that subjects learn from the task, without segmenting the sequence into logical blocks. Instead, they were sensitive to the relative frequency of small units, consisting of two or three successive locations. Some of the possible sequences of two or three locations were more frequent than others, because the rules determining the last three trials within each five-trial block debarred certain transitions. For instance, the rules precluded the back and forth movements of the target. In consequence, the other movements were more frequent *over the whole sequence*. The crucial point is that the more frequent events, which presumably elicit shorter reaction times, tended to be located in those trials that could be predicted by the underlying rules. This does not result not from an unfortunate bias in randomization, but from a logical principle: the rules determine both the relative frequency of certain events in the overall sequence and the selective occurrence of these events in specific trials. The validity of this interpretation was tested by deriving specific predictions relating to specific features of fine-grained performance from an abstractionist model on the one hand, and from our alternative model on the other. Empirical data, confirmed by a connectionist modelling approach (Cleeremans 1993), unambiguously supported the latter interpretation. In addition, the frequency distribution of the small relevant units was available to subjects' consciousness, as revealed by a subsequent prediction task.

These studies have crucial implications. Reber (1993) endorses the validity of our reappraisal of many results initially adduced as support for an abstractionist position (and notably our reappraisal of the Lewicki *et al.* results). However, according to him, our criticisms address only, to quote, 'a slightly embarrassing methodological glitch' (Reber 1993, p. 115) because, he argues, subjects are still engaged in implicit learning in these experiments. We agree with this latter statement. That, however, is not the point. The point is that we must choose between two opposite interpretations of the implicit learning of complex structures. Let us take the Lewicki *et al.* situation as a case in point. In the first interpretation, subjects unconsciously abstract the structure of the sequence into logical blocks, an operation involving an incredibly sophisticated autonomous processor (we suggest that skeptical readers try to design a general-purpose algorithm intended for the discovery of the kind of logical block involved in Lewicki *et al.*'s task), then they unconsciously learn the second-order sequential dependency rules embedded within the logical blocks. In our account, subjects form units from some salient subsequences of events, such as the back and forth movements of the target, and become sensitive to the relative frequency of these units, with the result that some sequences of events are consciously experienced as more familiar than other sequences. The issue is not methodological. It lies at the root of our theoretical understanding of implicit learning.

This research gives us cause for optimism about the generality of a unit formation model of implicit learning. It shows that the formation of experiential units, and the sensitivity to their frequency of occurrence,[1] is able to account for performance in situations that are not structured by low-order sequential dependency rules, but instead by relational or other non-salient rules normally implying an abstract recoding of the input data to be captured.

6.6.3 *Learning out of the laboratory*

The question we must now ask concerns the validity of a unit formation account for learning in natural situations. Although an exhaustive, or even approximate, list of these situations is obviously impossible to draw up given the current state of our knowledge, we assume that any reader who does not adhere to a strict nativism will agree that implicit learning is active in at least some aspects of first language (e.g. Chandler 1993) and second language learning (e.g. Carr and Curran 1994), the formation of natural categories, the acquisition of reading and writing abilities, adaptation to the physical constraints of the world (e.g. Krist *et al.* 1993), sensitivity to musical structures, and many other phenomena which testify to successful adaptation to a complexly structured environment. It should be noted that most of these acquisitions take place during childhood, so that any discussion

1 Note that the Lewicki *et al.* situation presents a case in which the relevant information consists only of the relative frequency of units, not of the selective occurrence of a subset of these units. Indeed, in contrast to the situations described earlier, all the possible associations between two and three events occurred.

about the 'ecological' relevance of implicit learning theories is more or less coextensive with developmental issues. The implications of our account for the theories of child development are examined extensively in Perruchet and Vinter (in press).

In this chapter, we will briefly examine how the explanatory power of the present account may be considerably extended by considering the possibility it offers of hierarchical processing. Hierarchical processing occurs whenever the units built from a given set of primitives themselves become the primitives for higher level units. The possibility of hierarchical coding directly derives from our conception of implicit learning as being due to prior attentional processing, and shaping subsequent attentional processing. Once a new perceptual unit is formed from the integration of initially separate features, this unit can be processed as a primitive of a higher-level unit. We have not so far commented on this aspect, because it is not required to account for most experimental data. However, this aspect deserves to be emphasized in the context of very complex learning.

Let us consider reading as a case in point. We do not intend to deal here with the controversial issue of the nature of the linguistic units; our only postulate is that such units exist, and are hierarchically organized. Let us assume that the letters, the syllables, and the words are fundamental units in reading. Our proposal is that children exposed to a graphemic corpus primarily learn to associate primitive graphic features in order to build letters. The letters then become the new primitives for the formation of syllables, and so on. This means that the conscious perceptual experience proceeds from the perception of a subset of isolated features to the direct perception of words. It is worth stressing that the construction of intermediary units is necessary. Indeed, at any given level of learning, the probability of transition between the last primitives of a given unit and the first primitive of the subsequent unit is irrelevant. What is relevant is the probability of transition *between the new units*. For instance, it is clear that considering the probability of transitions between the last letter of a syllable and the first letter of the next syllable is meaningless: what is useful when the syllable is processed as an entity is the probability of transitions between syllables.

This possibility of hierarchical coding makes the explanatory power of a unit formation account far stronger than that of alternative accounts, although it relies on seemingly less powerful processes. For instance, the Cleeremans model (e.g. 1993) posits that subjects compute predictions, successively for any event of a sequence, from the set of prior events. At first glance, this way of processing the input is richer than forming small units, because forming units entails the loss of the probability of transitions between primitives belonging to different units. However, it should be borne in mind that what is lost is ignored in any case: children do not learn to read by becoming increasingly sensitive to the probability of transitions between, say, a subset of the primitive features of two consecutive letters. Likewise, the generation of a holistic representation of specific exemplars, as maintained by the supporters of episodic accounts, surely involves more sophisticated abilities than parsing the original material into small units. However, this processing mode, which relies exclusively on memory, renders the subject helpless in very complexly structured environment, because processing does not benefit from the hierarchical structure.

To conclude, the formation of units through low-level associative processes can be thought of as contributing to the general recoding mechanisms that make it possible to transform a set of data possessing barely detectable regularities into a tractable problem for research (Clark and Thornton in press). It is interesting to note that this functional recoding stems from a property of the attentional perceptual systems that can be construed as a limitation, since it constrains the size of new units to a few primitive elements. This kind of apparent paradox has been targeted in other contexts as 'the importance of starting small' considered by Elman (e.g. 1994) and the 'less is more' hypothesis advanced by Newport (1990).

These optimistic considerations do not imply that a unit formation account is able to explain all forms of learning. We believe that subjects learn only co-occurrences between those few primitive elements that are processed concurrently at any given moment. Learning will be possible insofar as the structure of the rules underlying the domain in question can be 'at least partially' captured by the nature (or the frequency) of the simple associations between the primitive components of the domain. Hence, in order to determine whether implicit learning really occurs, it is first necessary to analyse the task domain itself. Thus, the final word in this debate lies more properly with the specialists in each domain than with researchers primarily concerned with the experimental approach. When considering this issue, domain specialists should be aware that componential information may be relevant even in situations where the structuring rules have no apparent relationship with this kind of information (as shown, for instance, in our reanalysis of the Lewicki *et al.* data presented above).

It is obvious that purely associative mechanisms are unable to account for certain types of rule. For instance, we subscribe to the view, cogently advocated by Lockhart and Blackburn (1993), that implicit learning of probability is not possible. But, subjects learn probabilities. We contend that the processes involved in learning probabilities and, more generally, processes such as rule abstraction, hypothesis testing, and logical reasoning, are linked to conscious forms of thought. Although this question goes beyond the scope of this chapter, it can be seen that our account conveniently integrates the phenomena described in the implicit learning literature within more general thinking and reasoning processes. Indeed, associative mechanisms, by shaping the perception and the internal representation of the material, constitute an optimal preparation for a subsequent analytical approach. This happens because associative processes recode the material in a way that tends to make its deep (possibly hierarchical) structure directly available to conscious thought.

6.7 CONCLUSIONS

Implicit learning is defined as an adaptive mode in which subjects' behaviour is sensitive to the structural features of a previously presented situation, without this adaptation being due to the intentional exploitation of subjects' explicit knowledge of these features. Our account of implicit learning may be condensed into a few key points, which are listed below.

1. Subjects process complex material by parsing it into small and disjunctive (i.e. not overlapping) units. Under the effect of training, these units result in the optimal parsing of the material in the sense that, for a given size of unit, the material is segmented in a way that yields the smallest possible number of different units.
2. This phenomenon is a mandatory consequence of the attentional processing of the input data. Units emerge from the association of the primitive features that are jointly processed in the attentional focus. Optimal parsing is the end product of automatic associative phenomena.
3. Implicit learning generates conscious perceptual units, and not implicit representations. The change in coding directly affects the way the external world is attentionally perceived and processed. *Implicit learning shapes the phenomenal experience of the world.*
4. Conscious perceptual units, because they facilitate subjects' interaction with their environment, are directly responsible for the improvement in performance observed in conventional implicit learning tasks.
5. With training, conscious perceptual units become increasingly independent of the sensory input, and hence, form the explicit representations that are tapped by conventional tests of explicit knowledge.
6. With training, conscious perceptual units can form the new primitives of subsequent attentional processing, and can therefore be used to build higher level units. This possibility for hierarchical processing provides great explanatory power for learning in some very complex, real-world settings.
7. However, learning situations involving genuine abstraction cannot be accounted for by implicit processes. Abstraction requires conscious forms of thought. In general, however, by making possible the perception and the explicit representation of the relevant units (possibly in a hierarchical format), unconscious associative processes prepare us for an intentional analysis of the structure of the world.

When compared with alternative accounts, this framework provides a very different picture of the relationships between conscious and unconscious phenomena. In the conception of the mind underpinning most of the literature on implicit learning, implicit and explicit knowledge are two separate, presumably independent, forms of knowledge, but both forms share a common architecture. Indeed, both consist of mental representations which can be stored, manipulated, transformed, and so on. The main difference lies in the availability or the non-availability of these representations to consciousness. In our view, there is no place for implicit knowledge or implicit representations. The perceptual contents and the internal representations built from the perceptual contents are conscious. However, the processes shaping these conscious contents are unconscious.[2] These processes are intrinsically unavailable to consciousness, because they are purely physiological phenomena. These phenomena can be identified as associative mechanisms, whose very existence provides our most reliable knowledge of the functioning of the brain.

Although this conception posits a radical difference in nature between conscious and unconscious phenomena, it also gives us cause to view them as closely interconnected, undissociable categories of events. Unconscious associative processes shape the way we consciously perceive the outside world, and because subjective perceptual units are transformed into internal representations as they gain in independence from the sensory input, unconscious associative processes determine the way we conceive and consciously think about the world. In this perspective, the recurrent question about the relationships between conscious and unconscious phenomena is like asking whether the image perceived on the television screen is dependent on or independent of the variations in the electromagnetic voltages that guide the beam of the electrons. The two categories of phenomena are both functionally indissociable and intrinsically different. There is no gradual transition from unconscious to conscious phenomena, as claimed by Cleeremans (this book), and no possible 'redescription' of implicit knowledge as explicit knowledge, as claimed by Karmiloff-Smith (1992). What is gradual is the decontextualization process leading from conscious perceptual units linked to the presence of the sensory input, to conscious internal representation available with minimal contextual cues. Karmiloff-Smith's redescription is better viewed as an attributional inference on conscious contents (Perruchet and Vinter in press).

ACKNOWLEDGEMENTS

This work was supported by the CNRS and the University of Bourgogne. We thank Axel Cleeremans, David Shanks, and Annie Vinter for the useful suggestions they provided on an earlier version of this chapter.

REFERENCES

Altmann, G.T.M., Dienes, Z. and Goode, A. (1995). On the modality-independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **21**, 899–912.

Berry, D.C. (1994). Implicit learning: Twenty-five years on. In *Attention and performance XV: Conscious and nonconscious information processing* (ed. C. Umilta and M. Moscovitch) pp. 755–82. MIT Press, Cambridge, MA.

Berry, D.C. and Broadbent, D.E. (1984). On the relationship between task performance and associated verbalisable knowledge. *Quarterly Journal of Experimental Psychology*, **36**, 585–609.

Berry, D.C. and Dienes, Z. (1993). *Implicit Learning: Theoretical and empirical issues*. Lawrence Erlbaum Associates, Hove, UK.

Bowers, J.S. (1994). Does implicit memory extend to legal and illegal nonwords? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **20**, 534–49.

Brooks, L.R. (1978). Nonanalytic concept formation and memory for instances. In *Cognition and categorization* (ed. E. Rosch and B.B. Lloyd), pp. 169–215. Erlbaum Hillsdale, NJ.

Brooks, L.R. and Vokey, J.R. (1991). Abstract analogies and abstracted grammars: A comment on Reber, and Mathews *et al. Journal of Experimental Psychology: General*, **120**, 316–23.

Buchner, A., Funke, J. and Berry, D. (1995). Negative correlations between control performance and verbalisable knowledge: Indicators for implicit learning in process control tasks. *Quarterly Journal of Experimental Psychology*, **48**, 166–87.

Carr, T.H. and Curran, T. (1994). Cognitive factors in learning about structured sequences. *SSLA*, **16**, 205–30.

Ceraso, J. (1985). Unit formation in perception and memory. In *The psychology of learning and motivation* (ed. G. Bower), pp. 179–210. Academic Press, New York.

Chandler, S. (1993). Are rules and modules really necessary for explaining language? *Journal of Psycholinguistic Research*, **22**, 593–606.

Clark, A. and Thornton, C. (in press). Trading spaces: Computation, representation and the limits of uninformed learning. *Behavioral and Brain Sciences*.

Cleeremans, A. (1993). *Mechanisms of Implicit Learning: A connectionnist model of sequence processing*. Bradford Books, MIT Press, Cambridge, MA.

Cleeremans, A. (1994). The representation of structure in sequence prediction tasks. In *Attention and performance XV: Conscious and nonconscious information processing* (ed. C. Umilta and M. Moscovitch), pp. 783–809. MIT Press, Cambridge, MA.

Cleeremans, A. and McClelland, J.L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, **120**, 235–53.

Cohen, A., Ivry, R.I. and Keele, S.W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 17–30.

Dienes, Z. and Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 848–62.

Dienes, Z., Broadbent, D. and Berry, D. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **17**, 875–87.

Dorfman, J. (1994). Sublexical components in implicit memory for novel words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1108–25.

Dulany, D.E. (1961). Hypothesis and habits in verbal "operant conditioning". *Journal of Abnormal and Social Psychology*, **63**, 251–3.

Dulany, D.E. (1991). Conscious representation and thought systems. In *Advances in social cognition* (ed. R.S. Wyer and T.K. Srull). Vol. 4, pp. 97–120, Erlbaum, Hillsdale, N.J.

2 The idea that the processes are unconscious, while their products are available to consciousness, is far from new in psychology. Related claims have been made, from different perspectives, by Miller (1962), Neisser (1967), Mandler (1975), Nisbett and Wilson (1977), among others. More recently, Dulany (1991, 1996) has reflected on similar views.

Dulany, D.E. (1996). Consciousness in the explicit (deliberative) and implicit (evocative). In *Scientific approaches to the study of consciousness* (ed. J.D. Cohen and J.W. Schooler), pp. 179–212, Erlbaum, Mahwah, NJ.
Dulany, D.E., Carlson, A. and Dewey, G.I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, **113**, 541–55.
Elman, J.L. (1994). Implicit learning in neural networks: The importance of starting small. In *Attention and performance XV: Conscious and nonconscious information processing* (ed. C. Umilta and M. Moscovitch), pp. 961–888. MIT Press, Cambridge, MA.
Gomez, R.L. (submitted). How implicit is implicit learning?
Gomez, R.L. and Schvanevelt, R.W. (1994). What is learned from artificial grammar? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 396–410.
Green, R.E.A. and Shanks, D.R. (1993). On the existence of independent learning systems: An examination of some evidence. *Memory and Cognition*, 21, 304–17.
Hayes, N. and Broadbent, D.E. (1988). Two modes of learning for interactive tasks. *Cognition*, **28**, 249–76.
Higham, P.A. and Vokey, J.R. (1994). Recourse to stored exemplars is not necessarily explicit: A comment on Knowlton, Ramus, and Squire (1992). *Psychological Science*, **5**, 59–60.
Howard, J.H. and Ballas, J.A. (1982). Acquisition of acoustic pattern categories by exemplar observation. *Organizational Behavior and Human Performance*, 30, 157–73.
Jimenez, L., Mendez, C., and Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **22**, 948–69.
Kamin, L.J. (1969). Predictability, surprise, attention and conditioning. In *Punishment and aversive behavior* (ed. R. Campbell and R. Church), pp. 279–96. Appleton-Century-Crofts New York.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Bradford/MIT press, Cambridge, MA.
Knowlton, B.J. and Squire, L.R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 79–91.
Kolers, P.A. and Roediger, H.L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, **23**, 425–49.
Krist, H., Fieberg, E.L., and Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 952–66.
Kushner, M., Cleeremans, A., and Reber A. (1991). Implicit detection of event interdependencies, and a PDP model of the process. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, Chicago (IL)*. Erlbaum, Hillsdale, NJ.
Lewicki, P., Hill, T., and Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, 20, 24–37.
Lewicki, P., Hill, T., and Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist*, 47, 796–801.
Ling, C.X. and Marinov, M. (1994). A symbolic model of the nonconscious acquisition of information. *Cognitive Science*, **18**, 595–621.
Lockhart, R.S. and Blackburn, A.B. (1993). Implicit processes in problem solving. In *Implicit memory: New directions in cognition, development, and neuropsychology*, (ed. P. Graf and M.E.J. Masson), pp. 95–116. Lawrence Erlbaum Associates Hilldale, N.J.

Logan, G.G. and Etherton, J.L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**, 1022–50.
MacKintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276–98.
Mandler, G. (1975). Consciousness: Respectable, useful, and probably necessary. In *Information processing and cognition: The Loyola symposium* (ed. R. Solso), pp. 229–54. Erlbaum Hillsdale, NJ.
Manza, L. and Reber, A.S. (1997). Representing artificial grammars: Transfer across stimulus forms and modalities, this volume.
McAndrews, M.P. and Moscovitch, M. (1985). Rule-based and exemplar-based classification in artificial grammar learning. *Memory and Cognition*, **13**, 469–75.
McGeorge, P. and Burton, A.M. (1990). Semantic processing in an incidental learning task. *Quarterly Journal of Experimental Psychology*, **42A**, 597–609.
Marescaux, P-J. and Karnas, G. (1991). *The implicit versus explicit knowledge distinction revisited: When finding associations between verbalizable knowledge and some performance criteria*. Report 4PR3GK of the KAUDYTE project (ESPRIT BRA 3219).
Marescaux, P-J, Dejean, K, and Karnas, G. (1990). *Acquisition of specific or general knowledge at the control of a dynamic simulated system: an evaluation through a static situations questionnaire and a transfer control task*. Report 2PR2GK of the KAUDYTE project (ESPRIT BRA #3219).
Mathews, R.C. (1990). Abstractness of implicit grammar knowledge: Comments on Perruchet and Pacteau's analysis of synthetic grammar learning. *Journal of Experimental Psychology: General*, **119**, 412–16.
Mathews, R.C., Buss, R.R., Stanley, W.B., Blanchard-Fields, F., Cho, J.-R., and 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**, 1083–100.
Neal, A. and Hesketh, B. (1997). Episodic knowledge and implicit learning. *Psychonomic Bulletin and Review*, **4**, 24–37.
Neisser, U. (1967). *Cognitive Psychology*. Appleton-Century-Crofts, New York.
Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science*, **14**, 11–28.
Nisbett, R.E. and Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, **84**, 231–59.
Nissen, M.J. and Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, **19**, 1–32.
Perruchet, P. (1984). Dual nature of anticipatory classically conditioned reactions. In *Preparatory states and processes* (ed. S. Kornblum and J. Requin), pp. 179–98. Lawrence Erlbaum Associates, Hillsdale, N.J.
Perruchet, P. (1994*a*). 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. Umilta and M. Moscovitch), pp. 811–35. MIT Press Cambridge, MA.
Perruchet, P. (1994*b*). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 223–8.

Perruchet, P. and 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.

Perruchet, P. and Gallego, J. (1993). *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 1438–44.

Perruchet, P. and Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119, 264–75.

Perruchet, P. and Vinter, A.(in press). Implicit learning and development. In *Handbook of implicit learning*, (ed. M. Stadler and P. Frensch). Sage Publications, Thousand Oaks, CA.

Perruchet, P., Gallego, J., and Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology*, **22**, 493–516.

Perruchet, P., Gallego, J., and Pacteau, C. (1992). A reinterpretation of some earlier evidence for abstractiveness of implicitly acquired knowledge. *Quarterly Journal of Experimental Psychology*, **44A**, 193–210.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–63.

Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, **81**, 115–19.

Reber, A.S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, **2**, 88–94.

Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, **118**, 219–35.

Reber, A.S. (1993). *Implicit learning and tacit knowledge*. Oxford University Press, New York.

Reber, A.S. and Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition*, **5**, 331–61.

Redington, M. and Chater, N. (1996). Transfer in artificial grammar learning: A re-evaluation. *Journal of Experimental Psychology: General*, **125**, 123–38.

Reed, J. and Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Human Learning and Memory*, **20**, 585–94.

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.

Seger, C.A. (1994). Implicit learning. *Psychological Bulletin*, **115**, 163–96.

Servan-Schreiber, D. and Anderson, J.R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 592–608.

Shanks, D.R. and Johnstone, T. (in press). Implicit knowledge in sequential learning tasks. In *Handbook of Implicit learning* (ed. M. Stadler and P. Frensch). Sage Publications, Thousand Oaks. C.A.

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

Stadler, M.A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memory and Cognition*, **21**, 819–27.

Stadler, M.A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 367–447.

Turner, C.W. and Fischler, I.S. (1993). Speeded tests of implicit knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1165–77.

Vokey, J.R. and Brooks, L.R. (1992). Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **18**, 328–44.

Vokey, J.R. and Brooks, L.R. (1994). Fragmentary knowledge and the processing-specific control of structural sensitivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**, 1504–10.

Whittlesea, B.W.A. and Dorken, M.D. (1993). Incidentally, things in general are incidentally determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General*, **122**, 227–48.

Willingham, D.B., Greeley, T., and Bardone, A.M. (1993). Dissociation in a serial reaction time task using a recognition measure: Comment on Perruchet and Amorim (1992). *Journal of Experimental Psychology: Learning, Memory and Cognition*, **19**, 1424–30.

Wright, R.L. and Burton, A.M. (1995). Implicit learning of an invariant: Just say no. *Quarterly Journal of Experimental Psychology*, **48A**, 783–96.