2.32 Implicit Learning

P. Perruchet, University of Bourgogne, Dijon, France

2.32.1 Introduction

All of us have learned much without parental supervision and outside of any form of planned academic instruction, and more generally without any intentional attempts to acquire information about the surrounding world. Countless examples could be found in domains as diverse as first-language acquisition, category elaboration, sensitivity to musical structure, acquisition of knowledge about the physical world, and various social skills. All of these domains have several features in common. In particular, they are commonly described as governed by complex abstract rules by scientists, whether they would be linguists, musicologists, physicists, or sociologists. Also, learning in those situations mainly proceeds through the learner's exposure to a structured environment, without negative evidence (i.e., without direct information about what would contradict the rules underlying the domain).

Despite the pervasiveness of these forms of learning in real-world settings, it is worth stressing that they have been virtually ignored by experimental psychology for decades. At the beginning of the cognitive era, the study of learning was essentially devoted to classical and operant conditioning on the one hand, and to the formation of concepts or problem solving processes on the other. The above phenomena seem hardly reducible to simple conditioning effects in regards to their complexity, and research on concept learning and problem solving does not provide *a priori* a better account, primarily due to the fact that the hypothesis testing strategies essential in these research domains do not seem applicable in situations where negative evidence is lacking. This empirical and conceptual vacuum opened the door to the upsurge of the nativist perspective, which characterized the cognitive approach from its outset.

This chapter presents a stream of research that is primarily aimed at exploring the forms of learning illustrated in the examples above through laboratory situations involving arbitrary materials (for overviews, see Berry and Dienes, 1993; Berry, 1997; Cleeremans et al., 1998; French and Cleeremans, 2002; Jimenez, 2003; Perruchet and Pacton, 2006; Reber, 1993; Seger, 1994; Shanks, 2005; Stadler and Frensch, 1998). This field of research evolved essentially from the end of the 1980s, although its roots are in the pioneering studies of Arthur Reber, who coined the term 'implicit learning' (IL) about 40 years ago (Reber, 1967). The implications of the results issued from IL research for the nativist/empiricist debate will be addressed in the final discussion, after having examined what is learned in this context, how 'implicit' is implicit learning, and the relations of laboratory research with real-world situations of learning.

2.32.2 Rules, Instance-Based Processing, or Sensitivity to Statistical Regularities?

2.32.2.1 Learning Rules

A large part of the literature on IL exploits the artificial grammar learning paradigm, initially proposed by Reber (1967). Participants first study a set of letter strings generated from a finite-state grammar that defines legal letters and permissible transitions between them (**Figure 1**). Typical instructions do not mention the existence of a grammar and are framed so as to discourage participants from engaging in explicit, intentional analysis of the material. Participants are then subsequently informed about the rule-governed nature of the strings and asked to categorize new grammatical and nongrammatical letter strings. Participants are typically able to perform this task with better-than-chance accuracy, while remaining unable to articulate the rules used to generate the material. This empirical outcome has been unambiguously confirmed by a vast number of subsequent experimental studies involving many variants of the situation.

Reber's (1967) original proposal was that participants have internalized the constraints embodied by

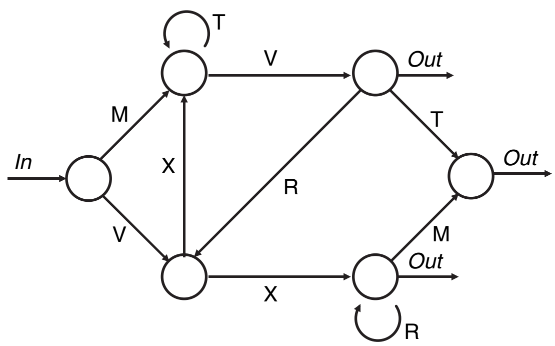

Figure 1 The artificial grammar used by Reber and Allen (1978), Dulany et al. (1984), and Perruchet and Pacteau (1990) among others. For example, MTTV and VXVRXVT are grammatical, whereas MXVT is not grammatical.

the generation of rules during training. Rule abstraction is assumed to occur during the study phase, when participants are exposed to a sample of letter strings generated from the grammar. During the test phase, participants are assumed to use the acquired knowledge, stored in an abstract format, to judge the grammaticality of new items. Other illustrations of this reasoning can be found in many subsequent studies. Let us consider those by Lewicki et al. (1988) and McGeorge and Burton (1990).

In the Lewicki et al. (1988) paradigm, participants were asked to perform a four-choice reaction-time task, with the targets appearing in one of four quadrants on a computer screen. They were simply asked to track the targets on the numeric keypad of the computer as fast as possible. The sequence looked like a long and continuous series of randomly located targets. However, this sequence was organized on the basis of subtle, nonsalient rules. Indeed, unbeknown to participants, the sequence was divided into a succession of 'logical' blocks of five trials each. In each block, the first two target locations were random, while the last three were determined by rules. The participants were unable to verbalize the nature of the manipulation and, in particular, they had no explicit knowledge of the subdivision into logical blocks of five trials, which was a precondition that had to be satisfied if they were to grasp the other rules. However, performance on the final trials of each block, the locations of which were predictable from the rules, improved at a faster rate and was better overall than performance on the first, random, trials. Lewicki et al. (1988) accounted for these results by postulating that the structuring rules were discovered by a powerful, multipurpose unconscious algorithm abstractor.

In the McGeorge and Burton (1990) study, which initiated a stream of research on the so-called 'invariant learning,' participants were asked to perform an arithmetic task on a set of four-digit numbers. Unbeknown to them, each four-digit number contained one '3' digit (the 'invariant'). In a subsequent forced-choice recognition test, participants were shown 10 pairs of four-digit numbers. They were told that one of the numbers in each pair was seen during the study phase, and that they had to find it. In fact, all the numbers were new, but half of them contained one '3' as in the study strings, while no '3' occurred in the other half. Participants choose above chance the numbers containing a '3,' although they were unable to report anything pertinent to the invariant digit. The authors inferred that participants had learned the critical rule unconsciously.

These results, and most of the others in the early IL literature, have been shown to be empirically robust in subsequent studies. However, two other interpretations have been proposed. Their common intuition is that people do not abstract the rules of the domain, but instead learn about the product of the rules.

2.32.2.2 The Instance-Based or 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 specific 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.

Vokey and Brooks (1992) made independent the usually confounded factors of specific similarity and grammaticality, in order to assess the size of the effect of each factor on grammaticality judgments. Test items were classified as similar when they differed by only one letter from one study item, and dissimilar when they differed by two or more letters from any study items. The authors obtained a reliable effect of specific similarity on grammaticality judgments (see also McAndrews and Moscovitch, 1985). As expected, similar items were more often classified as grammatical than dissimilar items when their grammatical status (i.e., their consistency with the grammar) was kept constant. However, the grammaticality factor also had a significant, and usually additive, effect. Similar evidence was collected by Cock et al. (1994) in invariant learning. The authors demonstrated that similarity to instances in the study phase was even a more important factor than apparent knowledge of the invariant feature in the McGeorge and Burton (1990) paradigm.

To account for the fact that the similarity to a specific training item fails to account for all the variance in performance, Vokey and Brooks (1992; Brooks and Vokey, 1991) suggested that the similarity may also be computed with the whole set of study items instead of a single one (see also Pothos and Bailey, 2000, for another measure of similarity). The currently prevalent interpretation keeps the idea of some kind of pooling or summation over multiple episodes, but privileges a formulation in terms of statistical regularities.

2.32.2.3 The Sensitivity to Statistical Regularities

While the instance-based model considers whole episodes (e.g., VXVRXVT in artificial grammar learning), this alternative account considers elementary components (e.g., the individual letters). The consequences are considerable. Indeed, large episodes are idiosyncratic, and hence they generate distinctive and independent memory traces. By contrast, elementary components occur iteratively under the same or other combinations, and hence it makes sense to describe the to-be-learned stimuli using statistical concepts, such as frequency, probability, or contingency. For this reason, this approach is designated here as the statistical account, even though it is most commonly referred to as the 'fragmentary' approach in the conventional IL literature. It is worth noting that the term 'statistics' does not necessary entail that learners perform statistical computations, an issue that will be addressed later (see the section titled 'Statistical computations and chunk formation').

The general principle of this account is straightforward: Organizing rules generates statistical regularities in the world, and people adjust their behavior to those regularities. Understanding how this account works in artificial grammar learning is easy. Looking at the grammar shown in **Figure 1**, it appears that that some associations between letters are possible (e.g., MV), and other impossible (e.g., MX), and that among the legal associations, some are more frequent than others (e.g., RX presumably occurs more often than RM). If participants learn something about the frequency distribution of the pairs of letters (bigrams) that compose the study strings, they should perform subsequent grammaticality judgments better than chance. Perruchet and Pacteau (1990) tested this hypothesis. They reasoned that, if subjects learn only bigram information when faced with the whole strings, the direct presentation of the bigrams, which precludes the use of any high-level rules, should not change the final performance. The prediction was confirmed; the performance of participants who had learned using the complete grammatical strings (as usual) and those who were trained using the bigrams from which these strings were composed were statistically indistinguishable. Other experiments from the same study and other studies (Dienes et al., 1991; Gomez and Schvaneveldt, 1994) confirmed the importance of bigrams knowledge, although they showed that participants also learn other piecemeal information, such as the location of permissible bigrams and the first and last letters of the strings.

The question is now: Does this interpretation work in general? It could be argued indeed that artificial grammar learning is especially well-fitted to a statistical interpretation, because the rules can be easily translated in terms of statistical regularities. To address this question, let us consider how the Lewicki et al. (1988) study presented above can be reinterpreted. Recall that a precondition to grasp the complex second-order dependency rules structuring the sequence was a parsing of the whole sequence into logical blocks of five trials, and that participants were fully unconscious of doing so. However, Perruchet et al. (1990) demonstrated that participants could learn the task without ever performing the segmentation of the sequence into logical blocks. Instead, they could become sensitive to the relative frequency of small units, comprising 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 prohibited certain transitions from occurring. In particular, an examination of the rules shows that they never generated back-and-forth movements. As a consequence, the back-and-forth transitions were less frequent *on the whole sequence* than the other possible movements. The crucial point is that these less frequent events, which presumably elicit longer reaction times, were exclusively located on the random trials. This stems not from an unfortunate bias in randomization, but from a logical principle: The rules determined both the relative frequency of certain events within the entire sequence and the selective occurrence of these events in specific trials (for an alternative interpretation based on connectionist modeling, see Cleeremans and Jimenez, 1998).

A similar reanalysis was performed by Wright and Burton (1995) on the McGeorge and Burton (1990) invariant task. Wright and Burton observed that a by-product of the invariant rule was to modify the probability of occurrence of observing a digit repetition in the strings. More precisely, the strings that contain one '3' include, on the mean, a smaller proportion of repeated digits than the strings in which no '3' occurs, all simply because the chances of generating repeated digits are lesser over three than over four successive drawings. The authors showed that at least a part of the participants' above-chance performance during the test was due to the fact that they tended to reject the items containing repetitions, rather than the items violating the invariant rule.

What is new in these examples (for a similar illustration, see the reinterpretation of Kushner et al. (1991) by Perruchet (1994b)) with regard to the artificial grammar-learning situation is the fact that the link between the generating rules and the distributional statistics of simple and salient events is far from obvious. The fact that rules may have remote consequences, the learning of which having effects similar to the learning of the rules themselves, may obviously be thought of as a drawback in the experimental designs, without any implication out of the laboratory studies. However, it may also be thought of as a quite fundamental outcome, essential to understand the power of the statistical approach in the natural situation of learning, as the section titled 'Discussion: about nativism and empiricism' will emphasize.

2.32.2.4 Rules versus Statistics: A Crucial Test

How can rule-based and statistical interpretations be discriminated? When the rules of a domain generate a set of events so restricted that all the possible events can be exhaustively experienced by a subject, it may be impossible to discriminate between the two types of interpretations. However, this case is largely deprived of interest. Indeed, the power of the rules is that they make people able to adapt to new situations from previous exposure to a subset of the events the rules can generate. Here is the hint for a crucial test.

For a first example, let us consider an argument for rules put forth by Reber and Lewis (1977) in the context of artificial grammar learning. In a given experiment, participants are exposed to a subset of the virtual full set of strings generated by the grammar, and this subset cannot be perfectly representative of the full set for all aspects. For instance, the frequency distribution of the observed bigrams has a high probability of departing to some extent from the frequency of the bigrams composing the full set of strings. Reber and Lewis argued that if participants abstract the rules of the grammar, they should be sensitive to the bigram frequency of the virtual full set of strings, and not the frequency of occurrence of the bigrams composing the strings actually displayed in the study phase. They provided empirical data supporting this hypothesis, and Reber (1989) construed these data as one of the main supports for his contention that studying grammatical letter strings gives access to the abstract structure of the grammar. The logic of the argument is indeed sound, but unfortunately, the supporting data turned out to be due to various methodological drawbacks inherent to the Reber and Lewis procedure (Perruchet et al., 1992). In fact, participants are sensitive to the frequency distribution of the bigrams they actually perceived.

In the preceding example, the possibility of discriminating interpretations based on rules and statistics stems from the fact that the subset of items to which participants are exposed are not representative of the whole set of items due to sampling biases. The same logic may be implemented in a more systematic way, by training participants with a given material and testing them with different material. The following section examines the findings obtained in these so-called 'transfer' situations, which have been heavily used in IL research.

2.32.2.5 The Phenomenon of Transfer: The Data

In the standard paradigm of transfer in artificial grammar learning, the letters forming the study items are changed in a consistent way for the test of grammaticality (e.g., M is always replaced by C, X by P, etc.). Reber (1969) and several subsequent studies (e.g., Mathews et al., 1989; Dienes and Altmann, 1997; Manza and Reber, 1997; Shanks et al., 1997; Whittlesea and Wright, 1997) have shown that participants still outperform chance level under these conditions. The principle underlying the 'changed letter procedure' has been extended to other surface changes. For instance, the training items and the test items may be, respectively, auditory items and visual items (Manza and Reber, 1997), color and color names, sounds and letters (Dienes and Altmann, 1997), or vice versa. Successful transfer was observed in each case.

The phenomenon of transfer has also been observed in invariant learning. McGeorge and Burton (1990) found that the selection of number strings containing the invariant digit persisted when study strings were presented as digits (e.g., 1234) and test strings as their word equivalents (e.g., one two three four; see Bright and Burton (1998) for similar examples of transfer in another invariant learning task).

Transfer has even been observed in infants. In Marcus et al. (1999), 7-month-old infants were exposed to a simplified, artificial language during a training phase. Then they were presented with a few test items, which were all composed of new syllables. For instance, in one experiment, infants heard 16 three-word sentences such as *gatiti, linana,* or *tanana,* during the study phase. All of these sentences were constructed on the basis of an ABB grammar. The infants were then presented with 12 other three-word sentences, such as *wofefe* and *wofewo.* The crucial point is that, although all of the test items were composed of new syllables, only half of the items were constructed from the grammar with which the infants had been familiarized. In the selected example, the grammatical item was *wofefe. Wofewo* introduces a structural novelty in that it is generated from a concurrent ABA grammar. The infants tended to listen more to the sentences generated by the ABA grammar, thus indicating their sensitivity to the structural novelty. In another experiment, infants were shown to be able to discriminate sentences generated by an AAB grammar. Similar studies using more complex material have been performed with 11-month-old infants (Gomez and Gerken, 1999).

2.32.2.6 The Phenomenon of Transfer: The Interpretations

2.32.2.6.1 Rules?

Marcus et al. concluded that infants have the capacity to represent algebraic-like rules and, in addition, "have the ability to extract those rules rapidly from small amounts of input and to generalize those rules to novel instances" (Marcus et al., 1999, p. 79). Demonstrations of transfer in more complex situations in adults have elicited similar comments. For instance, Reber, talking about performance in the changed letter procedure in artificial grammar learning studies, claimed that

> . . . the abstractive perspective is the only model of mental representation that can deal with the existence of transfer of knowledge across stimulus domains. (Reber, 1993: 121)

A rule-based interpretation may have difficulty accounting for the entire pattern of data, however. First, the traditional emphasis on positive results must not overshadow the fact that transfer failure has frequently been reported in the literature on IL. In the conclusion of their review on transfer in the most current IL paradigms, Berry and Dienes pointed out that

> . . . the knowledge underlying performance on numerous tasks . . . often fails to transfer to different tasks involving conceptually irrelevant perceptual changes. (Berry and Dienes, 1993: 180)

This empirical finding leads the authors to propose that limited transfer to related tasks is one of the important key features of performance in IL tasks. Moreover, in experiments where positive evidence of transfer is reported, performance levels on transfer situations are generally lower than performance levels on the original training situation. This so-called transfer decrement phenomenon raises a problem for a rule-based standpoint. In an authoritative discussion on the use of abstract rules, Smith et al. (1992) posit as the first of their eight criteria for rule use that "Performance on rule-governed items is as accurate with unfamiliar as with familiar material" (Smith et al., 1992, p. 7; see also Anderson, 1994, p. 35; Shanks, 1995, Ch. 5; Whittlesea and Dorken, 1997, p.66). Manza and Reber (1997) acknowledge this implication of their own abstractionist view. Clearly, this prediction of rule-based accounts has scarce experimental support at best.

However, observing that rule-based interpretation of transfer is, after all, not so well-fitted as might expected has limited interest until better interpretations are put forward. Are there alternatives?

2.32.2.6.2 Explicit inferences during the test?

In the standard situations of artificial grammar learning, most people are able to learn the abstract rules of the grammar when they are instructed to search for rules (Turner and Fischler, 1993) or when they are given incidental instructions which guide them toward the deep structure of the material (Wittlesea and Dorken, 1993). A first alternative possibility to account for transfer in IL studies is that transfer is due to the involvement of explicit reasoning, despite the instructions given to participants.

This account finds support in the examination of the tasks in which transfer routinely succeeds and tasks in which transfer fails. As noted by Newell and Bright (2002), the tasks that trigger transfer are those in which participants are instructed to use knowledge that they have acquired during training, such as artificial grammar learning tasks and invariant learning tasks. These instructions inevitably shift subjects to a rule-discovery mental set. The tasks in which subjects are not explicitly engaged to rely on what they saw in study phase, such as serial reaction-time (SRT) tasks and control interactive tasks, are far less prone to transfer. In SRT tasks, for instance, a target stimulus appears on successive trials at one of a few possible positions on the computer screen. Participants are simply asked to react to the appearance of the target by pressing a key that spatially matches the location of the target on a keyboard. Typically, the same sequence of trials is repeated throughout the session. Participants exhibit a decrease in reaction times with regard to a control condition, without ever being informed about the presence of a repeated structure. In this case, transfer to dissimilar surface feature typically fails (Stadler, 1989; Willingham et al., 1989).

The role of explicit reasoning in changed-letter transfer in artificial grammar learning is further suggested by the fact that transfer is performed better when the training session involves 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 (1997) showed that changed-letter transfer occurred only in subjects who had sufficient explicit knowledge of the rules.

Although these studies suggest that transfer performance partly depends on the involvement of conscious and deliberate processes, it is difficult to account for all the positive results in those terms. To evoke only one counterargument, the observation of transfer in infants (Gomez and Gerken, 1999; Markus et al., 1999) can hardly be explained by the recourse to intentional rule-breaking strategies. Is it possible to account for transfer in IL without any recourse to rules?

2.32.2.6.3 Disentangling rules and abstraction

There is no doubt that the evidence of transfer is indicative of abstraction. However, the view that abstraction is indicative of rule formation and rule use has been heavily challenged. As cogently argued by Redington and Chater (2002), "surface-independence and rule-based knowledge are orthogonal concepts."

To begin with a simple case, let us consider Manza and Reber's (1997) results, showing a transfer between auditory and visual modalities in the artificial grammar learning area. These authors interpret their findings as providing support for their abstractionist, rule-based view. However, the phenomenon can be easily explained otherwise. Any sequence – such as VXMTR – presented orally will be immediately recognized when displayed visually, irrespective of whether this sequence is generated by a grammar or not. This is because the perceptual primitives, namely the letters V, X, and so on, are processed to an abstract level that makes them partially independent of their sensory format. The differences between the two explanations is worth stressing. In the former case, a rule-governed pattern is assumed to be extracted from the auditory stimuli before being applied to the visual stimuli. In the latter case, matching is directly performed at the levels of the perceptual primitives. The same comment can be applied to many other studies. For example, the transfer between colors and the name of colors (Dienes and Altmann, 1997) and the transfer between digits and their word equivalents (McGeorge and Burton, 1990) can also be accounted for by the natural mapping between the primitives involved in the experiment.

At first glance, the above explanation does not apply to all transfer results. As a case in point, it does not seem to work for the studies by Marcus et al. (1999) in which transfer is observed between, say, *gatiti* and *wofefe*, because there is no natural mapping between *ga* and *wo*, or *ti* and *fe*. Reinterpretation of the Marcus et al. data is possible along the same line, however, if one assumes that the perceptual primitives can be relational in nature. The relation that needs to be coded is the relation 'same-different,' or, in other words, the only ability that infants need to possess is that of coding the repetition of an event. Indeed, as pointed out by McClelland and Plaut (1999), *gatiti, wofefe*, and more generally all the ABB items, can be coded as different-same, whereas none of the other items can be coded using the same schema.

As surprising as this conclusion may be, the demonstrations of transfer stemming from the more complex situations of artificial grammar learning in adults imply the coding of no more complex relations than event repetitions (e.g., Tunney and Altmann, 1999; Gomez et al., 2000; for an overview, see Perruchet and Vinter, 2002, Section 6). Lotz and Kinder (2006) confirmed and extended this conclusion. They showed that the sources of information used in transfer tasks in artificial grammar learning studies concern the local repetition between adjacent elements, as well as the repetition of nonadjacent elements in the whole items.

Overall, this analysis demonstrates that transfer, such as observed in IL settings, is in no way indicative of rule knowledge. It is fairly compatible with a statistical approach, provided one acknowledges the possibility that statistical processes operate not only on surface features (such as forms or colors), but also on more abstract properties and on simple relational features, such as the repetition of events. The idea that transfer does not imply rule abstraction has also gained support from the possibility of accounting for transfer performance within a connectionist framework (Altman and Dienes, 1999; McClelland and Plaut, 1999; Seidenberg and Elman, 1999; Christiansen and Curtin, 1999). Note also that transfer has been claimed to be compatible with an interpretation focusing on instance-based processing, thanks to the notion of abstract analogy (Brooks and Vokey, 1991), although Lotz and Kinder (2006) failed to find an empirical support for this account.

2.32.2.7 A Provisional Conclusion

There is evidence that the analogy with specific items may account for a specific part of variance in performance. The Vokey and Brooks (1992) demonstration, presented in the section titled 'The instance-based or episodic account,' has been challenged (Knowlton and Squire, 1994; Perruchet, 1994a), but additional evidence has been provided since then (Higham, 1997). The interest of the instance-based model resides in its highlighting the fact that behavior may be implicitly affected by individual episodes rather than simply by large amounts of training. However, there is a consensus on the idea that this account cannot be thought of as exclusive. It seems inevitable to jointly consider the pooled influence of a series of events to account for the whole pattern of data.

The two main views accounting for the influence of multiple past events are based on rules and statistics, respectively, but there is no symmetry between the two accounts. Indeed, no one disputes the existence of statistical learning. This consensus comes from the human ability to learn in the countless situations in which regularities cannot be described by a set of rules, as the concept has been defined above. As a consequence, the only possible question is: Do we need rules, in addition to statistical learning, to account for implicit learning in rule-governed situations?

Here is the end of the consensus. On the one hand, many authors respond "no." Their position is based on the fact that the sensitivity to statistical regularities is able to account for performance in most of the experimental situations that were initially devised to provide an existence proof for rule learning, including transfer settings. In addition, when a direct test has been performed to contrast the predictions of the two models, predictions of the statistical account have been unambiguously confirmed (see also Perruchet (1994b) on the situation devised by Kushner et al. (1991) and Channon et al. (2002) on the biconditional grammar). On the other hand, other authors (e.g., Knowlton and Squire, 1996) argue that empirical evidence requires a dual process account, mixing statistical learning and rule knowledge. Their position stems from experimental studies in which learning persists in test conditions where the simplest regularities – those that are presumably captured by statistical learning – have been made uninformative (e.g., Knowlton and Squire, 1996; Meulemans and Van Der Linden, 1997; but see Kinder and Assmann, 2000). This kind of evidence is not fully compelling, however, because it is not possible to ascertain that all the possible sources of statistical knowledge have been taken into account (see for instance the reanalysis of Meulemans and Van der Linden (1997) by Johnstone and Shanks (1999)). A more principled demonstration, in which some specified content of knowledge would fail to be approximated by statistical learning, would provide a far stronger argument.

The remaining of this chapter focuses on statistical learning. This does not mean that the possibility of implicit rule learning can be considered as definitely ruled out. This presumably will never be the case, because proof of nonexistence is beyond the scope of any empirical investigation. Needless to say, this approach does not mean either that rule learning does not exit at all; there is clear evidence that humans are able to infer and use abstract rules when conscious thought is involved. The very existence of science should provide a sufficient proof for the skeptic.

The implications of focusing on statistical learning are that the questions and their experimental approach will be considered irrespectively of whether the to-be-learned materials can be described in terms of rules or not. Note that this focus is in keeping with the recent literature on IL, which typically includes a number of situations that are not governed by rules such as SRT tasks with repeated sequences and word segmentation (e.g., Saffran et al., 1997), as well as other situations, such as contextual cuing (Chun and Jiang, 2003), that are not described in this chapter due to space limitations.

2.32.3 Learning about Statistical Regularities

Before examining the question of what processes underlie behavioral tuning to statistical regularities, one needs to identify the kinds of regularities to which humans are sensitive.

2.32.3.1 What Is Learnable?

2.32.3.1.1 Frequency, transitional probability, contingency

For many people, claiming that behavior is sensitive to statistical regularities amounts to saying that behavior is sensitive to the absolute or the relative frequency of events. For instance, participants in an artificial grammar experiment may have learned that MT occurred n times, or that a proportion p of the displayed bigrams were MT. Considering only frequency provides limited information, however. It may be interesting to know the probability for 'M' to be followed by 'T,' a measure called conditional or transitional probability. To assess whether 'M' is actually predictive of 'T,' this probability must be compared to the probability that another letter precedes 'T'. The difference between these two conditional probabilities is called DeltaP (Shanks, 1995). In addition, it may be worth considering the reverse relations, namely the probability for 'T' to be preceded by 'M.' This 'backward' transitional probability may be quite different from the standard, forward transitional probability. The normative definition of contingency in statistics (such as measured, for instance, by a χ^2 or a Pearson correlation) requires a consideration of the bidirectional relations. When data are dichotomized, for instance, Pearson correlation is the geometrical mean of the forward and backward DeltaP (for a more detailed presentation, see Perruchet and Peereman, 2004).

The focus on frequency in early studies on IL does not mean that human behavior is only sensitive to this variable. Indeed, all the measures of association are generally correlated, so evidence collected to support one specific measure is equivocal if no special care is taken for controlling the other measures. Aslin et al. (1998) demonstrated that participants were sensitive to the transitional probability in word-segmentation studies. These results have been replicated in visual tasks (e.g., Fiser and Aslin, 2001), so that most recent studies on statistical learning take for granted that the statistics to which people are sensitive are transitional probabilities. This conclusion could be premature, given the correlations between the different measures, and the paucity of studies including different measures. In fact, the literature on conditioning has long suggested that even animals such as rats or pigeons are sensitive to DeltaP (Rescorla, 1967). Perruchet and Peereman (2004) compared several measures of associations, and they found that participants were more sensitive to the bidirectional contingency than to simpler measures of associations (although in a specific context). A conservative conclusion could be that people are sensitive to more sophisticated measures of associations than co-occurrence frequency, and further study is needed for assessing more precisely which statistic is the more relevant in each context.

2.32.3.1.2 Adjacent and nonadjacent dependencies

A dimension orthogonal to the previous one concerns the distance between the to-be-related events. The early studies endorsing a statistical approach in the IL domain focused on adjacent elements (typically the bigrams of letters). The importance of adjacent relations, however, does not mean that it is impossible to learn more complex information. A number of studies in SRT tasks have investigated how reactions times to the event n improved due to the information brought out by the events n-1, n-2, n-3 (known as first-order, second-order, and third-order dependency rules, respectively), and so on. Second-order dependencies can be learned quite easily and are now used as a default in most SRT studies. Third-order dependency rules can also be learned, although less clearly (Remillard and Clark, 2001). However, higher-order dependency rules are seemingly much harder, or impossible to learn. For instance, even after 60 000 practice trials, Cleeremans and McClelland (1991) obtained no evidence for an effect of the event four steps away from the current trial.

In the situations discussed so far, the relations between distant events are not considered independently from the intervening events. By contrast, in the AXC structures investigated in several recent studies, a relation exists between A and C irrespective of the intervening event X, which is statistically independent from both A and C. Examining whether learning those nonadjacent relationships is possible was prompted by the fact that these relations are frequent in high-level domains such as language and music. The studies investigating the possibility of learning nonadjacent dependencies between syllables or words (Gomez, 2002; Newport and Aslin, 2004; Perruchet et al., 2004; Onnis et al., 2005), musical sounds (Creel et al., 2004; Kuhn and Dienes, 2005), digits (Pacton and Perruchet, in press), and visual shapes (Turk-Browne et al., 2005) report positive results. However, most of them conclude that learning nonadjacent dependencies presupposes more restrictive conditions than those required for learning the relations between contiguous events. Gomez (2002) showed that the degree to which the A_C relationships were learned depended on the variability of the middle element (X). For Newport, Aslin, and collaborators (e.g. Newport and Aslin, 2004), the crucial factor is the similarity between A and C. Learning seems also much easier in a situation where the successive AXC units are perceptually distinct (e.g., Gomez, 2002), than in situations where they are embedded in a continuous sequence (e.g., Perruchet et al., 2004). By contrast, Pacton and Perruchet (in press) provided support for a view in which nonadjacent dependencies can be learned as well as adjacent dependencies insofar as the relevant events are actively processed by participants to meet the task demands.

2.32.3.1.3 Processing multiple cues concurrently

Up to now, we have examined how the learner exploits one source of information, for instance, event repetition. However, taken in isolation, a source of information often has a limited value in real-world settings. The system efficiency would be considerably extended if various sources could be exploited in parallel. The number of studies exploring this issue is still tiny, but they provide converging evidence for a positive assessment, as well in artificial grammar learning (e.g., Kinder and Assmann, 2000; Conway and Christiansen, 2006) as in SRT tasks (e.g., Hunt and Aslin, 2001). Studies on word segmentation have also demonstrated the possibility of combining statistical and prosodic cues (e.g., Thiessen and Saffran, 2003). The concurrent exploitation of various information sources can be simulated by connectionist networks (e.g., Christiansen et al., 1998), a feature that strengthens the plausibility of such a possibility in humans.

2.32.3.1.4 Does learning depend on materials?

An impressive amount of data suggests that statistical learning mechanisms are domain general. For instance, although most studies in artificial grammar learning involve consonant letters, a large variety of other stimuli have been used occasionally, such as geometric forms, colors, and sounds differing by their timbre or their pitch, without noticeable difference. Conway and Christiansen (2005) directly compared touch, vision, and audition, and found many commonalities, although sequential learning in the auditory modality seemed easier than with the other two senses. Likewise, data on word segmentation have been successfully replicated with tones instead of syllables (e.g., Saffran et al., 1999; Saffran, 2002).

However, the fact that IL processes have a high level of generality across and within sensory modalities does not mean that they apply equally well to any stimuli, as if statistical learning mechanisms were blind to the nature of processed material. The well-known difficulty of publishing null results certainly accounts for a part of the apparent universality of IL mechanisms. A closer look shows, for instance, that learning may depend on aspects of the material that could seem *a priori* irrelevant. For instance, there is overwhelming evidence of rapid learning in standard SRT tasks, in which a target appears in successive trials at one of a few discrete locations. Chambaron et al. (2006) explored a similar situation, except that participants had to track a target that moved along a continuous dimension. They fail to obtain evidence of learning in several experiments, hence showing that, in spite of a close parallel between continuous tracking tasks and SRT tasks, taking benefit from the repetition of a segment in continuous tracking task appears to be considerably more difficult than taking benefit from the repetition of a sequence in SRT tasks. Moreover, recent research on statistical learning has shown that learning was highly dependent on low-level perceptual constraints. For instance, for a given statistical structure, the acoustic properties of the artificial speechflow have been shown to be determinant for learning to segment the speechflow into words (e.g., Onnis et al., 2005). Shukla et al. (2007) provides evidence that possible world-like sequences, namely chunks of three syllables with high transition probabilities, are not recognized as words if they straddle two prosodic constituents.

The efficiency of learning may also depend on high-level expectancies about the structure of the material. For instance, Pothos (2005) used an artificial grammar learning task in which the consonant letters were replaced by the name of cities. The sequences of cities were presented as the routes a salesman has to travel. In one group of participants, the training sequences matched the intuitive expectation that the salesman follows routes that link nearby cities, while in a second group, they conflicted with this intuition. Learning, as assessed by the comparison with an adequate control group, occurred only in the first group, as if a conflict between the knowledge acquired by processing the statistical structure of the stimuli and the intuitive expectations about stimulus structure prevented learning from occurring in the second group.

2.32.3.1.5 About the learners

Any discussion about learnability is meaningless without considering the characteristics of the learners. Most of the studies reported above have been performed on healthy adult participants.

Regarding first the effect of age, recent research on statistical learning has shown the surprising learning abilities of infants. These abilities have been initially revealed with auditory artificial languages (e.g., Saffran et al., 1996), but they are in no way limited to language-like stimuli. They concern sounds or combinations of visual features as well (e.g., Fiser and Aslin, 2001, 2002). Several studies have also investigated IL in children. They suggest that there is no noticeable evolution from 4- or 5-year-old children to young adults (e.g., Meulemans et al., 1998; Vinter and Detable, 2003; Karatekin et al., 2006). Other studies suggest that IL does not decline with healthy aging (e.g., Cherry and Stadler, 1995; Negash et al., 2003). Furthermore, a number of studies have reported impressive IL abilities in children (e.g., Detable and Vinter, 2004) and young adults (Atwell et al., 2003) with mental retardation, and in patients with psychiatric (e.g., Schwartz et al., 2003) and neurological disorders, including amnesia (e.g., Meulemans and Van der Linden, 2003; Shanks et al., 2006), Alzheimer's disease (Eldridge et al., 2002), Parkinson's disease (e.g., Smith and McDowall, 2006), closed-head injury (e.g., Vakil et al., 2002), and Williams Syndrome (Don et al., 2003). Statistical learning abilities have been shown in animals such as nonhuman primates (e.g., Hauser et al., 2001) or even rats (Toro and Trobalon, 2005).

These data have crucial implications for a number of fundamental and applied issues. They do not mean, however, that everyone shares equivalent abilities whenever statistical learning is concerned. In fact, comparative studies often select situations the difficulty of which is *a priori* well-suited for the full span of the investigated population. When the level of difficulty is increased, a difference often emerges. For instance, a deficit of performance has been observed in complex IL tasks in elderly people (e.g., Howard et al., 2004) and in amnesic patients (Curran, 1997; Channon et al., 2002). This dependency of IL with regard to learner's general competencies whenever the learning settings become complex enough is confirmed by studies on adult healthy people. For instance, Dienes and Longuet-Higgins (2004) observed that only participants experienced with atonal music were able to learn artificial regularities following the structures of serialist music.

2.32.3.2 Statistical Computations and Chunk Formation

2.32.3.2.1 Computing statistics?

Observing that performances in IL tasks conforms to statistical regularities may lead us to infer that learners compute statistics. Certainly, the idea that learners unconsciously compute statistics using the same algorithms as a statistician would use is somewhat implausible. However, the possibility of approximating the outcome of analytical computations through connectionist networks (e.g., Redington and Chater, 1998) offers a much more appealing alternative. Learning is performed by the progressive tuning of the connection weights between units within multilayer networks. Although different types of networks have been used (see Dienes, 1992), the Simple Recurrent Networks (SRN), initially proposed by Elman (1990), have been the most widely applied to IL. SRNs are typically trained to predict the next element of sequences presented one element at a time. Cleeremans and McClelland (1991) have shown that an SRN was able to simulate the performance of human learners in an SRT task in which the successive locations of the target were generated by a finite-state grammar, and the ability of SRNs to successfully account for performance in various IL paradigms has been confirmed since then in a number of studies (e.g., Kinder and Assmann, 2000). Certainly, due to the impressive ability of connectionist networks to simulate learner's performance, the idea that learners actually perform statistical computations is taken for granted by a number of authors. This idea is not compelling however. Inferring statistical computation from statistical sensitivity may amount to repeating the same error as the early researchers in IL, who inferred rule abstraction from the behavioral sensitivity to rules. An alternative interpretation emerges from the observation that IL generally leads to the formation of chunks.

2.32.3.2.2 The formation of chunks

The fact that learning leads to the formation of chunks is largely consensual. This is obviously the case in recent studies on word segmentation and object formation, in which performance is directly assessed through chunk formation. But this is also true in most of the other situations of IL, in which chunks are not the explicit end-product of learning. A number of studies have shown that participants learn small chunks of two or three elements in artificial grammar learning settings. Chunking the material, far from being a degraded procedure, is a highly efficient mode of coding. Indeed, dealing with small units facilitates transfer and generalization. This happens because, given the structure of finite-state grammars, new items are formed by recombining old components. However, this is true only if the chunks respect the statistical structure of the material. To put the matter simply, assuming five events (A, B, C, D, and E), forming the chunks 'AB' and 'CDE' is beneficial only if A and B on the one hand, and C, D, and E on the other hand, form cohesive structures. If 'AB' is frequently followed by C, and 'DE' frequently occurs in other contexts, then this mode of chunking would be ill suited (Perruchet et al., 2002).

The formation of chunks forming cohesive structures can be easily accounted for by the idea that learners compute statistics. For instance, for Saffran and Wilson (2003), verbal chunks are inferred from statistical computations and then serve as the stuff for further statistical computations. Fiser and Aslin (2005) also consider that the visual input is chunked into components according to the statistical coherence of their components. To use the five-event example, AB and CDE would emerge as from some kind of cluster or factorial analyses once the correlational structure of the events has been computed.

2.32.3.2.3 Are statistical computations a necessary prerequisite?

Chunks consistent with the statistical structure of the material can also emerge without prior statistical computations. Simple memory mechanisms could be sufficient. To begin with, let us consider the ubiquitous phenomenon of forgetting. Because frequently repeated events tends to be forgotten to a lesser extent than less frequent events, forgetting leads us to be sensitive to event frequency, without any statistical 'computation.' Several models of IL (Servan-Schreiber and Anderson, 1990; Knowlton and Squire, 1996; Perruchet and Vinter, 1998a) rest on this intuition. Again, in the example, the chunks 'AB' and 'CDE' would emerge from the fact that A and B on the one hand, and C, D, and E on the other hand, occur more often together than any other combinations of events. In Perruchet and Vinter's Parser model, those chunks emerge simply because other associations of events (such as ABC), if they occur, are forgotten due to their relative rarity. The difference with the statistical account is that, instead of being inferred from the results of statistical computation, chunk formation is the primary mechanism, and the cohesive chunks are those that are selected among a number of other ones due to well-known laws of associative memory, primarily forgetting.

Chunking is often thought of as exclusively sensitive to the raw frequency. This would indeed be the case if the strength of memory traces only depended on the repetitions of events. However, it has long been known that forgetting is due in large part to the interference generated by the prior or subsequent events that are related in some way to the target event. Now, and this is the crucial point, taking into account the effect of interference in chunk formation amounts to considering other measures of association than the raw frequency of co-occurrences. For instance, implementing forward interference is sufficient to make chunk strength sensitive to transitional probabilities (Perruchet and Pacton, 2006, Box 3). Moreover, Perruchet and Peereman (2004) have shown that the Perruchet and Vinter's (1998a) Parser model, thanks to the role ascribed to interference in chunk formation, was even sensitive to contingency, that is, to a measure of association more comprehensive than conditional probabilities.

To recap, the current debate is between those who argue that statistical computations are performed first, with the chunks inferred on the basis of their results, and those who argue that the chunks are formed from the outset, with the sensitivity to statistical regularities being a by-product of the selection of those chunks as a consequence of ubiquitous memory laws. Note that the two interpretations are equally consistent with associative learning principles. One advantage of the second option is its parsimony. Indeed, no additional computational devices have to be imagined to extract chunks from distributional information. In addition, the chunk model can be easily unified with the instance-based model, because both are grounded on standard memory laws. A unified view could find an integrative framework in the so-called processing account of IL. This account borrows, from research in memory, the idea that memory traces are 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 (e.g., Roediger et al., 1989). Support for this view in IL studies 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. Although the processing account is historically associated with the instance-based models of IL (e.g., Neal and Hesketh, 1997), its grounding principles could be applied to chunk-based models as well.

These considerations, however, cannot be considered as compelling. At this time, the available experimental studies intended to tease apart the predictions from statistical and chunk-based models have produced equivocal results (e.g., Boucher and Dienes, 2003). Clearly, the outcome of the debate is pending further empirical investigations.

2.32.4 How Implicit Is 'Implicit Learning'?

What defines implicitness in IL is far from being agreed upon. A distinction is made, in the following sections, between what occurs during the training phase and the test phase of an IL session. The study-test distinction has limited interest, insofar as in most real-world situations, and in several laboratory situations (such as SRT tasks) as well, any event both influences subsequent events and is itself influenced by the prior ones, hence serving the two functions simultaneously. However, this distinction provides a convenient means to tease apart different issues.

2.32.4.1 Implicitness during the Training Phase

2.32.4.1.1 Incidental and intentional learning

A feature which is a part of virtually all definitions of IL is the incidental nature of the acquisition process. IL proceeds without people's intention to learn. This characteristic is sometimes the only one to be retained, thus conflating the notion of IL and incidental learning (e.g., Stadler and French, 1994). The SRT tasks are often considered as those that offer the best guarantee of the incidental nature of learning, because this task is endowed with its own internal purpose, and it leaves quite limited time for thinking about the task structure. In most of the other IL tasks, instructions distract participants from thinking about the overall material structure, by focusing participants' attention on individual items. For instance, in artificial grammar learning, participants are generally asked for the rote learning of individual letter strings. In invariant learning, participants are asked to perform some arithmetic computation on each digit string. In other tasks, such as the word-segmentation task, participants are simply asked to listen to the artificial language, without specific demands.

2.32.4.1.2 Is attention necessary?

A question of major interest is whether performance improvement depends on the amount of attention paid to the study material during the familiarization phase. The main strategy consists in adding a concurrent secondary task during the training session, then observing whether performance improvement is equivalent to that observed in a standard procedure.

A few early studies claimed that the addition of a secondary task had no effect, or even could *facilitate* learning in very complex experimental settings. This leads to contrast the concepts of 'selective learning' and 'unselective learning' (e.g., Berry and Broadbent, 1988), with the latter being assumed to occur when the situation was too complex to be solved by attention-based mechanisms. The original results were not replicated, however (e.g., Green and Shanks, 1993), and to the best of our knowledge, the notion of unselective learning, as initially discussed in the studies conducted by Broadbent and colleagues, is no longer advocated.

The idea of two forms of learning, differing according to whether attention is required or not, has also been proposed in another context, but with the opposite stance. The hypothesis was that attention is required for learning complex sequences in SRT tasks, while nonattentional learning is efficient for the simplest forms of sequential dependencies (Cohen et al., 1990). However, observing learning under dual-tasks conditions does not imply the existence of a nonattentional form of learning, because the secondary task might not deplete the attentional resources completely (Stadler, 1995). Closing their survey on the role of attention in implicit sequence learning, Hsiao and Reber concluded:

> We view sequence learning as occurring in background of the residual attention after the cost of the tone-counting task [commonly used as a secondary task in this context] and the key-pressing task. If there is still sufficient attention available to the encoding of the sequence, learning will be successful; otherwise, failure will result. (Hsiao and Reber, 1998: 487)

Regarding artificial grammars, Reber (e.g., 1993) has also acknowledged that attention to the study material is necessary for learning to occur. In support of this claim, Dienes et al. (1991) have shown that the accuracy of grammaticality judgments was lowered when subjects had to perform a concurrent random number generation task during the familiarization phase.

Note also that other studies have shown that, without at least minimal attentional involvement, even simple covariations or regularities turn out to be impossible to learn (Jimenez and Mendez, 1999; Hoffmann and Sebald, 2005; Pacton and Perruchet, in press; Rowland and Shanks, 2006b). The conclusion according to which improved performance in IL situation requires attention has been recently supported by studies on statistical learning using continuous speech flow (Toro et al., 2005) or visual displays (Baker et al., 2005; Turk-Browne et al., 2005). This conclusion comes as no surprise, because the major role played by selective attention in acquisition processes is an old and robust empirical finding (for another approach that emphasizes the role of attention, see Frensch et al., 1994).

2.32.4.2 Implicitness during the Test Phase

2.32.4.2.1 The lack of conscious knowledge about the study material

Is it possible to improve his/her performance without being conscious of what has been learned? A considerable amount of studies have addressed this question by exploring participants' explicit knowledge through postexperimental tests. Overall, a number of studies report that participants are aware of the knowledge they have acquired. However, other studies report that participants fail in the test of explicit knowledge. The question is: Are those negative results reliable? A number of potential drawbacks have been raised.

2.32.4.2.2 The Shanks and St. John information criterion

The first problem is linked to the fact that exploring whether knowledge is consciously represented primarily requires that the knowledge relevant for performing the task has been correctly identified. In an influential synthesis of the literature, Shanks and St. John (1994) coined this requirement as the 'information criterion.' The information criterion stipulates that the information the experimenter is looking for in the awareness test needs to match the information responsible for the performance change.

Although the cogency of this criterion may seem obvious, it must be realized that it entails that any conclusion about implicitness entirely depends on the response given to the 'what is learned' question raised in the prior sections. Any error in the hypothesized content of knowledge, far from being a "slightly embarrassing methodological glitch" (Reber, 1993, note p. 44, 114-115), has dramatic consequences on the inference that one may drawn about the implicit/explicit status of the acquired knowledge. For instance, Reber and Allen correctly pointed out that:

> . . . clearly a considerable proportion of subjects' articulated knowledge can be characterized as an awareness of permissible and nonpermissible letter pairs. (Reber and Allen, 1978: 210).

However, the authors did not realize that this form of knowledge was sufficient to account for performance. Instead, they attributed performance improvement to rule knowledge, which they concluded to be the result of unconscious abstraction. A large part of the earlier claims for the lack of conscious knowledge about the study material seemingly stems from this problem, also known as the problem of the 'correlated hypotheses' after the seminal studies by Dulany (1961) and Dulany et al. (1984).

2.32.4.2.3 The Shanks and St. John sensitivity criterion

According to Shanks and St. John, a second criterion is that the test of explicit knowledge is sensitive to all of the relevant conscious knowledge. A test of free recall, such as used in the early studies on IL, is notoriously insensitive. For instance, participants may not report some knowledge they have about the material structure, because they have a conservative response criterion that makes them respond only when their knowledge is held with high confidence, or simply because they think this knowledge is irrelevant or trivial. For this reason, most studies now involve a test of recognition, in which participants have to discriminate items belonging to the training materials from new items. However, performing no better than chance in a recognition test is not necessarily a proof that participants lack any explicit knowledge about the task. For instance, Reed and Johnson (1994) used a recognition test after an SRT task and observed that recognition scores were at chance. In an attempted replication involving the same procedure, Shanks and Johnstone (1999) found instead very high levels of recognition. The only difference between the two studies was that participants in Shanks and Johnstone were rewarded by an extra sum of money for each correct decision. Performing the recognition test is somewhat tedious, and presumably participants in the Reed and Johnson study were not motivated enough to make the effort required to perform the task correctly.

2.32.4.2.4 The problems of forgetting

In the standard procedure, the explicit tests are postponed after the task of IL, thus raising the problem of the retention of the knowledge exploited during the implicit test. For instance, Destrebecqz and Cleeremans (2001) reported chance-level recognition in an SRT paradigm (at least for a group of participants). Notably, the test of recognition was administered after participants had performed another task, in which they had to generate sequences under various instructions (see below). Shanks et al. (2003) attempted to reproduce Destrebecqz and Cleeremans' dissociation between RT measures and recognition scores, but in conditions in which the two kinds of measures were taken concurrently. In three experiments, they failed to replicate the Destrebecqz and Cleeremans' dissociation and obtained instead clear evidence of recognition. Note that the problems of forgetting are made especially important due to the fact that a recognition test necessarily includes the exposure to new sequences (generally half of the test items). Because these new sequences are highly similar to old sequences, they are prone to generate interference for the subsequent test trials.

2.32.4.2.5 The problem of the reliability of measures

The scores in implicit and explicit tasks are often found to be correlated. For instance, in SRT tasks, Perruchet and Amorim (1992) reported that Pearson correlations over the sequence trials between RT and recognition scores ranged, in three experiments, from .63 to .98. However, some authors (e.g., Willingham et al., 1993) have argued that evidence for unconscious knowledge was given by the fact that learning could be still observed when the analysis was restricted to the subgroup of items (or the subgroup of participants) for which no evidence of explicit knowledge was gathered. This argument is questionable, however. As discussed in Perruchet and Amorim (1992), the method, in effect, dichotomizes the scores on the implicit measure on the one hand, and on the explicit measure on the other, to assign the items or the participants to a fourfold contingency table. Then inference for dissociation is drawn from the observation that some items or some participants fall into the discordant cells of the contingency table, or in other words, that the correlation is not perfect. The problem with this method is that a prerequisite for obtaining a perfect correlation is perfect reliability of measures. This condition is highly unrealistic for psychological measures, especially for the scores on implicit tests (Meier and Perrig, 2000; Buchner and Brandt, 2003). Shanks and Perruchet (2002) have developed this reasoning into a quantitative model, which assumes that the sources of error plaguing implicit and explicit measure are independent. Although the model involved a single underlying memory variable, it turned out to be able to generate a dissociation between RTs and recognition in SRT tasks that mimics fairly well the dissociation the authors reported themselves (despite the temporal synchrony of measures).

2.32.4.2.6 An intractable issue?

To sum up, the current evidence for the lack of conscious knowledge about the study material is weak at best. There is currently no identified condition allowing one to obtain a reproducible dissociation. Most of the experiments reporting above-chance performance in implicit measures and chance-level performance in explicit tests have been replicated in more stringent conditions, and it turns out that, as a rule, the dissociation no longer appears when appropriate controls are made.

These data do not allow clear conclusions. On the one hand, the preceding discussion makes it clear that it is impossible to conclude to the existence of learning without any conscious counterpart. But, on the other hand, it should be also unwarranted to infer from the current findings that conscious awareness of the material structure is necessary for performance improvement. The first reason is a logical one, which has been met with regard to rule abstraction, namely, it is not possible to prove that something does not exist. There is yet another reason, linked to the fact that no task is process-pure, as has been well documented in the literature on implicit memory. This is especially true for the most sensitive tests, such as recognition. Jacoby (1983), and many others since, have argued that the relative fluency of perception, which relies on implicit process, may be used as a cue for discriminating old from new items in a recognition task, thus making a variable contribution to recognition judgments over and beyond a directed memory search factor. This entails that above-chance recognition after an IL task does not provide a compelling evidence for explicit knowledge. To date, it is not clear how further studies could solve this conundrum. Some authors (e.g., Higham et al., 2000) have suggested that those problems are intractable and should prompt researchers to give up any attempts to demonstrate learning without concurrent consciousness.

2.32.4.2.7 The subjective measures

The measures discussed so far are often called 'objective,' because it is the experimenter that judges the level of awareness of participants from their performance in specific tests. Another way of defining implicitness starts from the consideration of the phenomenal state of the participants such as it may be directly expressed. Two such 'subjective' measures of implicitness have been proposed in the literature, the guessing criterion and the zero-correlation criterion (Dienes et al., 1995). In both cases, participants are submitted to a test of explicit knowledge such as a recognition test, and they have to rate how confident they are about each decision. To check whether the guessing criterion is filled, the scores on the recognition test are restricted to those of the decisions that are accompanied by a subjective experience of guessing. If participants achieve above-chance discrimination while they report to be guessing, the guessing criterion is met. The zero-correlation criterion rests on the idea that, if knowledge is implicit, participants must not be more confident when they are correct than when they are incorrect. As a consequence, if participants have no introspection into the bases of their decisions, the correlation between confidence and accuracy should be null.

Can performances on standard IL tasks be called implicit according to these criteria? The literature again does not provide a clear response, with some studies reporting positive results and others negative results. In fact, the general picture appears similar to that observed with objective measures, with initially positive findings being not replicated when more sensitive measures are used. For instance, Dienes and Altman (1997) reported a zero correlation between confidence and accuracy in an artificial grammar learning task involving a transfer paradigm. Notably, participants had to assess their confidence on a continuous scale ranging from 50 to 100, where 50 was a complete guess and 100 was absolutely sure. Using the same scale, Tunney and Shanks (2003b) replicated this result. However, based on a study by Kunimoto et al. (2001), Tunney and Shanks reasoned that a binary confidence judgment could be more sensitive, maybe because participants might find it easier to express subjective states on a binary than on a continuous scale. When participants had to express their confidence on a binary scale, they were found to be systematically more confident in their correct decision than in their incorrect decision in several independent experiments. To conclude, irrespective of the *a priori* validity that one decides to ascribe to subjective measure of implicitness, it appears that there is to date no identified procedure that fulfills subjective criteria in a reproducible way.

2.32.4.2.8 The lack of control

One possible meaning of 'implicitness' is that of 'automaticity.' One of the key features usually attributed to automatic behavior is that it is irrepressible, irrespective of people's intentions to do so. Although recent literature on automaticity has questioned the possibility that any learned behavior - even reading, which is often construed as prototypical of automaticity – could actually be outside of control (e.g., Tzelgov et al., 1992), the question of whether the expression of knowledge in IL tasks shares this property deserves to be raised. Such a demonstration was provided by Destrebecqz and Cleeremans (2001) in an SRT task. In an application of Jacoby's process dissociation procedure (Jacoby, 1991) to this task, the authors asked participants to generate a sequence under two successive conditions during the test phase of an otherwise standard SRT procedure. In the first condition, they were told to generate the sequences they were previously exposed to, and if they fail to remember them, to generate sequences as they come to their minds (the inclusion instructions). Then participants had to produce a sequence of key-presses that did not overlap with the training sequence (the exclusion instructions). Crucially, participants - at least a subgroup of participants trained without any interval between the response to a target and the appearance of the next target - were influenced by the training sequences despite their intention to prevent this from happening. They performed in the same ways irrespective of the instructions, and under exclusion instructions, they generated the training sequence more than would be expected from an appropriate baseline.

These findings, however, have proven to be difficult to replicate. In the same conditions, Wilkinson and Shanks (2004) found that participants were more influenced by the training sequence under inclusion than under exclusion instructions (see also Destrebecqz and Cleeremans, 2003). In three experiments, Wilkinson and Shanks also failed to replicate the results according to which parts of the training sequence were generated more often under exclusion instructions than in the baseline, even after more extensive training than used by Destrebecqz and Cleeremans (2001) - although they did not get any negative difference either, as it could be expected if participants were able to withdraw the parts of the training sequence from influencing their production. Overall, these and others results (see for instance Dienes et al., 1995; Tunney and Shanks, 2003a) offer only quite limited evidence for the conclusion that knowledge gained in IL settings lies outside of intentional control.

2.32.4.2.9 The lack of intentional exploitation of acquired knowledge

The fact that participants are seemingly able to withdraw the influence of prior training when they are asked to do so does not mean that, under standard conditions, this influence is intentionally mediated. The lack of intentional exploitation of stored knowledge seems to be a hallmark of the real-world examples given at the outset of this chapter. Presumably nobody has the intuition of applying strategically a core of learned knowledge when speaking his maternal language, hearing music, or conforming to physical or social rules. Is this intuition confirmed in experimental studies?

The question is made difficult by the fact that, in most cases, influences expected from the intentional exploitation of conscious knowledge about the relevant aspects of the situation would have the very same effects as those induced by unconscious processes. As a consequence, it has been suggested that performance in IL tests can be accounted for by the use of explicit knowledge about various aspects of the experimental situation (Dulany et al., 1984; Shanks and St. John, 1994). It is certainly impossible to rule out this contention in general. However, it should be unwarranted to generalize it to all IL tasks. Indeed, there are cases in which the conscious exploitation of explicit knowledge does not coincide with the expected results of unconscious processing. One example is provided by the grammar learning studies involving preference judgments. Indeed, there should be *a priori* no reason for the knowledge about the material to be used to guide a preference judgment. However, participants consistently prefer grammatical items (e.g., Manza et al., 1998).

Vinter and Perruchet (2000) proposed a new task of IL that was especially devised to eliminate the potential influence of intentional control. When adults are asked to draw a closed geometrical figure such as a circle, their production exhibits a striking regularity. If they begin the circle in the lower half, they tend to rotate clockwise, and if they begin the circle in the upper half, they tend to rotate counter clockwise. In Vinter and Perruchet experiments, participants were guided to draw geometrical figures in such a way that this natural covariation was inverted. This training induced important and long-lasting modifications of subsequent free drawings. The point of interest is that, even if participants had become aware of the inverted covariation between the starting point and the rotation direction they experienced during the training session, they should have no reason to modify their usual mode of drawing as they did. This study provides clear evidence for an adaptive mode in which subjects' behavior becomes sensitive to the structural features of an experienced situation, without the adaptation being due to the intentional exploitation of subjects' explicit knowledge about these features.

2.32.4.3 Processing Fluency and Conscious Experience

Let us now reverse the direction of the potential relation between learning mechanisms and conscious thought, in order to examine the level of dependency of conscious thought with regard to IL.

An influential model of how training in IL settings leads to a change in performance posits that training induces a modification in the subjective experience of the learner. More specifically, the underlying idea is that training improves the fluency of perceptual processing for the studied materials. This account was initially proposed by Servan-Schreiber and Anderson (1990) in the context of their chunking theory. In fact, however, the improved processing fluency can also be attributed to other forms of knowledge, such as rules or memory for specific instances. Thus a fluency theory has been advocated as well by those researchers who maintain a role for rule-based processing (Zizak and Reber, 2004) and by those who consider that statistical computations are sufficient (e.g., Conway and Christiansen, 2005).

Let us examine how this account works in artificial grammar learning paradigms. The assumption is that, after training with a sample of grammatical strings of letters, new grammatical strings are processed fluently, or more precisely, more fluently than expected (Whittlesea and Williams, 2000), hence generating a feeling of familiarity leading itself to endorse the strings as grammatical. The two steps of this hypothesis have received experimental support. The fact that exposure to the training strings improved processing fluency has been shown by Buchner (1994). The test strings were displayed in such a way that they emerged progressively from noise, a procedure known as a perceptual clarification procedure. It turned out that grammatical strings were identified about 200 ms faster than ungrammatical ones. The fact that improved fluency in turn influences grammaticality judgments has been demonstrated using a method well documented in the literature on implicit memory, which consists in artificially enhancing the fluency of processing of selected items. During the test phase of an otherwise standard artificial grammar learning experiments, Kinder et al. (2003) exposed participants to test strings that did not differ in their grammatical status (they were all grammatical). The test strings were displayed in a perceptual clarification procedure as in Buchner (1994), except that some strings were clarified slightly faster than the others. Participants judged the former more often grammatical than the latter.

The fluency account suggests that IL modifies the subjective experience of the learner. However, the induced modifications appear to be quite minor, insofar as they are prompted by a gain of some fractions of second in processing speed. The frequent rapprochement of the concepts of IL and priming (e.g., Cleeremans et al., 1998; Conway and Christiansen, 2006; Kinder et al., 2003) is consonant with the idea that the training-induced modifications are relatively inconsequential. It is also possible to consider that the changes in the conscious experience of the learner are much more striking. For instance, in artificial grammar learning, participants normally learn to perceive the grammatical strings as a sequence of chunks the content of which is consonant with the structure of the grammar (e.g., the sequences of letters composing a recursive loop have high chance of being perceived as chunks, see Servan-Schreiber and Anderson, 1990; Perruchet et al., 2002). Likewise, in word-segmentation studies, the speechflow, which is initially perceived as an unorganized set of syllables, turns out to be perceived as a sequence of units, which match the words composing the language. More generally, an essential function of IL could be that of making the conscious perception and representation of the world isomorphic with world's deep structure. Because this change in subjective experience can be construed as a simple by-product of the attentional processing of the incoming information, Perruchet and Vinter (2002) have suggested the concept of 'self-organizing consciousness' to express the idea that IL shapes new conscious percepts and representations in a way which make them increasingly adaptive (see also Perruchet, 2005; Perruchet et al., 1997).

Neither the fluency account nor Perruchet and Vinter's (2002) self-organizing consciousness model is aimed to account for all behavioral changes observed in IL settings. For instance, although the fluency account is relatively consensual (partly due to the fact that it is mute with regard to the nature of knowledge inducing fluency), this account explains only a part of the performances observed in IL settings. Even in the artificial grammar learning paradigm, which is *a priori* a well-suited field of application, relative processing fluency does not seem to be able to account for the whole pattern of grammaticality judgments (Buchner, 1994; see also Zizak and Reber, 2004, p.23). However, these models point to the possibility of considering IL and consciousness not in terms of dissociation or independence, but rather in terms of dynamic interplay.

2.32.4.4 Summary and Discussion

To sum up, research of the last few decades has shown that it is surprisingly difficult to specify in what sense IL is implicit. The notion of unselective, nonattentional learning has vanished in light of studies demonstrating that learning requires at least some forms of attentional processing of the incoming information. Likewise, there are quite limited supports to claim that while they perform the implicit test participants (1) have no conscious knowledge about the study material, (2) have the subjective experience of guessing, or (3) have no control over the expression of their knowledge. Of course, it is possible to include one or the other of these features within a definition of IL, and some authors did so (for a sample of definitions, see Frensch, 1998). However, endorsing this kind of definition leads to the somewhat paradoxical consequence of giving to a research domain the objective of checking whether this domain actually exists. To date, there is no specified paradigm in which one or the other of these criteria can be asserted in a consensual and reproducible way.

A feature that can be retained with higher confidence is the lack of intentional exploitation of stored knowledge. This does not mean, obviously, that this condition is fulfilled in each and every study, but rather that the existence of the phenomenon can be reasonably asserted on the basis of reproducible evidence. Accepting the role of unconscious influences, however, does not lead us to conceive IL and conscious experiences as divorced one from each other. There is indeed extnsive evidence that these unconscious influences primarily affect the conscious experience of the learner.

2.32.5 Implicit Learning in Real-World Settings

2.32.5.1 Exploiting the Properties of Real-World Situations

Although most of research on IL uses artificial, laboratory situations, natural situations have been used on occasion to shed light on specific issues. In this case, only the test phase is carried out in well-controlled experimental conditions, while implicit training is assumed to have occurred previously in natural settings. For instance, Pacton et al. (2001) exploited the very extended time scale on which real-world learning takes place to examine whether transfer decrement (see the section titled 'The phenomenon of transfer: the interpretations') is a transitory or an enduring phenomenon. The issue is important, because it can be argued that the transfer decrement commonly observed in laboratory settings, which is one of the arguments used against a rule-based view, is simply due to the fact that training is not extensive enough to allow the full development of rule abstraction. Pacton et al. explored the development of the sensitivity to certain orthographic regularities not explicitly taught at school. They showed that the decrement in performance due to transfer persisted without any trend toward fading over the 5 years of experience with printed language that they examined, hence strengthening the claim that IL is not mediated by rule knowledge.

2.32.5.2 Exploiting our Knowledge about Implicit Learning

The knowledge gained in laboratory studies is aimed at improving our understanding of world-sized issues. Explicit loans from the IL literature have been made occasionally in a number of domains, including child development (Perruchet and Vinter, 1998b), second-language acquisition (e.g., Ellis, 1994; Robinson, 2002), spelling acquisition (Kemp and Bryant, 2003; Pacton et al., 2005; Pacton and Deacon, in press), and the development of gustatory preferences (Brunstrom, 2004). To various degrees, the concepts and the methodology of laboratory studies have inspired researchers to progress in the understanding of these domains. In regard of the potential relevance of IL mechanisms in these and other domains, much more could be made in this direction, however. The only domain in which a sizeable amount of literature has emerged concerns the relationships between IL and natural language acquisition (e.g., Gomez and Gerken, 2000). This rapprochement is partly due to the fact that research on language has evolved on its own toward methods – the use of artificial languages – and concepts – notably around the notion of statistical learning – that are also at the heart of IL research.

The practical applications of IL, for instance for educative purposes or the reeducation of neurological patients, appear to be still sparser. Some methods have evolved that exploit principles which can be *a posteriori* related to IL principles, such as using conditions as similar as possible to natural learning to teach second language (after Krashen, 1981) or reading (for a review, see Graham, 2000). An extensive literature also concern the use of errorless learning for reeducative purposes in a neuropsychological perspective (see review in Fillingham et al., 2003). But most of these attempts have been conducted without considering the possible contribution of IL research (for a recent exception, see Saetrevik et al., 2006). The explanations for this relative paucity are certainly manifold. One of them may be that learning in real-world situations most often involve some mixture of implicit (or incidental) and explicit (or intentional) learning. Now, the interactions between these forms of learning have not been at focus in the literature on IL, because, except in a few studies (e.g., Matthews et al., 1989), the objective has been to isolate implicit processes to examine them in their maximum state of purity. Further studies are needed to assess how, for instance, the learning of rules in explicit conditions may be combined with implicit statistical learning.

2.32.6 Discussion: About Nativism and Empiricism

Let us return to a question raised at the outset of this chapter, which stemmed from the lack of consideration during the behaviorist era of issues such as first-language acquisition, category elaboration, sensitivity to musical structure, acquisition of knowledge about the physical world, and various social skills. It was pointed out that this situation opened the door to the upsurge of a nativist perspective. Where do the studies reported in this chapter leave us?

At first glance, the mechanisms of IL, as they are revealed in laboratory studies, appear as definitely underpowered. The picture given by recent research stands far from the idea of the extraordinarily powerful processes that were imagined once, for instance by Lewicki et al., when they contended that "our nonconscious operating processing algorithms can do instantly and without external help" the same job as our conscious thinking achieves in relying on "notes (with flowchart or lists of if-then statements) or computer" (Lewicki et al., 1992, p.798). In fact, IL processes are probably unable to bring out to genuine rule knowledge, and the possibility of transfer are limited. In addition, the involvement of these processes seems to be dependent on selective attention. As pointed out above, the experimental study of learning around the 1960s was essentially devoted to classical and operant conditioning on the one hand and to the formation of concepts or problem-solving processes on the other hand. To make a long story short, IL mechanisms seem to be much nearer to the former than to the latter.

To be sure, experimental studies show that participants generally perform above chance in complex experimental settings. However, above-chance performance is generally attributable to the learning of some indirect, correlated aspects, which can be easily captured by elementary mechanisms. Everything happens as if IL often captured only nonessential aspects of the task. In experimental contexts, these correlated features are often considered as potential drawbacks, which need to be eliminated to reach the deep substance of the training material. For instance, studies in artificial grammar learning are often designed in such a way that bigram distribution becomes noninformative, studies in invariant learning often are controlled in such a way that the repetition of digits brings out no information about the invariant, and so on.

On the face of it, these data seem to provide fuel for a perspective in which the role of learning is minimized with regard to innate abilities. This is indeed the case if one considers that the knowledge base underlying the mastery of language and of the other high-level abilities alluded to above should be of the same form as the knowledge base that the scientist - for instance the linguist - acquires from an analytic investigation into his or her domain, that is, a formal, rule-based set of principles. This form of knowledge seems indeed to be definitely out of reach of IL processes.

However, a quite different perspective is possible. The general idea consists in assuming that learning in real-world setting proceeds as in the laboratory, that is, through the capture of correlated, apparently secondary aspects that can be grasped by elementary associative processes and that allow a good approximation of the behavior that would result from the knowledge of the formal structure of the domain. In order to be viable, such a perspective requires that the objective analysis of specific domains provides evidence for such correlated features. Quite interestingly, recent research on language has revealed a number of such features. The best-documented example is certainly the past-tense formation in English, in which it has been shown that regular and irregular verbs differ according to the distribution of their phonological and semantic features. Connectionist simulations have shown that exploiting those correlated cues leads to a very good approximation of the performance that would result from the formal knowledge of the *-ed* suffix rule, along with the knowledge of the exceptions (e.g., McClelland and Patterson, 2002). To consider another illustration, it has been shown that simple co-occurrence statistics (e.g., Redington et al., 1998) as well as phonological cues (e.g., Monaghan et al.,

2005) turn out to be highly informative about grammatical categories. These and other studies suggest that, as far as language is concerned, abstract classes and categories are often associated with simple statistical properties that make them tractable by general-purpose statistical learning mechanisms.

If further studies on language corpuses confirm and extend this kind of findings, and if the same kind of analysis proves to be successful in other high-level domains of competence, then IL mechanisms would appear extraordinarily powerful to promote behavioral adaptation. Indeed, those mechanisms are remarkably well-suited to exploit a massive amount of correlated cues. This approach appears to provide the first viable alternative to the nativist perspective that is still prevalent in the cognitive approach starting from Chomsky. The development of a full-blown empiricist alternative depends obviously upon further investigations on human learning processes, but also on the development of a nonconventional mode of description of the world humans are faced with.

References

Altmann GTM and Dienes Z (1999) Rule learning by seven-month-old infants and neural networks. *Science* 284: 875.

Anderson JR (1994) *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Aslin RN, Saffran JR, and Newport EL (1998) Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* 9: 321–324.

Atwell JA, Conners FA, and Merrill EC (2003) Implicit and explicit learning in young adults with mental retardation. *Amer. J. Ment. Retard.* 108: 56–68.

Baker CI, Olson CR, and Behrmann M (2005) Role of attention and perceptual grouping in visual statistical learning. *Psychol. Sci.* 15: 460–466.

Berry D (ed.) (1997) *How Implicit Is Implicit Learning?* Oxford: Oxford University Press.

Berry DC and Broadbent DE (1988) Interactive tasks and the implicit-explicit distinction. *Br. J. Psychol.* 79: 251–272.

Berry DC and Dienes Z (1993) *Implicit Learning: Theoretical and Empirical Issues*. Hillsdale NJ: Lawrence Erlbaum Associates.

Boucher L and Dienes Z (2003) Two ways of learning associations. *Cogn. Sci.* 27: 807–842.

Bright JEH and Burton M (1998) Ringing the changes: Where abstraction occurs in implicit learning. *Eur. J. Cogn. Psychol.* 10: 113–130.

Brooks LR (1978) Nonanalytic concept formation and memory for instances. In: Rosch E and Lloyd BB (eds.) *Cognition and Categorization*, pp. 169–211. Hillsdale NJ: Lawrence Erlbaum Associates.

Brooks LR and Vokey JR (1991) Abstract analogies and abstracted grammars: A comment on Reber and Mathews et al. *J. Exp. Psychol. Gen.* 120: 316–323.

Buchner A (1994) Indirect effects of synthetic grammar learning in an identification task. *J. Exp. Psychol. Learn. Mem. Cogn.* 20: 550–565.

Buchner A and Brandt M (2003) Further evidence for systematic reliability differences between explicit and implicit memory tests. *Q. J. Exp. Psychol.* 56A: 193–209.

Brunstrom JM (2004) Does dietary learning occur outside awareness? *Conscious. Cogn.* 13: 453–470.

Chambaron S, Ginhac D, Ferrel-Chapus C, and Perruchet P (2006) Implicit learning of a repeated segment in continuous tracking: A reappraisal. *Q. J. Exp. Psychol.* 59A: 845–854.

Channon S, Shanks D, Johnstone T, Vakili K, Chin J, and Sinclair E (2002) Implicit learning spared in amnesia? Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia* 1404: 1–13.

Cherry KE and Stadler MA (1995) Implicit learning of a nonverbal sequence in younger and older adults. *Psychol. Aging* 10: 3793–3794.

Christiansen MH, Allen J, and Seidenberg MS (1998) Learning to segment speech using multiple cues: A connectionist model. *Lang. Cogn. Processes* 13: 221–268.

Christiansen MH and Curtin S (1999) Transfer of learning: Rule acquisition or statistical learning? *Trends Cogn. Sci.* 3: 289–290.

Chun MM and Jiang Y (2003) Implicit, long-term spatial contextual memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 29: 224–234.

Cleeremans A and Jimenez L (1998) Implicit sequence learning: The truth is in the details. In: Stadler M and Frensch P (eds.) *Handbook of Implicit Learning*, pp. 323–364. Thousand Oaks, CA: Sage Publications.

Cleeremans A, Destrebecqz A, and Boyer M (1998) Implicit learning: News from the front. *Trends Cogn. Sci.* 2: 406–416.

Cleeremans A and McClelland JL (1991) Learning the structure of event sequences. *J. Exp. Psychol. Gen.* 120: 235–253.

Cock JJ, Berry DC, and Gaffan EA (1994) New strings for old: The role of similarity processing in an incidental learning task. *Q. J. Exp. Psychol.* 47: 1015–1034.

Cohen A, Ivry RI, and Keele SW (1990) Attention and structure in sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 16: 17–30.

Conway CM and Christiansen MH (2005) Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 31: 24–39.

Conway CM and Christiansen MH (2006) Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychol. Sci.* 17: 905–912.

Creel SC, Newport EL, and Aslin RN (2004) Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 30: 1119–1130.

Curran T (1997) Higher-order associative learning in amnesia: Evidence from the serial reaction time task. *J. Cogn. Neurosci.* 9: 522–533.

Destrebecqz A and Cleeremans A (2001) Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychon. Bull. Rev.* 8: 343–350.

Destrebecqz A and Cleeremans A (2003) Temporal effects in sequence learning. In: Jimenez L (ed.) *Attention and Implicit Learning*, pp. 181–213. Amsterdam: John Benjamins.

Detable C and Vinter A (2004) The preservation in time of implicit learning effects in children and adolescents with mental retardation. *Annee Psychol.* 104: 751–770.

Dienes Z (1992) Connectionist and memory-array models of artificial grammar learning. *Cogn. Sci.* 16: 41–79.

Dienes Z and Altmann G (1997) Transfer of implicit knowledge across domains: How implicit and how abstract? In: Berry D (ed.) *How Implicit Is Implicit Learning?* pp. 107–123. Oxford: Oxford University Press.

Dienes Z, Altmann GTM, Kwan L, and Goode A (1995) Unconscious knowledge of artificial grammars is applied

strategically. *J. Exp. Psychol. Learn. Mem. Cogn.* 21: 1322–1338.
Dienes Z, Broadbent D, and Berry D (1991) Implicit and explicit knowledge bases in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 17: 875–887.
Dienes Z and Longuet-Higgins C (2004) Can musical transformations be implicitly learned? *Cogn. Sci.* 28: 531–558.
Don AJ, Schellenberg EG, Reber AS, DiGirolamo KM, and Wang PP (2003) Implicit learning in children and adults with Williams syndrome. *Dev. Neuropsychol.* 23: 201–225.
Dulany DE (1961) Hypotheses and habits in verbal ''operant conditioning''. *J. Abnorm. Soc. Psychol.* 63: 251–263.
Dulany DE, Carlson A, and Dewey GI (1984) A case of syntactical learning and judgment: How conscious and how abstract? *J. Exp. Psychol. Gen.* 113: 541–555.
Eldridge LL, Masterman D, and Knowlton BJ (2002) Intact implicit habit learning in Alzheimer's disease. *Behav. Neurosci.* 116: 722–726.
Ellis NC (*ed*.) (1994) *Implicit and Explicit Learning of Languages*. New York: Academic Press.
Elman JL (1990) Finding structure in time. *Cogn. Sci.* 14: 179–211.
Fillingham JK, Hodgson C, Sage K, and Lambon Ralph MA (2003) The application of errorless learning to aphasic disorders: A review of theory and practice. *Neuropsychol. Rehabil.* 13: 337–363.
Fiser J and Aslin RN (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12: 499–504.
Fiser J and Aslin RN (2002) Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol. Learn. Mem. Cognit.* 28: 458–467.
Fiser J and Aslin RN (2005) Encoding multielement scenes: Statistical learning of visual features hierarchies. *J. Exp. Psychol. Gen.* 134: 521–537.
French RA and Cleeremans A (eds.) (2002) *Implicit Learning*, pp. 121–143. London: Routledge, Psychology Press.
Frensch PA (1998) One concept, multiple meanings. In: Stadler M and Frensch P (eds.) *Handbook of Implicit Learning*, pp. 47–104. Thousand Oaks, CA: Sage Publications.
Frensch PA, Buchner A, and Lin J (1994) Implicit learning of unique and ambiguous serial transitions in the presence and absence of a distractor task. *J. Exp. Psychol. Learn. Mem. Cogn.* 20: 567–584.
Gomez RL (1997) Transfer and complexity in artificial grammar learning. *Cogn. Psychol.* 33: 154–207.
Gomez R (2002) Variability and detection of invariant structure. *Psychol. Sci.* 13: 431–436.
Gomez RL and Gerken LA (1999) Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition* 70: 109–135.
Gomez RL and Gerken LA (2000) Infant artificial language learning and language acquisition. *Trends Cogn. Sci.* 4: 178–186.
Gomez RL, Gerken LA, and Schvaneveldt RW (2000) The basis of transfer in artificial grammar learning. *Mem. Cognit.* 28: 253–263.
Gomez RL and Schvanevelt RW (1994) What is learned from artificial grammar? Transfer tests of simple association. *J. Exp. Psychol. Learn. Mem. Cogn.* 20: 396–410.
Graham S (2000) Should the natural learning approach replace spelling instruction? *J. Educ. Psychol.* 92: 235–247.
Green REA and Shanks DR (1993) On the existence of independent learning systems: An examination of some evidence. *Mem. Cognit.* 21: 304–317.
Hauser MD, Newport EL, and Aslin RN (2001) Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition* 78: B53–B64.
Higham PA (1997) Dissociations of grammaticality and specific similarity effects in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 23: 1029–1045.
Higham PA, Vokey JR, and Pritchard JL (2000) Beyond dissociation logic: Evidence for controlled and automatic influences in artificial grammar learning. *J. Exp. Psychol. Gen.* 129: 457–470.
Hoffmann J and Sebald A (2005) When obvious covariations are not even learned implicitly. *Eur. J. Cogn. Psychol.* 17: 449–480.
Howard DV, Howard JH, Jakipse K, DiYanni C, Thompson A, and Somberg R (2004) Implicit sequence learning: Effects of level of structure, adult age, and extended practice. *Psychol. Aging* 19: 79–92.
Hsiao AT and Reber A (1998) The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious. In: Stadler M and Frensch P (eds.) *Handbook of Implicit Learning*, pp. 471–494. Thousand Oaks, CA: Sage Publications.
Hunt RH and Aslin RN (2001) Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *J. Exp. Psychol. Gen.* 130: 658–680.
Jacoby LL (1983) Perceptual enhancement: Persistent effects of an experience. *J. Exp. Psychol. Learn. Mem. Cogn.* 9: 21–38.
Jacoby LL (1991) A process dissociation framework: Separating automatic from intentional uses of memory. *J. Mem. Lang.* 30: 513–541.
Jimenez L (2003) *Attention and Implicit Learning*. Amsterdam: John Benjamins.
Jimenez L and Mendez C (1999) Which attention is needed for implicit sequence learning? *J. Exp. Psychol. Learn. Mem. Cogn.* 25: 236–259.
Johnstone T and Shanks DR (1999) Two mechanisms in implicit artificial grammar learning? Comment on Meulemans and Van der Linden (1997) *J. Exp. Psychol. Learn. Mem. Cogn.* 25: 524–531.
Karatekin C, Marcus DJ, and White T (2006) Oculomotor and manual indexes of incidental and intentional spatial sequence learning during middle childhood and adolescence. *J. Exp. Child Psychol.* 96: 107–130.
Kemp N and Bryant P (2003) Do beez buzz? Rule-based and frequency-based knowledge in learning to spell plural -s. *Child Dev.* 74: 63–74.
Kinder A and Assmann A (2000) Learning artificial grammars: No evidence for the acquisition of rules. *Mem. Cognit.* 28: 1321–1332.
Kinder A, Shanks DR, Cock J, and Tunney RJ (2003) Recollection, fluency, and the explicit/implicit distinction in artificial grammar learning. *J. Exp. Psychol. Gen.* 132: 551–565.
Knowlton BJ and Squire LR (1994) The information acquired during artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 20: 79–91.
Knowlton BJ and Squire LR (1996) Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *J. Exp. Psychol. Learn. Mem. Cogn.* 22: 169–181.
Krashen S (1981) *Second Language Acquisition and Second Language Learning*. New York: Prentice-Hall international.
Kuhn G and Dienes Z (2005) Implicit learning of non-local musical rules. *J. Exp. Psychol. Learn. Mem. Cogn.* 31: 1417–1432.
Kunimoto C, Miller J, and Pashler H (2001) Confidence and accuracy of near-threshold discrimination responses. *Conscious. Cogn.* 10: 294–340.
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: 7–10 August*

1991, Chicago, Illinois, pp. 215–221. Hillsdale, NJ: Lawrence Erlbaum Associates.
Lewicki P, Hill T, and Bizot E (1988) Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognit. Psychol.* 20: 24–37.
Lewicki P, Hill T, and Czyzewska M (1992) Nonconscious acquisition of information. *Am. Psychol.* 47: 796–801.
Lotz A and Kinder A (2006) Transfer in artificial grammar learning: The role of repetition information. *J. Exp. Psychol. Learn. Mem. Cogn.* 32: 707–715.
Manza L and Reber AS (1997) Representing artificial grammars: Transfer across stimulus forms and modalities. In: Berry DC (ed.) *How Implicit Is Implicit Learning?* pp. 73–106. Oxford: Oxford University Press.
Manza L, Zizak D, and Reber AS (1998) Artificial grammar learning and the Mere Exposure Effect: Emotional preference tasks and the implicit learning process. In: Stadler M and Frensch P (eds.) *Handbook of Implicit Learning*, pp. 201–222. Thousand Oaks, CA: Sage Publications.
Marcus GF, Vijayan S, Rao SB, and Vishton PM (1999) Rule learning by Seven-Month-Old Infants. *Science* 283: 77–80.
Mathews RC, Buss RR, Stanley WB, Blancard-Fields F, Cho J-R, and Druhan B (1989) Role of implicit and explicit processes in learning from examples: A synergistic effect. *J. Exp. Psychol. Learn. Mem. Cogn.* 15: 1083–1100.
McAndrews MP and Moscovitch M (1985) Rule-based and exemplar-based classification in artificial grammar learning. *Mem. Cognit.* 13: 468–475.
McClelland JL and Patterson K (2002) Rules or connections in past-tense inflections: What does the evidence rule out? *Trends Cogn. Sci.* 6: 465–472.
McClelland JL and Plaut DC (1999) Does generalization in infant learning implicate abstract algebra-like rules? *Trends Cogn. Sci.* 3: 166–168.
McGeorge P and Burton AM (1990) Semantic processing in an incidental learning task. *Q. J. Exp. Psychol.* 42A: 597–609.
Meier B and Perrig WJ (2000) Low reliability of perceptual priming: Consequences for the interpretation of functional dissociations between explicit and implicit memory. *Q. J. Exp. Psychol.* 53A: 211–233.
Meulemans T and Van der Linden M (1997) Chunk strength and artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 23: 1007–1028.
Meulemans T and Van der Linden M (2003) Implicit learning of complex information in amnesia. *Brain Cogn.* 52: 250–257.
Meulemans T, Van Der Linden M, and Perruchet P (1998) Implicit sequence learning in children. *J. Exp. Child Psychol.* 69: 199–221.
Monaghan P, Chater N, and Christiansen MH (2005) The differential role of phonological and distributional cues in grammatical categorisation. *Cognition* 96: 143–182.
Neal A and Hesketh B (1997) Episodic knowledge and implicit learning. *Psychon. Bull. Rev.* 4: 24–37.
Negash S, Howard DV, Japikse KC, and Howard JH, Jr. (2003) Age-related differences in implicit learning of non-spatial sequential patterns. *Aging Neuropsychol. Cognit.* 10: 108–121.
Newell BR and Bright JEH (2002) Well past midnight: Calling time on implicit invariant learning? *Eur. J. Cogn. Psychol.* 14: 185–205.
Newport EL and Aslin RN (2004) Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognit. Psychol.* 48: 127–162.
Onnis L, Monaghan P, Chater N, and Richmond (2005) Phonology impacts segmentation in online speech processing. *J. Mem. Lang.* 53: 225–237.
Pacton S and Deacon H (in press) The timing and mechanism of children's use of morphological information in spelling: Evidence from English and French. *Cogn. Dev.*
Pacton S, Fayol M, and Perruchet P (2005) Children's implicit learning of Graphotactic and Morphological regularities in French. *Child Dev.* 76: 324–339.
Pacton S and Perruchet P (in press) An attention-based associative account of adjacent and nonadjacent dependency learning. *J. Exp. Psychol. Learn. Mem. Cogn.*
Pacton S, Perruchet P, Fayol M, and Cleeremans A (2001) Implicit learning out of the lab: The case of orthographic regularities. *J. Exp. Psychol. Gen.* 130: 401–426.
Perruchet P (1994a) Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992) *J. Exp. Psychol. Learn. Mem. Cog.* 20: 223–228.
Perruchet P (1994b) Learning from complex rule-governed environments: On the proper functions of nonconscious and conscious processes. In: Umilta C and Moscovitch M (eds.) *Attention and Performance XV: Conscious and Nonconscious Information Processing*, pp. 811–835. Cambridge, MA: MIT Press.
Perruchet P (2005) Statistical approaches to language acquisition and the self-organizing consciousness: A reversal of perspective. *Psychol. Res.* 69: 316–329.
Perruchet P and Amorim MA (1992) Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *J. Exp. Psychol. Learn. Mem. Cog.* 18: 785–800.
Perruchet P, Gallego J, and Pacteau C (1992) A reinterpretation of some earlier evidence for abstractiveness of implicitly acquired knowledge. *Q. J. Exp. Psychol.* 44A: 193–210.
Perruchet P, Gallego J, and Savy I (1990) A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognit. Psychol.* 22: 493–516.
Perruchet P and Pacteau C (1990) Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *J. Exp. Psychol. Gen.* 119: 264–75.
Perruchet P and Pacton S (2006) Implicit learning and statistical learning: Two approaches, one phenomenon. *Trends Cogn. Sci.* 10: 233–238.
Perruchet P and Peereman R (2004) The exploitation of distributional information in syllable processing. *J. Neurolinguists.* 17: 97–119.
Perruchet P, Tyler MD, Galland N, and Peereman R (2004) Learning non-adjacent dependencies: No need for algebraic-like computations. *J. Exp. Psychol. Gen.* 133: 573–583.
Perruchet P and Vinter A (1998a) PARSER: A model for word segmentation. *J. Mem. Lang.* 39: 246–263.
Perruchet P and Vinter A (1998b) Learning and development: The implicit knowledge assumption reconsidered. In: Stadler M and Frensch P (eds.), *Handbook of Implicit Learning*, pp. 495–531. Thousand Oaks, CA: Sage Publications.
Perruchet P and Vinter A (2002) The self-organized consciousness. *Behav. Brain Sci.* 25: 297–388.
Perruchet P, Vinter A, and Gallego J (1997) Implicit learning shapes new conscious percepts and representations. *Psychon. Bull. Rev.* 4: 43–48.
Perruchet P, Vinter A, Pacteau C, and Gallego J (2002) The formation of structurally relevant units in artificial grammar learning. *Q. J. Exp. Psychol.* 55A: 485–503.
Pothos EM (2005) Expectations about stimulus structure in implicit learning. *Mem. Cognit.* 33: 171–181.
Pothos EM and Bailey TM (2000) The role of similarity in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 26: 847–862.
Reber AS (1967) Implicit learning of artificial grammars. *J. Verbal Learn. Verbal Behav.* 6: 855–863.
Reber AS (1969) Transfer of syntactic structure in synthetic languages. *J. Exp. Psychol.* 81: 115–119.
Reber AS (1989) Implicit learning and tacit knowledge. *J. Exp. Psychol. Gen.* 118: 219–35.

Reber AS (1993) *Implicit Learning and Tacit Knowledge: an Essay on the Cognitive Unconscious*. Oxford University Press, New York.

Reber AS and Allen R (1978) Analogy and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition* 6: 189–221.

Reber P and Lewis S (1977) Toward a theory of implicit learning: The analysis of the form and structure of a body of tacit knowledge. *Cognition* 5: 333–361.

Redington M and Chater N (1998) Connectionist and statistical approaches to language acquisition: A distributional perspective. *Lang. Cogn. Processes* 13: 129–191.

Redington M and Chater N (2002) Knowledge representation and transfer in artificial grammar learning. In: French RA and Cleeremans A (eds.) *Implicit Learning*, pp. 121–143. London: Routledge, Psychology Press.

Redington M, Chater N, and Finch S (1998) Distributional information: A powerful cue for acquiring syntactic categories. *Cogn. Sci.* 22: 425–469.

Reed J and Johnson P (1994) Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *J. Exp. Psychol. Learn. Mem. Cogn.* 20: 585–594.

Remillard G and Clark JM (2001) Implicit learning of first-, second-, and third-order transition probabilities. *J. Exp. Psychol. Learn. Mem. Cogn.* 27: 483–498.

Rescorla RA (1967) Pavlovian conditioning and its proper control procedures. *Psychol. Rev.* 74: 71–80.

Robinson P (2002) Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA. In: Robinson P (ed.) *Individual Differences and Instructed Language Learning*, pp. 211–266. Amsterdam: John Benjamins.

Roediger HL, Weldon MS, and Challis BH (1989) Explaining dissociations between implicit and explicit measures of retention: A processing account. In: Roediger HL and Craik FIM (eds.) *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving*, pp. 3–39. Hillsdale, NJ: Erlbaum.

Rowland LA and Shanks DR (2006a) Sequence learning and selection difficulty. *J. Exp. Psychol. Hum. Percept. Perform.* 32: 287–299.

Rowland LA and Shanks DR (2006b) Attention modulates the learning of multiple contingencies. *Psychon. Bull. Rev.* 13: 643–648.

Saetrevik B, Reber R, and Saunnum P (2006) The utility of implicit learning in the teaching of rules. *Learn. Instruct.* 16: 363–373.

Saffran JR (2002) Constraints on statistical language learning. *J. Mem. Lang.* 47: 172–196.

Saffran JR, Aslin RN, and Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274: 1926–1928.

Saffran JR, Johnson EK, Aslin RN, and Newport EL (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70: 27–52.

Saffran JR, Newport EL, Aslin RN, Tunick RA, and Barrueco S (1997) Incidental language learning. *Psychol. Sci.* 8: 101–105.

Saffran JR and Wilson DP (2003) From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy* 4: 273–284.

Schwartz BL, Howard DV, Howard JH, Hovaguimian A, and Deutsch SI (2003) Implicit learning of visuospatial sequences in schizophrenia. *Neuropsychology* 17: 517–533.

Seger CA (1994) Implicit learning. *Psychol. Bull.* 115: 163–196.

Seidenberg MS and Elman JL (1999) Do infants learn grammar with algebra or statistics? *Science* 284: 433.

Servan-Schreiber D and Anderson JR (1990) Learning artificial grammars with competitive chunking. *J. Exp. Psychol. Learn. Mem. Cogn.* 16: 592–608.

Shanks DR (1995) *The Psychology of Associative Learning*. Cambridge: Cambridge University Press.

Shanks DR (2005) Implicit learning. In: Lamberts K and Goldstone R (eds.) *Handbook of Cognition*, pp. 202–220. London: Sage.

Shanks DR, Channon S, Wilkinson L, and Curran HV (2006) Disruption of sequential priming in organic and pharmacological amnesia: A role for the medial temporal lobes in implicit contextual learning. *Neuropsychopharmacology* 31: 1768–1776.

Shanks DR and Johnstone T (1999) Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *J. Exp. Psychol. Learn. Mem. Cogn.* 25: 1435–1451.

Shanks DR, Johnstone T, and Staggs L (1997) Abstraction processes in artificial grammar learning. *Q. J. Exp. Psychol.* 50A: 216–252.

Shanks DR and Perruchet P (2002) Dissociation between priming and recognition in the expression of sequential knowledge. *Psychon. Bull. Rev.* 9: 362–367.

Shanks DR and St. John MF (1994) Characteristics of dissociable human learning systems. *Behav. Brain Sci.* 17: 367–447.

Shanks DR, Wilkinson L, and Channon S (2003) Relationship between priming and recognition in deterministic and probabilistic sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 29: 248–261.

Shanks DR and St. John MF (1994) Characteristics of dissociable human learning systems. *Behav. Brain Sci.* 17: 367–447.

Shukla M, Nespor M, and Mehler J (2007) An interaction between prosody and statistics in the segmentation of fluent speech. *Cognit. Psychol.* 54: 1–32.

Smith EE, Langston C, and Nisbett RE (1992) The case for rules in reasoning. *Cogn. Sci.* 16: 1–40.

Smith JG and McDowall J (2006) When artificial grammar acquisition in Parkinson's disease is impaired: The case of learning via trial-by-trial feedback. *Brain Res.* 1067: 216–228.

Stadler MA (1989) On learning complex procedural knowledge. *J. Exp. Psychol. Learn. Mem. Cogn.* 15: 1061–1069.

Stadler MA (1995) Role of attention in implicit learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 21: 674–685.

Stadler MA and Frensch P (1994) Whither learning, whither memory? *Behav. Brain Sci.* 17: 423–424.

Stadler MA and Frensch P (1998) *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage Publications.

Thiessen ED and Saffran JR (2003) When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev. Psychol.* 39: 706–716.

Toro JM, Sinnett S, and Soto-Faraco S (2005) Speech segmentation by statistical learning depends on attention. *Cognition* 97: B25–B34.

Toro JM and Trobalon JB (2005) Statistical computation over a speech stream in a rodent. *Percept. Psychophys.* 67: 867–875.

Tunney RJ and Altmann GTM (1999) The transfer effect in artificial grammar learning: Re-appraising the evidence of transfer of sequential dependencies. *J. Exp. Psychol. Learn. Mem. Cogn.* 25: 1322–1333.

Tunney RJ and Shanks DR (2003a) Does opposition logic provide evidence for conscious and unconscious processes in artificial grammar learning? *Conscious. Cogn.* 12: 201–218.

Tunney RJ and Shanks DR (2003b) Subjective measures of awareness and implicit cognition. *Mem. Cognit.* 31: 1060–1071.

Turk-Browne NB, Jungé J, and Scholl BJ (2005) The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* 134: 552–564.

Turner CW and Fischler IS (1993) Speeded test of implicit knowledge. *J. Exp. Psychol. Learn. Mem. Cogn.* 19: 1165–1177.

Tzelgov J, Henik A, and Berger J (1992) Controlling Stroop effects by manipulating expectations for color words. *Mem. Cognit.* 20: 727–735.
Vakil E, Kraus A, Bor B, and Groswasser Z (2002) Impaired skill learning in patients with severe closed-head injury as demonstrated by the serial reaction time (SRT) task. *Brain Cogn.* 50: 304–315.
Vinter A and Detable C (2003) Implicit learning in children and adolescent mental retardation. *Amer. J. Ment. Retard.* 108: 94–107.
Vinter A and Perruchet P (2000) Implicit learning in children is not related to age : Evidence from drawing behavior. *Child Dev.* 71: 1223–1240.
Vokey JR and Brooks LR (1992) Salience of item knowledgein learning artificial grammar. *J. Exp. Psychol. Learn. Mem. Cogn.* 18: 328–344.
Whittlesea BWA and Dorken MD (1993) Incidentally, things in general are incidentally determined: An episodic-processing account of implicit learning. *J. Exp. Psychol. Gen.* 122: 227–248.
Whittlesea BWA and Dorken MD (1997) Implicit learning: Indirect, not unconscious. *Psychon. Bull. Rev.* 4: 63–67.
Whittlesea BW and Williams LD (2000) The discrepancy-attribution hypothesis: II Expectation, uncertainty, surprise, and feelings of familiarity. *J. Exp. Psychol. Learn. Mem. Cogn.* 27: 14–33.
Whittlesea BWA and Wright RL (1997) Implicit (and explicit) learning: Acting adaptively without knowing the consequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 23: 181–200.
Wilkinson L and Shanks DR (2004) Intentional control and implicit sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 30: 354–369.
Willingham DB, Greeley T, and Bardone AM (1993) Dissociation in a serial response time task using a recognition measure: Comment on Perruchet and Amorim (1992). *J. Exp. Psychol. Learn. Mem. Cogn.* 19: 1424–1430.
Willingham DB, Nissen MJ, and Bullemer P (1989) On the development of procedural knowledge. *J. Exp. Psychol. Learn. Mem. Cogn.* 15: 1047–1060.
Wright RL and Burton AM (1995) Implicit learning of an invariant: Just say no. *Q. J. Exp. Psychol.* 48A: 783–796.
Zizak DM and Reber AR (2004) Implicit preferences: The role(s) of familiarity in the structural mere exposure effect. *Conscious. Cogn.* 13: 336–362.