Psychological Research (2005) 69: 383–398
DOI 10.1007/s00426-004-0207-4

ORIGINAL ARTICLE

Maud Boyer · Arnaud Destrebecqz · Axel Cleeremans

Processing abstract sequence structure: learning without knowing, or knowing without learning?

Received: 18 October 2004 / Accepted: 26 November 2004 / Published online: 15 March 2005

Abstract Constant interaction with a dynamic environment—from riding a bicycle to segmenting speech—makes sensitivity to the sequential structure of the world a fundamental dimension of information processing. Accounts of sequence learning vary widely, with some authors arguing that parsing and segmentation processes are central, and others proposing that sequence learning involves mere memorization. In this paper, we argue that sequence knowledge is essentially statistical in nature, and that sequence learning involves simple associative prediction mechanisms. We focus on a choice reaction situation introduced by Lee (1997), in which participants were exposed to material that follows a single abstract rule, namely that stimuli are selected randomly, but never appear more than once in a legal sequence. Perhaps surprisingly, people can learn this rule very well. Or can they? We offer a conceptual replication of the original finding, but a very different interpretation of the results, as well as simulation work that makes it clear how highly abstract dimensions of the stimulus material can in fact be learned based on elementary associative mechanisms. We conclude that, when relevant, memory is optimized to facilitate responding to events that have not occurred recently, and that sequence learning in general always involves sensitivity to repetition distance.

M. Boyer (✉)
Cognitive Neuroscience Research Unit,
Université Libre de Bruxelles,
Avenue F.-D. Roosevelt, 50 CP 191,
1050 Bruxelles, Belgium
E-mail: mboyer@ulb.ac.be
Tel.: +322-6504228
Fax: +322-6502209

A. Destrebecqz
LEAD, Université de Bourgogne,
UMR 5022 CNRS, Dijon, France

A. Cleeremans
Cognitive Science Research Unit,
Université Libre de Bruxelles, Avenue F.-D. Roosevelt,
50 CP 191, 1050, Bruxelles, Belgium

Introduction

While there is widespread agreement that each mental state is necessarily caused by a neural state (this is the basic principle that underlies the ongoing "search for the neural correlates of consciousness" (see e.g., Frith, Perry, & Lumer, 1999), the reverse claim that each neural state necessarily has a phenomenal correlate is unlikely to be true. Exploring cases where there are demonstrable changes in neural activity or in behavior without concomitant changes in subjective experience is what the study of implicit cognition is about. Understanding the mechanisms of implicit cognition thus entails that a stand is taken on what it means to be conscious. However, thinking about and conducting research on consciousness is clearly hard: Nobody agrees on exactly what it is that we are trying to understand; different, often competing theories exist; there is continuing debate about how one should interpret the increasingly abundant empirical evidence. The difficulty, of course, comes from the fact that conscious experience is a private phenomenon/process that depends on your current state and on your history: your perception of a toad will not be exactly the same as mine, and it will not even be exactly the same for you today as it was yesterday. Furthermore, I do not have direct access to your mental states, and, some would argue, neither do I have perfect access to my own mental states (or if I do, I am often likely to be mistaken in different ways, see e.g., Dennett, 1991; Nisbett & Wilson, 1977; Wegner, 2002). What is it, then, that we can hope to explain? How are we to proceed?

This assessment will strike many as overly grim, and yet, after several years of thinking about these issues, we do not see an easy way out. One solution is to adopt a thoroughly pragmatic approach to the relationships between conscious and unconscious cognition: Focus on the mechanisms now, worry about the difficult conceptual issues later. This is essentially the perspective we adopt in this paper. However, it nevertheless remains important to spell out our general perspective on the larger issues, for fear of losing sight of what we are ultimately trying to achieve. In this respect, it is worth pointing out that the rather pessimistic tone of our introductory comments stems not only from the now familiar challenges that reconciling subjective and objective approaches to the mind entails, but also from the rather unrealistic expectations that renewed interest in consciousness has triggered in the research community. Somehow, deep down, we continue to expect that there will be a single ‘‘aha’’ moment when an obscure neuroscientist suddenly comes up with ‘‘the’’ explanation for consciousness. Needless to say, this is not going to happen: Functional accounts of consciousness that take as a starting point that it is a single, static property associated with some mental states are doomed to fail, for consciousness is neither ‘‘a single thing’’ nor is it static. Instead, consciousness refers to several, possibly dissociable, aspects of information processing, and it is a fundamentally dynamic, graded, process. This point is worth emphasizing because for a long while, research on implicit learning (or any other domain where evidence of implicit cognition has been collected, such as subliminal priming) has been characterized by similar lines of thought: Learning is either implicit or explicit, consciousness of some stimulus is necessarily all-or-none, pure measures of implicit or explicit influences can be devised, and so on.

Recently, Cleeremans and Jiménez (2002) introduced a conceptual framework that attempts to acknowledge the graded character of conscious experience as well as the graded effects that learned knowledge can exert on behavior (Cleeremans, forthcoming). The framework takes as its starting points:

1. That the main function of consciousness is to make flexible, adaptive control over behavior possible
2. That consciousness is best viewed as involving a graded continuum expressed over properties if representation
3. That learning is a mandatory process that always accompanies information processing, and through which our conscious representations of the world are made to reflect those contents which are most in need of control at some point in time

The framework takes its roots in connectionist (or more generally, sub-symbolic) models of the mind (Rumelhart & McClelland, 1986).

Decades of exploration of such models and of related exemplar- and memory-based frameworks (e.g., Brooks, 1978; Hintzmann, 1986; Perruchet & Vinter, 2003) have yielded a number of insights that are particularly relevant for the study of implicit learning (see Cleeremans, forthcoming).

Chief among those is the notion that sensitivity to some regularity does not necessarily imply that the regularity itself is represented as an object of representation. We call this the ‘‘principle of emergent representation’’ (Cleeremans, 1997, forthcoming) and it simply means that it should not be concluded, from observing sensitivity to some regularity in human participants, that these participants possess propositional, explicit, symbolic, and conscious knowledge of the regularity. The case is made most clear by the fact that native speakers of a natural language are perfectly capable of uttering grammatically correct expressions without even coming close to possessing verbalizable knowledge of the underlying grammar. Yet, some representation of the underlying set of regularities must nevertheless exist; it is just that this knowledge is not accompanied by relevant meta-knowledge, and that it is stored in a form that is not easily expressed. In other words, observing rule-like behavior does not necessarily imply that the subject's knowledge is rule-based: A system can act as though it were following a rule, yet nothing like a rule is actually stored in the system's memory (see e.g., Hinton, 1986; Pacton, Perruchet, Fayol, & Cleeremans, 2001; Redington & Chater, 1996). Instead, and along with others (e.g., Shanks & John, 1994), we think that genuine rule-based knowledge is necessarily conscious knowledge (see also Cleeremans & Destrebecqz, in press).

What, then, makes a representation a good candidate for availability to conscious experience? Cleeremans and Jiménez (2002) suggested that this depends in a graded manner on several properties of representations, namely their relative strength, their distinctiveness (i.e., whether they refer to specific instances or to features shared by many exemplars; roughly episodic versus semantic memory), and their stability in time. Strong, stable, and distinctive representations of some content are accessible in a manner that weaker representations are not. Crucially, such weak representations are nevertheless susceptible of influencing behavior, for instance through priming effects emerging as a result of the involvement of several such weak traces.

In this paper, we would like to illustrate the principle of ‘‘emergent representation’’ in action once again, through a reinterpretation of sequence learning data initially reported by Lee (1997). We will show how Lee's conclusion that her data suggested unconscious knowledge of a highly abstract rule can in fact be dismissed in favor of an account that is entirely based on associative learning mechanisms operating on exemplars. We show how both human subjects and models can appear to act as though they possess rule-based knowledge of the material, yet neither subjects nor the models, of course, are able to verbalize anything about the simple regularity to which they are nevertheless sensitive. In the next section, we briefly overview sequence learning situations, upon which the experimental and simulation work we describe in this paper is based.

Sequence learning

Sequence learning is a fundamental process involved in the many different cognitive skills required for successful interaction with an intensely dynamic environment. Among those skills, language is probably the most complex, and the role that elementary associative sequence learning processes may play in its development has recently begun to be explored anew. For instance, Saffran, Newport, Aslin, and Barrueco (1997) showed how incidental exposure to artificial language-like auditory material (e.g., *bupadapatubitutibu*.... was sufficient to enable participants to segment the continuous sequence of sounds they had heard into artificial words (e.g., *bupada*, *patubi*, etc.) of which it consisted, as evidenced by their performance in a subsequent recognition test. Based on these data, Saffran et al. suggested that word segmentation abilities develop based on mechanisms that exploit the statistical regularities present in sequences of events, such as the fact that the transitional probabilities of successive syllables are higher within words than between words. Further studies (Saffran, Johnson, Aslin, & Newport, 1999) showed that this statistical learning ability was not uniquely tied to linguistic materials: Both adults and 8-month-old infants were able to segment a continuous non-linguistic auditory sequence (made up of "tone words"). Interestingly, Saffran et al. rooted their interpretation of such findings in the apparently remote literature dedicated to implicit learning.

The connection is obvious as soon as we recognize that language acquisition, like implicit learning (for reviews, see Berry & Dienes, 1993; Cleeremans, Destrebecqz, & Boyer, 1998), is likely to involve, at least in part, incidental learning of complex information organized at different levels. In particular, research on sequence learning (see Clegg, DiGirolamo, & Keele, 1998, for an overview) has, over the past decade or so, provided a steady stream of relevant evidence suggesting that participants exhibit detailed sensitivity to the sequential structure through differences in their reaction times to stimuli that are or are not predictable based on the temporal context set by previous elements. In typical sequence-learning situations, participants are asked to react to each element of sequentially structured and typically visual sequences of events (Nissen & Bullemer, 1987). Several variants of this basic paradigm exist. In *rule*-based paradigms, sequences either conform or fail to conform to an abstract rule that describes permissible transitions between successive stimuli. Rule-based paradigms can in turn involve either deterministic (Lewicki, Hill, & Bizot, 1988) or probabilistic rules, as when the stimulus material is generated based on the output of finite-state grammars (Cleeremans, 1993; Cleeremans & McClelland, 1991). By contrast, in the more common *simple repeating sequence* paradigm, a single sequence containing fixed regularities is repeated many times to produce the training set (Nissen & Bullemer, 1987; Reed & Johnson, 1994). Sequence learning paradigms now constitute one of the main experimental situations through which to explore the mechanisms of implicit learning.

A perennial question in this context is to determine exactly what people learn about when exposed to sequentially-structured stimulus material. Perhaps unsurprisingly, it is often the case that different accounts are partially or completely consistent with the data. Consider for instance a sequence learning situation in which the stimulus material consists of a simple repeating sequence such as "ABCDBA". When exposed to this material in the context of a choice reaction situation, participants could:

1. Learn something about the generation rules (the "abstractionist" position, see e.g., Reber, 1967)
2. Memorize the entire sequence (Brooks, 1978)
3. Become sensitive to the frequency of specific repeating fragments of the sequence (Perruchet & Gallego, 1997; Servan-Schreiber & Anderson, 1990)
4. Learn something about the conditional probability of occurrence of each element in the context of the previous elements (Cleeremans & McClelland, 1991; Jiménez, Mendez, & Cleeremans, 1996)
5. Learn about other aspects of the material such as specific motor patterns (e.g., alternations, trills, or more abstract patterns; see Koch & Hoffmann, 2000) or repetition distance between successive occurrences of the same stimulus (Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998)

Cleeremans and Jiménez (1998) suggested that these different accounts may in fact turn out to be descriptively equivalent, and concluded that the core processes involved in sequence learning are best thought of as involving elementary associative learning processes that result in a progressively developing sensitivity to the statistical constraints contained in the material (see also Stadler, 1992). Such processes are well instantiated by connectionist models such as the Simple Recurrent Network (Cleeremans & McClelland, 1991; Elman, 1990).

In this context, Lee (1997) described an interesting sequence learning situation which, at first sight, seems to challenge traditional accounts of sequence learning. Indeed, Lee's material consisted of a random selection of the 720 (6!) sequences of six elements that are consistent with the following simple constraint: Each of the six different elements can only appear once in each six-element sequence. For instance, the sequences "123456" or "236145" are both legal because each stimulus appears only once. The sequence "235451", however, does not follow the rule because element '5' appears twice and element '6' is missing. This rule thus results in a probability gradient across the six positions within each sequence, such that the first element of any legal sequence is always completely unpredictable, and such that the subsequent elements become increasingly predictable based on the context set by the previous elements. The final element of each legal sequence is thus always completely predictable based on the first five elements.

Lee's material therefore contains almost no structure but for the single highly abstract structural property described by the generation rule. Nevertheless, Lee showed that participants trained in this material tend to respond faster to stimuli that occur in serial position 6 than to stimuli that appear in serial position 1, thereby indicating that they had learned something about the structure of the material. Of course, participants ignored the existence of the six-element cycles and nothing marked the beginning or end of the sequences.

As Lee indicated, traditional theories of sequence learning (such as those listed above) may have a hard time accounting for the data. Indeed, theories that assume instance memorization would have difficulty in this case because the material simply fails to contain repeated instances. Fragment-based accounts appear likewise implausible because even three-element fragments fail to convey much information about the relevant regularities (but see below). For instance, the fragment "123" may end in any of the six serial positions and be followed by any of the six possible elements but "3" (stimulus repetitions were forbidden). No salient patterns are detectable in the material. Lee concluded that "both parsing and short-term memory mechanisms must be involved" (p. 428), and that models based on simple associative learning mechanisms, such as the simple recurrent network (SRN), were probably incapable of learning this stimulus material.

We present two experiments and accompanying simulations in the rest of this article. Experiment 1 was first described in Boyer, Destrebecqz, and Cleeremans (1998). Here, we report on additional analyses and show that participants' sensitivity to the rule used to generate the material can actually be understood based on the operation of elementary associative mechanisms that do not involve parsing of any kind. We also challenge the idea that any learning is involved in this situation, and suggest instead that the structure of the material merely reinforces existing "negative recency" biases (Jarvik, 1951).

In the second experiment, we suggest, through new empirical and simulation work, that both Lee's and Boyer et al.'s data can be accounted for based on a preexisting tendency to prepare responses to stimuli that have not occurred recently, and further speculate that this bias may result from continuously reinforced preexperimental exposure to environmental regularities for which negative recency has predictive value.

Experiment 1

Method

Participants

Twelve participants aged 18–24, all undergraduate students in the Psychology Department of the Université Libre de Bruxelles, took part in the experiment. They were paid a fee of about $14 and could earn an additional bonus of up to $9 based on performance of the task (see below).

Apparatus and display

The experiment was run on Macintosh computers. The display consisted of six dots arranged in a horizontal line on the computer's screen and separated by intervals of 3 cm. Each screen position corresponded to a key on the computer's keyboard. The stimulus was a small black circle 0.35 cm in diameter that appeared on a white background, centered 1 cm below one of the six dots.

Procedure

Participants were exposed to 24 blocks of a six-choice reaction time (RT) task. Each block consisted of 180 trials, for a total of 4,320 trials. In each trial, a stimulus appeared in one of the six positions, and participants responded as fast and as accurately as possible by pressing the corresponding key. The target was removed as soon as a key had been pressed, and the next stimulus appeared after a 120-ms interval. Erroneous responses were signaled by means of a tone. Participants were exposed to two practice blocks of 18 trials each before the onset of the experiment. Short rest breaks occurred between any two experimental blocks. During these breaks, participants were informed about their performance and bonus money earned so far. This amount was computed for each block based on both accuracy and speed. A longer rest break occurred after 12 experimental blocks.

Stimulus material

The stimulus set consisted of the 720 (6!) sequences of six elements that were consistent with the following simple constraint: Each element could only appear once in each sequence. Each of the 24 training blocks was produced by randomly selecting (without replacement) 30 legal sequences and by concatenating them in random order with the only constraint that the last element of any sequence could not be identical to the first element of the next sequence. The boundaries between sequences of six elements were not marked. Each participant was exposed to a different random order of the 24 training blocks. To control for finger and hand effects, sequence elements were assigned to different response keys in a 6×6 Latin square design, so that each sequence element was assigned to each key exactly once. Each of the six mappings was then used for two participants.

Results

Reaction time performance

Outliers (RTs above or below two standard deviations from the mean) and incorrect responses were discarded from subsequent analyses and represented 10.6% of all responses. To assess whether participants were sensitive to sequential structure, we first examined whether RTs reflected the serial position effect described by Lee (1997). Recall that the stimulus material was such that stimuli appearing as the first element of a sequence were completely random according to the generation rules, and that stimuli appearing in subsequent serial positions were increasingly predictable, up to serial position 6, where the stimulus was completely determined.

Figure 1a shows average RTs over the entire experiment, plotted separately for each serial position. We replicate Lee's original finding: The figure makes it clear that participants' responses are strongly influenced by serial position within each sequence. Indeed, RTs decrease linearly from the first to the sixth serial position, with a difference of about 30 ms between the first and last positions. These impressions were confirmed by a two-way ANOVA with Block (24 levels) and Serial Position (six levels) as repeated measures factors, which revealed a significant main effect of Block, $F(23, 253) = 48.14$, $p < .0001$, MSE = 2,997.7, and of Serial Position, $F(5, 55) = 22.26$, $p < .0001$, MSE = 1,554.3—significant linear tendency was obtained, $F(1, 11) = 29.85$, $p < .001$. The interaction also reached significance, $F(115, 1{,}265) = 1.26$, $p < .05$, MSE = 754.4, albeit more detailed analyses (see below) do not confirm that this interaction should be taken as evidence of learning.

Learning

Figure 1b shows how the serial position effect described above changes over training. It is clear that the effect is already present early on, and that the slope of the curves corresponding to different moments during training does not appear to change much. An ANOVA with Block (four levels) and Serial Position (six levels) on this aggregate data again produced a significant interaction between Blocks and Position, $F(15, 165) = 1.77$, $p < .05$, MSE = 185.3. However, this significant interaction offers little evidence of learning. Indeed, planned comparisons showed that the difference between RTs to positions 1 and 6 stimuli (30 ms) is already significant over the first two blocks, $F(1, 11) = 16.87$, $p < .001$, MSE = 309.5, and that it stays relatively constant up until the last two blocks, during which it averages 41 ms, $F(1, 11) = 23.56$, $p < .001$, MSE = 443.74). In short, there is in fact little evidence of any learning in this situation, short of unspecific practice effects: The serial position effect emerges very early in training and remains quite constant over the entire experiment.

Finally, participants were also asked questions about whether they had noticed anything about the structure of the stimulus material. All participants indicated that they thought that the stimulus material was completely random and none noticed lag structure. Admittedly, these are qualitative data at best, but in this case, the fact that none of the participants (throughout this and other

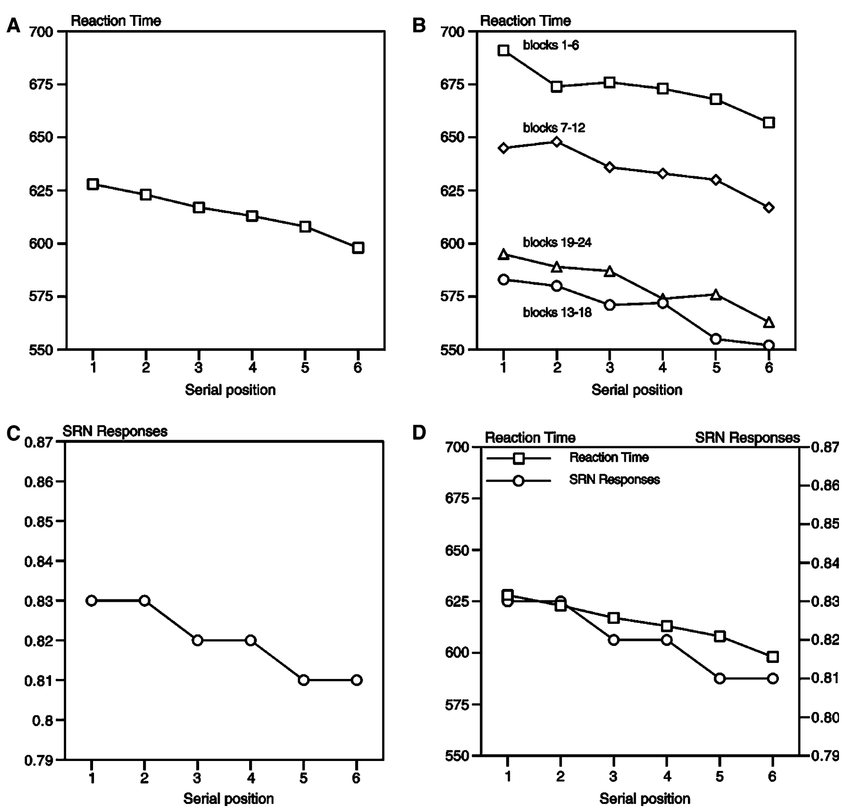

Fig. 1 The position effect. **a** Mean reaction times as a function of serial position over the entire experiment. **b** Mean reaction times as a function of serial position, plotted separately for blocks 1–6, 7–12, 13–18, and 19–24. **c** Mean simple recurrent network (*SRN*) responses as a function of serial position over the entire experiment. **d** Both human subjects and the SRN model exhibit a linear serial position effect.

experiments) volunteered anything at all when asked about the material's structure convinced us that there is little doubt that participants indeed have very little or no introspective awareness of the structure contained in the material. Note also that the very nature of the material makes it rather challenging to devise effective tests of awareness in this situation. It is unclear, for instance, which sequences could be contrasted in the context of a recognition task, or what we would look for in generation data as evidence of participants' ability to successfully deploy learned knowledge in a direct test.

Discussion

Taken at face value, the data we obtained seem to confirm that participants are able to somehow segment the stimulus material in chunks of six elements—the finding that prompted Lee to conclude that parsing mechanisms are involved. However, and somewhat surprisingly, our data also indicate that participants appeared to be sensitive to sequential structure right from the beginning of the experiment. A re-examination of Lee's results likewise seems to suggest that the serial position effect is already present very early during training. In the following, we will show that parsing mechanisms of any kind turn out to be unnecessary to account for performance, and suggest an account of how participants' sensitivity to the sequential structure may in fact develop before they are first exposed to the task.

Lee's analysis rests on the assumption that it is necessary for participants to encode the serial position for them to exhibit faster RTs to serial position 6 stimuli than to other positions, and therefore to somehow parse the material in successive chunks of six elements with the

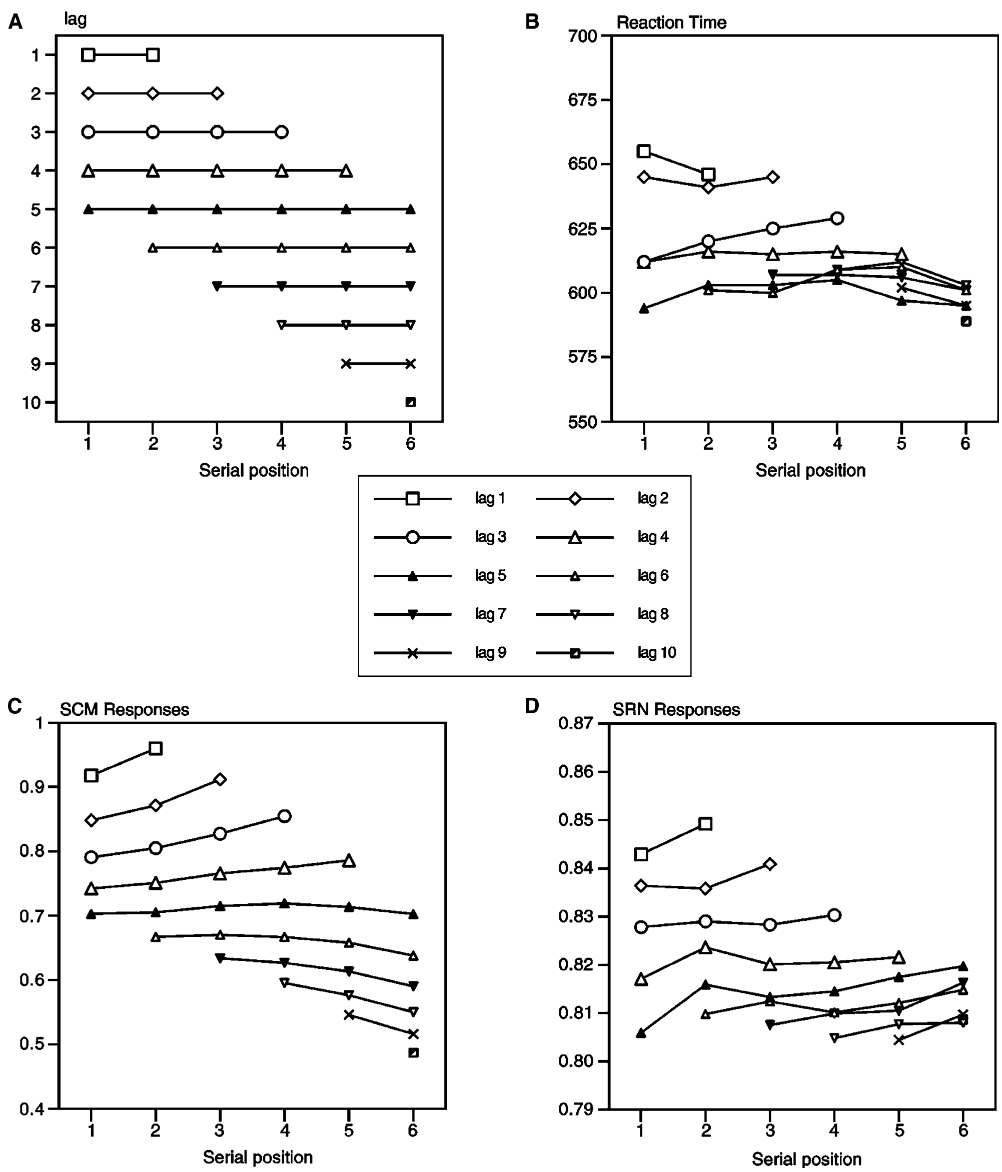

Fig. 2 **a** Distribution of lags of lengths 1–10 between successive occurrences of the same stimulus plotted as a function of serial position (Experiment 1). **b** Mean reaction times plotted separately for stimuli associated with lags 1–10, as a function of serial position. **c** Mean SRN responses plotted separately for stimuli associated with lags 1–10, as a function of serial position. **d** Mean simple condensator model (*SCM*) responses plotted separately for stimuli associated with lags 1–10, as a function of serial position

correct boundaries. This, however, need not be the case: Participants in fact merely need to be sensitive to the lag that separates two occurrences of the same stimulus, and to produce faster responses to stimuli associated with a long lag.

To see this, consider the fact that any stimulus that occurs in serial position 6, that is, as the final element of an experimenter-generated sequence, is necessarily associated with a lag of at least length 5, in that, by construction, the same stimulus could not have occurred within the same experimenter-generated sequence. In contrast, stimuli that occur in serial position 1 could have previously occurred as recently as two trials ago (in the previous sequence), and thus be associated with a lag of length 1. This state of affairs is depicted in Fig. 2a, which clearly shows that the different serial positions are associated with ranges of lags of increasing length. For instance, position 1 is associated with lags of lengths 1–5, and position 6 with lags of lengths 5–10. From this perspective, then, the position effect described by Lee (1997) and replicated in this experiment, merely emerges out of more elementary features of the material, namely that in each trial, the probability of any stimulus increases linearly with the lag that separates the current trial from the stimulus's previous occurrence, and that different serial positions in the experimenter-generated sequences are, by construction, associated with distributions of increasing lags.

If our account is correct, then we should observe a lag effect in the data. Figure 2b represents the data plotted separately for each lag, and shows that position, by itself, seems to have little impact on performance. Indeed, each curve (corresponding to stimuli with a given lag length as in Fig. 2a) is relatively flat across serial positions. An ANOVA with Position (six levels) applied to these data, and restricted to stimuli with a lag of length 5 (the only case where position and lag are completely crossed) confirmed this impression and showed no significant effect of position ($p = .09$). To further assess the relative contribution of Position and Lag to performance, we conducted a series a regression analyses on the set of mean RTs obtained from crossing both factors (see Fig. 3a). First, simple linear regressions with either Position or Lag as predictors indicated that both factors influence performance (Position: $p < .01$; $r^2 = .30$; Lag: $p < .01$; $r^2 = .64$). Next, a stepwise analysis with Position entered at step 0 and Lag at step 1 showed that adding Lag as a predictor significantly increased r^2 (r^2 Change = .34, $p < .01$). When the order of predictors was reversed, however, r^2 did not change significantly (r^2 Change = .00001, $p > .05$), thus confirming that lag, rather than position, accounts for the distribution of RTs in this experiment.

Hence, it should be clear that participants do not need to, and in fact do not, parse the material in order to exhibit the observed serial position effect. It would appear that sensitivity to the serial position, far from indicating learning of an abstract rule, may in fact reflect elementary knowledge that participants already possess before being exposed to the task. This knowledge may consist of a tendency to prepare for responses that have not been used recently, in a way similar to the biases involved in the well-known fact that spontaneously generated random sequences are in fact much more uniform (i.e., involve more alternations) than true random distributions (see, e.g., Lopes, 1982; Rapoport & Budescu, 1997; Wagenaar, 1972). How might this knowledge be established? This is the issue on which we focus in the rest of this paper. We first present two models that are both equally capable of simulating the empirical data, but that differ considerably in their assumptions about how sensitivity to repetition distance is achieved. Next, in Experiment 2, we present new empirical and simulation results aimed specifically at making it possible to contrast predictions based on each model.

Simulation

A simple "condensator" model

A first hypothesis is that sensitivity to the lag that separates successive occurrences of the same event reflects fundamental properties of the motor system or fundamental, hard-wired properties of the attentional system. For instance, Dominey's (1998) model of sequence learning explicitly includes mechanisms that are directly sensitive to repetition distance. To find out whether such direct sensitivity was sufficient to account for the data, we constructed a simple model (which we dubbed the "Simple Condensator Model"; SCM) consisting of six units, each corresponding to a given response. Each unit is associated with an activation level ranging from .0 to 1.0. Units are initialized to activation values of .166 (1.0 divided by 6) to reflect the fact that there are six response alternatives. When a stimulus is presented, the corresponding response unit fires and its activation upon firing is distributed equally among the other five units. The summed activation of all units is therefore always equal to 1.0. The RT is assumed to be inversely proportional to the activation of the firing unit. During processing, the activation of each unit will thus increase with time up until the point when the corresponding stimulus is presented. The dynamics of this simple accumulate-and-fire model thus depend directly on repetition distance, because at any point in time, the most active unit always corresponds to the stimulus associated with the longest lag. Note that this model never learns anything about the stimulus material or its sequential structure.

To find out how well this simple model fared in accounting for the data, 12 models were each exposed to the same stimulus series as experienced by each of the 12 human participants, and their responses recorded and then averaged together. Incorrect human responses were eliminated from the corresponding model's data, and the models' responses were subtracted from 1.0 prior to averaging so as to make increases in activation compatible with decreases in RT. In the following, we will simply call these transformed responses "SCM responses."

We assessed the model's performance by conducting the same analyses as for human participants. First, Fig. 2c shows that the model's responses correspond almost perfectly with the actual distribution of lags (Fig. 2a), and very closely with the distribution of RTs (Fig. 2b). Second, regressions using Position and Lag as predictors indicated significant effects in both cases (respectively $p < .01$; $r^2 = .43$, and $p < .01$; $r^2 = .97$). Figure 3b makes it clear that SCM responses decrease linearly from the first to the tenth lag. Third, stepwise analyses with Position (step 0) and Lag (step 1) indicated that r^2 is significantly increased by the addition of Lag (r^2 Change $= .54$, $p < .01$), whereas entering Position on step 1 in the converse analysis failed to increase r^2 (r^2 Change $= .0005$, $p > .05$). Finally, the high correlation between human and simulated performance ($r = .84$; $p < .05$) confirmed that the model was successful in accounting for the distribution of RTs, and hence that sensitivity to the lag provides a very good account of the serial position effect.

However, while this simple model provides a good descriptive account of the data, it also takes it as a starting point that participants are directly sensitive to repetition distance. An important issue, however, is whether we should regard this sensitivity as a hard-wired bias, or whether this bias should itself better be taken as the result of prior learning. In the following, we suggest that sensitivity to repetition distance is a bias that emerges as the result of life-long exposure to sequences of environmental stimuli structured in a way that is consistent with the material used in this experiment. Events do not occur randomly. On the contrary, they tend to repeat at different predictable intervals. For instance, the probability of your having a meal is close to zero if you have just had one, and subsequently increases monotonically with the time elapsed since your last meal. The same situation is true for myriads of real-world events, from daily, personal events to the movements of planetary bodies. Any

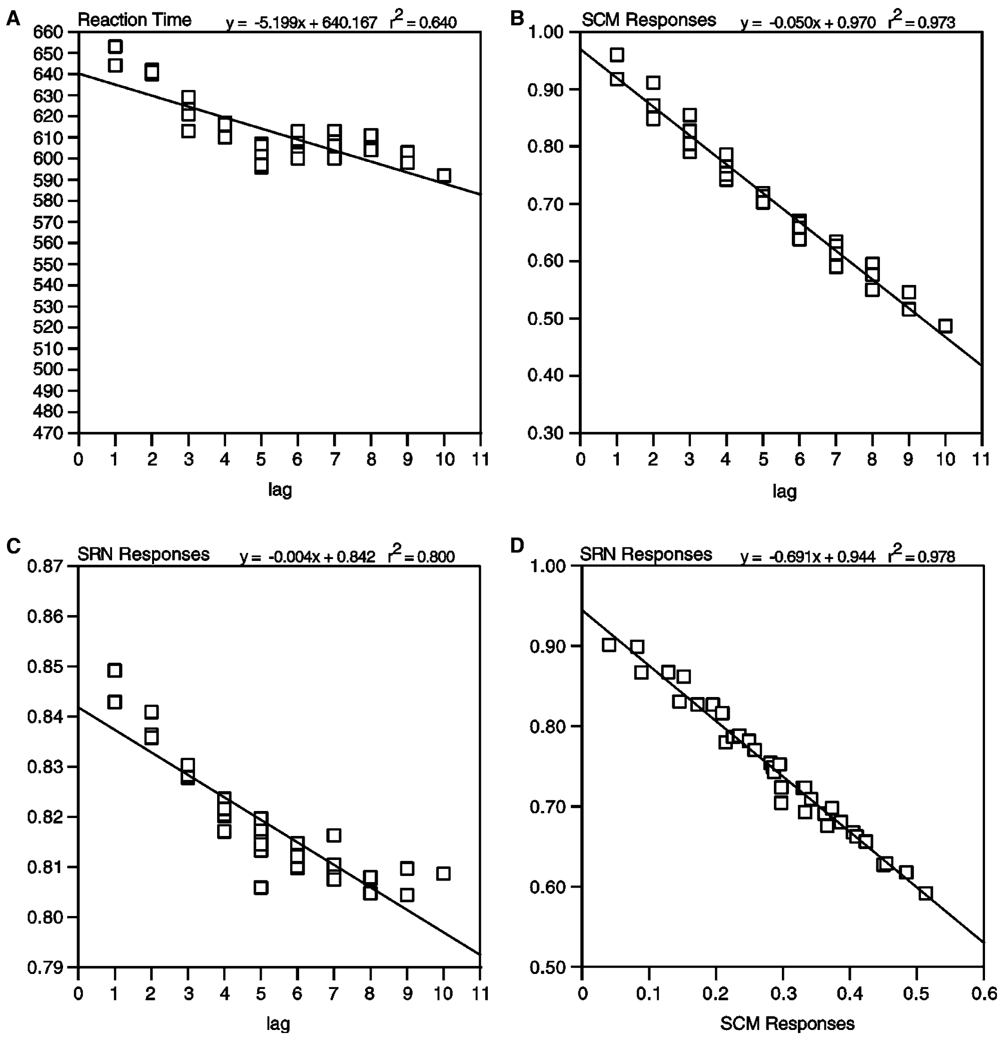

Fig. 3 Simple linear regressions for the Lag effect: Comparisons between human and simulated performance (Experiment 1). **a** Mean reaction times. **b** SCM responses. **c** SRN responses. **d** Association between SRN and SCM responses. The network has learned to emulate the SCM almost perfectly

system that attempts to predict future events based on past experience is likely to become sensitive to these temporal regularities, precisely because repetition distance has predictive value in these contexts. Our suggestion is thus that sensitivity to repetition distance is a learned bias with which participants enter our experiment, and that this bias itself emerges out of even more elementary prediction-based associative learning mechanisms. To test these ideas, we explored whether the data could also be explained by adaptive models such as the SRN, which constitutes one of the best models of sequence learning performance.

The simple recurrent network model

The SRN (see Fig. 4) uses back-propagation to learn to *predict* the next element of a sequence based only on the current element and on a representation of the temporal context that the network has elaborated itself. To do so, it uses information provided by so-called context units which, at every step, contain a copy of the network's hidden unit activation vector at the previous time step. Over training, the relative activation of the output units representing each possible successor come to approximate the optimal conditional probabilities associated with their appearance in the current context, and can thus be interpreted as representing implicit preparation for the next element when the network is used as a model of human sequence learning performance. Previous work (Cleeremans, 1993; Cleeremans & McClelland, 1991) has shown that the SRN accounts for about 80% of the variance in similar tasks.

Simulation parameters and procedure

To assess how well the SRN could capture RT performance in this experiment, we trained the model on the same material as human participants. The network consisted of 80 hidden units and local representations of both the input and output pools (i.e., each unit corresponded to one of the six stimuli)[1]. The network was trained to predict each element of a continuous sequence of stimuli generated exactly as for human participants. On each step, a stimulus was presented to the network by setting the activation of the corresponding input unit to 1.0. Activation was then allowed to spread to the other units of the network, and the error between its response and the actual successor of the current stimulus was used to modify the connection weights, using standard back-propagation. During training, the activation of each output unit was recorded in every trial and normalized according to Luce's choice rule (Luce, 1963). For the purpose of comparing simulated and observed responses, we assumed that the normalized activations (i.e., strength) of the output units represent response tendencies, and that there is a linear reduction in RT proportional to the strength of the unit corresponding to the correct response. The network's responses were finally subtracted from 1.0 to make increases in response strength compatible with reduction in RT. In the following, we will simply call these transformed responses "SRN responses."

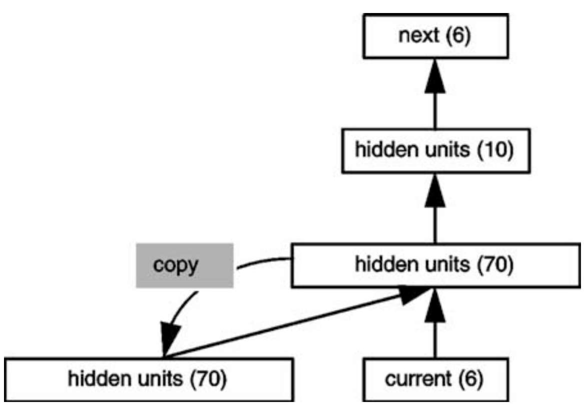

Fig. 4 The simple recurrent network

The lag effect is emergent

We first conducted exploration of the parameter space using the network illustrated in Fig. 4. We found that the network was able to master the material perfectly, but only after extensive training consisting of seven presentations of the entire training set. Figure 5 provides a more detailed view of the network's performance as it changes over training, and shows that the network becomes progressively able to predict which elements are possible at each serial position (bottom row). For instance, after seven epochs of training, the network perfectly predicts that "6" is the only possible successor of "12345."

As described in Servan-Schreiber, Cleeremans, and McClelland (1991), the development of sequence knowledge in the SRN involves gradually increasing sensitivity to the sequential constraints contained in an increasingly large and self-developed representation of the temporal context defined by previous elements of the sequence. Initially, the network learns to associate each element with the distribution of its possible successors, and essentially ignores the context information. In this material, each element is associated with a unique distribution of successors because, by construction, an element cannot be followed by itself. Hence, after one epoch of training (see Fig. 5, top row), the network

[1]All networks involved dual connection weights as described in Cleeremans and McClelland (1991). Slow and fast weights were associated with learning rates of .04 and .45 respectively. Momentum was .9, and the fast weights decayed at a rate of .4. For each simulation study described in the text, twelve networks initialized with different random weights selected in the $-.5$–$+.5$ range were each trained in a total of 30,240 trials (720 sequences $\times$ 6 elements $\times$ 7 epochs), and their responses averaged

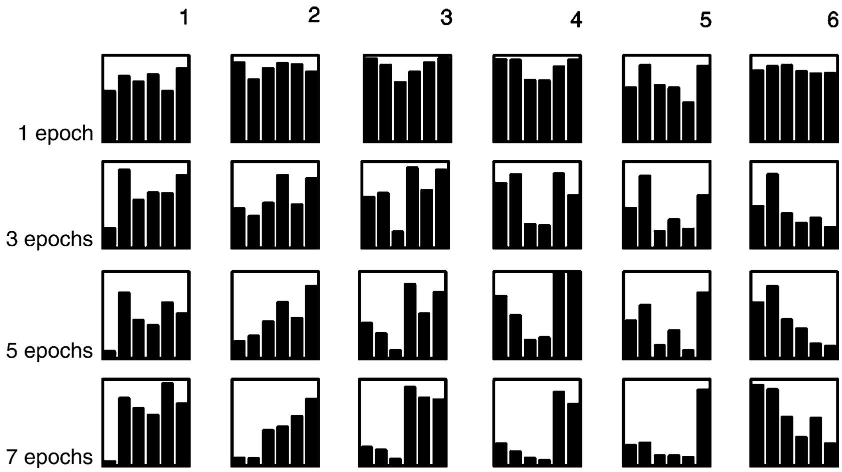

Fig. 5 SRN prediction responses upon presentation of each element of the sequence "123456" (*columns*) at different points in training (*rows*). Each *box* in the figure represents the activation of each of the network's six output units upon presentation of one of the six possible stimuli. The *figure* shows that after only one epoch of training (*top row*), the network fails to make discriminative predictions (all of the six possible successors tend to be equally activated). In contrast, after seven epochs of training (*bottom row*), it can clearly be seen that the network has learned to strongly inhibit those responses that would correspond to repetitions of a recent response. For instance, when presented with stimulus "1," the network predicts that all stimuli except stimulus "1" are likely to occur in the next trial. When presented with "2" immediately thereafter, it predicts that neither "1" nor "2" are likely to occur next, and so on. This sensitivity to the material's repetition structure develops gradually (epochs 3–5) over training

tends to predict that all the elements except the stimulus itself are possible successors. The patterns of activation over its hidden units now represent these associations. When fed back onto the context units, these patterns can now be used by the network as representations of the previous element, and it can then start basing its predictions on two elements. It is when the network uses three previous elements that the material's structure starts conveying information about the lag that separates occurrences of the same stimulus.

Consider for instance the fragment "123." It can never be followed by 3 by construction, but it is also more often associated with "1" as a successor than it is with "2," regardless of the serial position at which it ends. This is simply because there are more ways for "1231" to occur in the stimulus set than there are ways for "1232" to occur. Indeed, whereas neither "1231" nor "1232" can occur within any legal sequence, "1231" can span two legal sequences in three different ways ("1-231," "12-31," and "123-1"), whereas "1232" can only do so in two different ways ("12-32," and "123-2"). Hence, the lag effect emerges out of the network's prediction-based sensitivity to the statistical structure of the material. Furthermore, this sensitivity to the lag is itself the basis for the emerging serial position effect characteristic of human performance. As Fig. 3d shows, the network's responses in fact come to correspond almost exactly ($r^2 = .978$, $p < .001$) to the responses produced by the SCM described above. Over training, the network has thus learned to emulate the SCM perfectly.

Simulating the human data

The simulation work described above—when considered together with the fact that human participants do not appear to learn much beyond unspecific practice effects—suggests that sensitivity to repetition distance is a learned bias that participants enter the experiment with. To find out how well the SRN was able to account for human performance, we therefore assumed that the network had previously experienced stimulus material containing predictive lag structure, and assessed how well its responses captured human performance when exposed to Lee's material.

To do so, we pre-trained networks for six epochs on a degraded version of Lee's material. The degradation was achieved by adding normally distributed random noise (m = .0; s = .5) to the *net input* of each unit in the network. The networks were then exposed to the material of Experiment 1, and their performance assessed as described above.

Simulation results

When exposed to the stimulus material after pre-training as described above, the network is able to master the training set almost perfectly, as illustrated in Fig. 1c, and exhibits a comparable linear serial position effect (Fig. 1d) confirmed by an effect of Position $F(5, 55) = 4.19$, $p < .01$, MSE = .0029 obtained from the ANOVA with Block (24 levels) and Serial Position (six levels) as repeated measures factors. Furthermore, Fig. 2c shows that the network's distribution of responses is remarkably similar to—and indeed almost identical with—the actual distribution of lags over the six serial positions within the stimulus material (compare Fig. 2a, c). Regression analyses showed that Position, in itself, only accounts for about 24% of the variance in the network's responses. ($p < .01$; $r^2 = .24$). Lag, however, accounts for about 80% of the variance ($p < .01$; $r^2 = .80$), as illustrated in Fig. 3c. When Position is entered at step 0 into the equation of a stepwise analysis, r^2 increases significantly with the addition of Lag at step 1 (r^2 Changed = .59, $p < .01$). Conversely, when Lag is entered first, r^2 only increases very moderately (r^2 Changed = .028, $p < .05$), thereby again suggesting that Position alone makes almost no contribution to overall variance. Finally, the correlation between human and SRN responses was .88 ($p < .01$).

Discussion

Our simulation work indicated that both the SRN and the SCM are quite capable of accounting for the distribution of human responses. Both models suggest that sensitivity to repetition distance is the key factor influencing human responses. However, the two models make very different assumptions about the origin of sensitivity to repetition distance. The SCM assumes that this sensitivity is hard-wired and not subject to modification through experience, whereas the SRN assumes that repetition distance is a learned, adaptive bias. To contrast the two models and to determine whether sensitivity to repetition distance depends on predictive value or not, we conducted a second experiment in which both human participants and models were exposed to random material that fails to contain prediction-relevant lag structure. With such sequential material, stimuli cannot be predicted or anticipated through lag structure. So, exposed to such material, participants would not be sensitive to repetition distance and their results would not indicate any lag effect. As for models, the SCM would predict continued sensitivity to repetition distance, whereas prediction-based learning models such as the SRN would predict the extinction of lag effects.

Experiment 2

Experiment 2 followed exactly the same design as Experiment 1, but involved pseudo-random stimulus material. Twelve new participants aged 18–26 were paid as in Experiment 1 to participate. The sequences of stimuli (720 sequences of six elements) were chosen randomly with the only constraints that runs of more than two identical elements were forbidden and that the maximum lag between two occurrences of the same stimulus could be no larger than 20[2]. Unlike Experiment 1's material, the stimulus sequence used here fails to contain any association between lag and serial position because the stimuli occurring in each serial position could now be associated with any lag. We chose to increase the maximum lag between successive occurrences of the same stimulus from 10 to 20 in Experiment 2 so as to test the idea that mere sensitivity to the lag accounts for the distribution of reaction times. Indeed, the SCM would predict even better performance on stimuli associated with lag 20 than on stimuli associated with lag 10. On the other hand, since the material of Experiment 2 is such that it fails to contain any prediction-relevant lag structure, the SRN model is not expected to exhibit any sensitivity to lag structure. Each participant received the stimuli in a different random order.

[2]An error in the stimulus generation program resulted in one instance of a stimulus associated with a lag of 21.

Results and discussion

The results presented below were analyzed by simple linear regressions based on the means obtained by crossing Serial Position (six levels) and Lag (21 levels). Errors and outliers were again eliminated and amounted to 10.9% of the human data.

Figure 6a shows the human data. Each data point represents the mean RT obtained for a specific combination of Lag and Position. The figure shows that unlike Experiment 1, participants do not appear to be sensitive to repetition distance when exposed to material for which lag structure fails to be predictive. These impressions were confirmed by a two-way ANOVA with Block (24 levels) and Serial Position (six levels) as repeated measures, which failed to reveal a significant effect of position, $F(5, 55) = 1.75$, $p > .05$. Regression analyses applied to the set of mean RTs obtained from crossing the factors Lag (20 levels) and Serial Position (six levels) likewise failed to reveal significant effects of either Lag or Position (respectively $p > .05$, $r^2 = .006$; $p > .05$, $r^2 = .003$).

Simulations

To find out about the extent to which either the SCM or the SRN model may account for our results, we again conducted simulations of our empirical data. Twelve instances of both models were again exposed to the same material, as described above. The SRN was pre-trained for six epochs on a noisy version of the stimulus material of Experiment 1, and subsequently exposed to the random material.

SCM results

Figure 6b shows the results obtained for the SCM. A two-way ANOVA with Block (24 levels) and Serial Position (six levels), as repeated measures was first applied to SMC responses. Unlike human participants, we observed a significant effect of serial position, $F(5, 55) = 20.98$, $p < .0001$, suggesting that the model continues to exhibit sensitivity to repetition distance: Indeed, SCM responses decreased linearly from the first to the sixth serial positions. Figure 6b shows the results of regression analyses obtained for the SCM. A simple linear regression conducted on these data with Lag as predictor confirms this impression ($p < .01$, $r^2 = .96$). In contrast, the same analysis with Position as predictor failed to reach significance ($p > .05$, $r^2 = .0004$). This was expected since the random material was constructed to break down the association between lags and serial position as SCM is built-in sensitivity of repetition distance whatever lag structure. These results therefore indicate that the SCM cannot be taken as an adequate model of human performance, for its continued sensitivity to repetition distance is inconsistent with the human data.

SRN results

Figure 6c shows SRN performance. A simple regression analysis conducted on these data showed that the network, like human participants, fails to be sensitive to Position ($p > .05$, $r^2 = .004$). A similar analysis using Lag as predictor was unexpectedly significant ($p < .01$, $r^2 = .09$), but it should be pointed out that the percentage of variance explained is very small (9%), unlike that observed for the SCM (96%). This weak sensitivity to repetition distance can be attributed to the residual effects of the pre-training to which the SRN was exposed. This idea is confirmed by the results of separate simple linear regressions conducted either on the data corresponding to lags 1–10 or on the data corresponding to lags 11–21. While the first such analysis showed a significant effect of lag ($p < .01$, $r^2 = .74$), the second indicated no effect ($p > .05$, $r^2 = .0006$), thus confirming that the very weak continued sensitivity to repetition distance exhibited by the model is a result of its pre-training. While it cannot be excluded that a better choice of parameters would further increase the contrast between the two models, we did conduct extensive "manual" exploration of the parameter space and almost systematically always found the same pattern of results.

The results of this second experiment thus confirmed our expectations: When repetition distance no longer has predictive value, both human participants and the SRN fail to exhibit continued sensitivity to it, in contrast to the inflexible SCM.

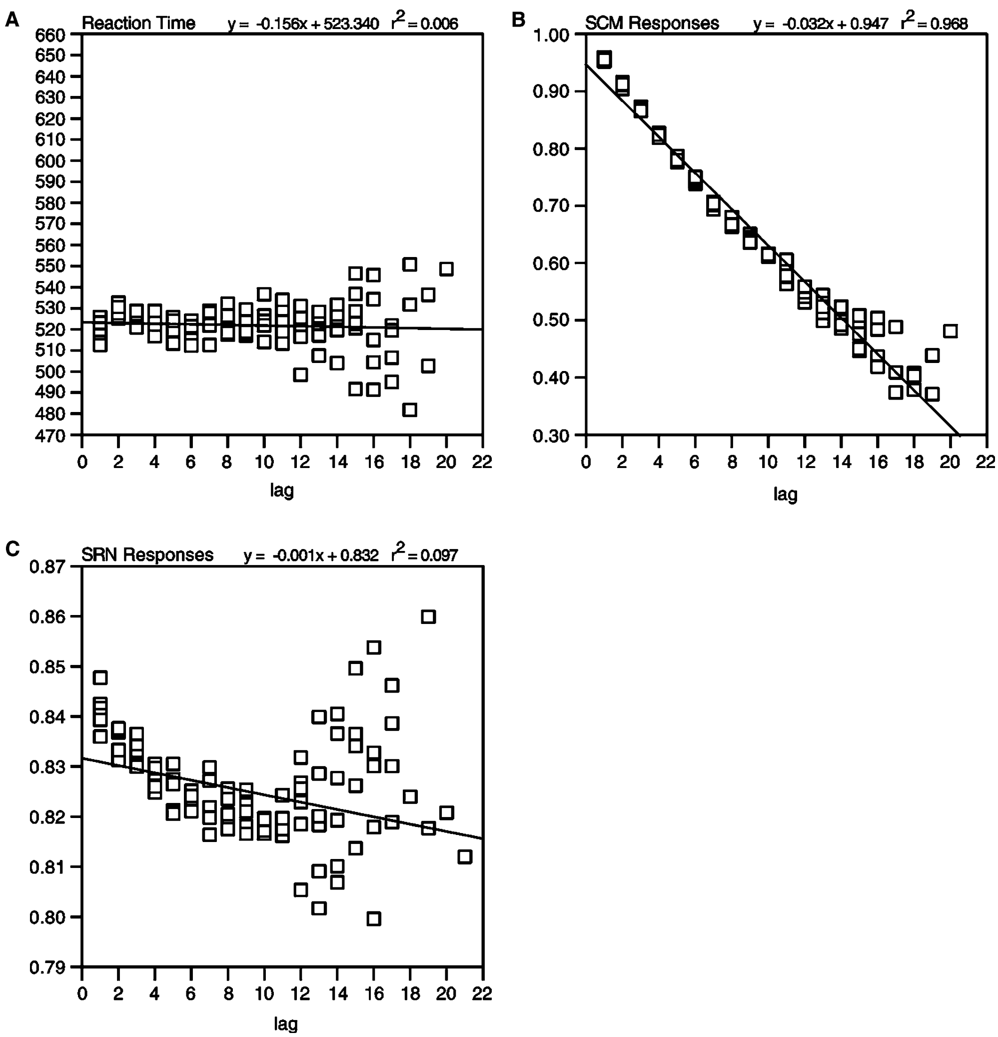

Fig. 6 Simple linear regressions for the Lag effect: Comparisons between human and simulated performance (Experiment 2). **a** Mean reaction times. **b** SCM responses. **c** SRN responses

General discussion

In this paper, we first suggested that the core mechanism involved in sequence learning is statistical in nature, and rooted in the development of distributed representations of the temporal context acquired through elementary associative learning processes that operate on exemplars. We showed how such mechanisms are in fact sufficient to understand how sensitivity to very abstract features of the material, such as the serial position effect described by Lee (1997), can emerge out of a sensitivity to more elementary features of the material, such as the lag that separates successive occurrences of the same stimulus. Neither memory for previous specific elements nor parsing mechanisms of any kind are in fact necessary to understand human performance in this situation. More surprisingly, perhaps, our results suggest that performance in this experiment does not involve actual learning of the sequential regularities, but merely reflects continued reinforcement of knowledge that participants already possess before being exposed to the task. This is not to say that no learning whatsoever occurs in this situation. Indeed, we have suggested that this pre-experimental knowledge—which amounts to a bias to preferentially prepare responses that have not been produced recently—is itself the result of learning, achieved through the operation of continuously operating prediction mechanisms exposed to real-world contingencies in which repetition distance has predictive value. Furthermore, Experiment 2—in which the material is such that the lag structure fails to be predictive—shows that only an adaptive model (the SRN) is able to account for human performance. This is turn suggests that sensitivity to repetition distance depends on continued reinforcement, as in Experiment 1. Hence the SRN model does learn over the course of Experiment 1, but this learning appears limited to the reinforcement of existing knowledge. In other words (and this will be our main conclusion), it appears that performance in this task might be more a matter of knowing without learning than a matter of learning without knowing.

Second, our findings suggest that phenomena such as negative recency—the tendency to respond more efficiently to stimuli that have not occurred recently (Baddeley, 1966; Budescu, 1987; Jarvik, 1951; Lopes, 1982; Rapoport & Budescu, 1997; Wagenaar, 1972), inhibition of return—which in its most basic form results in a slowing of responses to recently attended visual location (Klein, 2000; Lupiañez, Milliken, Solano, Weaver, & Tipper, 2001; Posner & Cohen, 1984; Posner, Rafal, Choate, & Vaughan, 1985; Taylor & Klein, 1998), and the "gambler's fallacy"—the erroneous belief that alternations should occur more frequently than repetitions in random sequences of events (Anderson, 1960; Edwards, 1961; Feldman, 1959; Keren & Lewis, 1994) may all find their roots in real-world experience with sequential structure characterized by the fact that repetition distance has predictive value.

Sensitivity to repetition distance (that is, negative recency) is prima facie inconsistent with Anderson and Schooler's (1991) characterization of memory as a system in which retrieval probability depends only on how recently and on how frequently the corresponding traces were needed. Negative recency findings instead suggest that memory might also be optimized to facilitate responding to events that have not occurred recently. Nevertheless, we would like to suggest that memory, in general, might be optimized to facilitate both the retrieval of recent and earlier traces, perhaps based on differences in the time scale over which these traces occur. Our main suggestion is that the key to reconciling these two seemingly incompatible computational objectives is to consider the conditions in which positive and negative recency each have predictive value. In other words, memory continuously serves attention by directing it towards the items that are most likely to occur next. In most environments, these items will be those that have been most recently experienced. In other environments, however, these items will be ones that have not occurred recently. In both cases, adaptive responding is made possible by automatic encoding of the predictive value of the lag that separates recurring events. This illustrates another important principle of information processing brought forward by connectionist modeling, and which we discussed briefly in the Introduction: The "principle of mandatory plasticity," that is, the notion that learning occurs whenever information processing takes place, whether subjects are aware of learning or not.

We also suggested that negative recency effects are rooted in the sort of elementary-based mechanisms that characterize adaptive sequence learning models such as the SRN. Indeed, our simulation work suggests that the SRN learns, through exposure to environments in which repetition distance has predictive value, to behave just as though it were directly sensitive to repetition distance, as with the SCM model. In other words, the SRN learns to emulate the SCM over training. An important implication of this finding is that negative recency is likely to play a role in any sequence learning situation. If this is indeed the case, it would be important to assess its effects on performance, particularly when such effects can be confounded with other aspects of the material's structure.

The simulation work we described is illustrative of the power of elementary associative learning mechanisms, and in this sense makes it clear how behavior can be "rule-like" without necessarily being "rule-based": An agent's decisions can thus be influenced by the regularities shared by exemplars of a domain in such a manner that these regularities are never tokened as "knowledge" by the agent, but instead remain implicit (distributed) over the representations associated with each exemplar. These regularities might further depend on the functional similarity between exemplars rather than on their physical similarity. In other words, the regularities may reflect abstract similarity relationships between exemplars (McClelland & Plaut, 1999)—how different exemplars can be processed together because they are associated with the same class of responses—rather than on their physical similarity—the number of features their input representations have in common. In this case, these abstract relationships consist of the association between the serial position at which stimuli occur and the lag that separates each occurrence of a particular stimulus from its previous occurrence. While these associations can easily be described by a simple rule, the current experiments provide no evidence that participants (or networks, for that matter) induce this rule. Yet, they behave as though they had.

The rules versus similarity issue continues to be an object of lively debate in various literature, ranging from language acquisition (Marcus, Vijayan, Bandi Rao, & Vishton, 1999) to categorization (Pothos, in press). It is also an issue that has proven particularly challenging to settle convincingly using laboratory settings, for it can always be argued, for instance, that people would end up developing rule-based knowledge if given enough time to fully induce the corresponding rules. In this respect, recent findings described by Pacton et al. (2001) are illuminating. Pacton et al. took it as a starting point that rules are typically defined as being absolute: When a rule applies, it does so regardless of the surface features with which the stimulus is instantiated. Anderson (1993), for instance, states that "abstraction refers to the generality of production rules. Production rules do not require that a specific stimulus be present; the rules will apply in any stimulus condition that satisfies the pattern specification of the condition" (p. 35, see also Manza & Reber, 1997; Smith, Langston, & Nisbett, 1992). This definition therefore predicts that, if a rule is at play, we should not observe any loss of performance in a transfer situation in which the surface features of the material are changed between transfer and test. Pacton et al. (2001) applied this logic to a natural situation—the learning of orthographic regularities by children between grades 1 and 5. Over a series of empirical studies, Pacton et al. explored children's sensitivity to untaught—and hence presumably implicit—regularities that can easily be described as rules, such as the fact that vowels are never doubled in French, or the fact that double consonants never occur at the beginning or endings of words. Using carefully controlled preference tasks, Pacton et al. were able to show not only that sensitivity to such rules tended to develop between grades 1 and 5, but also, and more importantly, that this sensitivity transferred to instances that never occur in French (i.e., some consonants are never doubled in French).

While such transfer performance would be predicted by rule-based approaches, it would also be expected, in virtue of the fact that rules are defined to be absolute, that this transfer should be perfect after several years of exposure. The results, however, failed to confirm this prediction: What was observed instead was that performance on familiar and novel material exhibited a "transfer decrement" that remained stable across grade levels. This persistence of transfer decrement, even after several years of exposure to relevant material, clearly suggests that rule abstraction does not occur in this natural context. Just as importantly, simulations based on the Simple Recurrent Network confirmed all aspects of these data. Based on such findings, we would tend to concur with Pothos (in press) that the rules vs. similarity distinction is best viewed as a graded continuum—a perspective that is fully consistent with the conceptual framework introduced by Cleeremans and Jiménez (2002).

In closing, it is interesting to speculate on the reasons why participants remain unaware of the single abstract rule through which the structure of the material can be easily described. As we indicated in the Discussion of Experiment 1, not a single participant reported being aware of any structure in the stimulus material, despite several hours of experience with the task. While skeptics will be quick to point out the limitations of verbal report as a measure of conscious awareness, the result remains striking, particularly in light of the fact that the structure of the material is so simple. We take this (admittedly scant) evidence as providing further support for the notion that performance in this task is not rule-based, for we think that genuine rule-based knowledge is necessarily conscious (Cleeremans & Destrebecqz, in press).

Interestingly, Perruchet (1985) reported on a similar dissociation in the context of an eye-blink conditioning situation that made it possible to obtain comparable quantitative measures of priming and awareness. In this experiment, people were exposed to a series of identical tones, 50% of which could be followed after a short interval by an air puff directed to the left cornea. Immediately after each tone was presented (and before the puff occurred in reinforced trials), people were asked to indicate (using a 0–7 points scale) the extent to which they expected the tone to be reinforced. A trial-by-trial analysis of the results indicated that eye blink responses were increasingly more likely to occur after presentation of a tone if the corresponding trial had been preceded by a series of reinforced trials (i.e., trials during which the tone had indeed been followed by an air puff). In stark contrast, however, people's subjective expectancy of the occurrence of an air puff tended to decrease with the number of reinforced trials that preceded the trial under consideration. In other words, people's eye blink responses were completely dissociated from their conscious expectations about when each tone would be followed by an air puff.

The two situations—Perruchet's eye-blink experiment and Lee's scenario—differ in interesting ways. In Perruchet's scenario, people follow the gambler's fallacy through their conscious expectancy judgments (without realizing that they do so), yet their behavior fails to be congruent and appears instead to be exclusively sensitive to automatic priming. In the experiments described in this paper, by contrast, it is people's behavior that reflects the gambler's fallacy, while their conscious apprehension of the material simply appears to be extremely limited. Interestingly, Perruchet, Cleeremans, and Destrebecqz (in preparation) recently replicated Perruchet (1985)'s original eye-blink conditioning results in a simple RT paradigm, that is, in a situation that involves voluntary responses.

These different results all point toward strong dissociations between behavior and conscious experience, and suggest that mere "quality-of-representation" accounts of the extent to which a representation is available to verbal report (such as Cleeremans and Jiménez's framework, briefly described in the Introduction) are not sufficient. What additionally appears to be crucial in making some representation available to conscious experience—above and beyond its overall quality (strength, stability, distinctiveness)—is the extent to which the representation is accompanied by relevant meta-knowledge, that is, the extent to which the representation can be redescribed using, for instance, a verbal code.

This redescription is made particularly difficult in our case because the relevant knowledge (i.e., Lee's rule) is not represented per se, but is instead highly distributed across the representations of the numerous exemplars or features thereof to which subjects were exposed. In other words, even though the representations associated with the relevant knowledge may be stable in time and relatively strong, they are not nevertheless not distinctive enough to be good candidates for such redescription, and hence, for availability to conscious awareness.

Exploring the conditions under which a piece of knowledge becomes or fails to become available to conscious awareness over the course of experience with a particular set of stimuli undoubtedly constitutes one of the most promising avenues of research in this domain (see e.g., Wagner, Gais, Haider, Verleger, & Born, 2004, for a fascinating account of the role of sleep in insight).

Acknowledgements Maud Boyer is a post-doctoral researcher supported by Grant RPG-53 from the International Human Frontiers of Science Program. Arnaud Destrebecqz is a post-doctoral researcher supported by a grant from the Fyssen Foundation. Axel Cleeremans is a Senior Research Associate of the National Fund for Scientific Research (Belgium). This work was supported by FRFC Grant#2.4605.95 F, by a grant from the European Commission (HPRN-CT-1999-000065), and by a grant from the Fyssen Foundation to Maud Boyer. We thank Pierre Perruchet, Robert French, Tim Curran, David Shanks, and an anonymous referee for insightful comments on previous versions of this article.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 296–408.

Anderson, N. H. (1960). Effect of first-order conditional probability in a two-choice situation.*Journal of Experimental Psychology*, *56*, 73–93.

Baddeley, A. D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, *18*, 119–129.

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

Boyer, M., Destrebecqz, A., & Cleeremans, A. (1998). The serial reaction time task: Learning without knowing, or knowing without learning? In *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society* (pp. 167–172). Mahwah, NJ: Erlbaum.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), Cognition and concepts (pp. 16–211). Mahwah, NJ: Erlbaum.

Budescu, D. V. (1987). A Markov model for generation of random binary sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 25–39.

Cleeremans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing*. Cambridge, MA: MIT Press.

Cleeremans, A. (1997). Principles for implicit learning. In D. C. Berry (Ed.), *How implicit is implicit learning?* (pp. 195–234). Oxford: Oxford University Press.

Cleeremans, A. (forthcoming). *Being virtual*. Oxford: Oxford University Press.

Cleeremans, A., & Destrebecqz, A. (in press). Real rules are conscious: Open peer commentary to Pothos. *Behavioral and Brain Sciences*.

Cleeremans, A., & Jiménez, L. (1998). Implicit sequence learning: The truth is in the details. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 323–364). Thousand Oaks: Sage.

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, computational and philosophical consensus in the making?* (pp. 1–40). Hove, UK: Psychology Press.

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

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

Clegg, B. A., DiGirolamo, G. J., & Keele, S. W. (1998). Sequence learning. *Trends in Cognitive Sciences*, *2*, 275–281.

Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little, Brown

Dominey, P. F. (1998). Influences of temporal organization on sequence learning and transfer: Comments on Stadler (1995) and Curran and Keele (1993). *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 234–248.

Dominey, P. F., Lelekov, T., Ventre-Dominey, J., & Jeannerod, M. (1998). Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. *Journal of Cognitive Neuroscience*, *10*, 734–751.

Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology*, *62*, 597–599.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Feldman, J. (1959). On the negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, *41*, 291–297.

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

Hinton, G. E. (1986). Learning distributed representations of concepts. In *Proceedings of the 8th Annual Conference of the Cognitive Science Society* (pp. 1–12). Hillsdale, NJ: Erlbaum.

Hintzmann, D. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*,*93*, 411–428.

Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology*, *41*, 291–297.

Jiménez, L., Mendez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*, 948–969.

Keren, G., & Lewis, C. (1994). The two fallacies of gamblers: Type I and Type II. *Organizational Behavior and Human Decision Processes*, *60*, 75–89.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*, 138–146.

Koch, I., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychological Research*, *63*, 22–35.

Lee, Y. S. (1997). Learning and awareness in the serial reaction time task. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, *20*, 24–37.

Lopes, L. L. (1982). Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *8*, 626–636.

Luce, R. D. (1963). Detection and recognition. In R. R. Busch & E. Galanter (Eds), Handbook of mathematical psychology: Vol. 1. New York: Wiley.

Lupiañez, J., Milliken, B., Solano, C., Weaver, B., & Tipper, S. P. (2001). On the strategic modulation of the time course of facilitation and inhibition of return. *Quarterly Journal of Experimental Psychology*, *54A*, 753–773.

Manza, L., & Reber, A. S. (1997). Representing artificial grammar: Transfer across stimulus forms and modalities. In D. Berry (Ed.), *How implicit is implicit learning?* Oxford: Oxford University Press.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.

McClelland, J. L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules?*Trends in Cognitive Sciences*, *3*, 166–168.

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

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

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, *130*, 401–426.

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

Perruchet, P., & Gallego, G. (1997). A subjective unit formation account of implicit learning. In D. Berry (Ed.), *How implicit is implicit knowledge?* (pp. 124–161). Oxford: Oxford University Press.

Perruchet, P., & Vinter, A. (2003). The self-organizing consciousness. *Behavioral and Brain Sciences*, *25*, 297.

Perruchet, P., Cleeremans, A., & Destrebecqz, A. (in preparation). When automatic priming prevails over conscious expectancy in a simple reaction time task. *Journal of Experimental Psychology: Human Perception and Performance*.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance X* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*, 211–228.

Pothos, E. M. (in press). The rules versus similarity distinction. *Behavioral and Brain Sciences*.

Rapoport, A., & Budescu, D. V. (1997). Randomization in individual choice behavior. *Psychological Review*, *104*, 603–617.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *5*, 855–863.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A reevaluation. *Journal of Experimental Psychology: General*, *125*, 123–138.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure.*Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 585–594.

Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations. Cambridge, MA: MIT Press.

Saffran, J. R., Newport, E. L., Aslin, R. N., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.

Servan-Schreiber, D., Cleeremans, A., & McClelland, J. L. (1991). Graded state machines: The representation of temporal contingencies in simple recurrent networks. *Machine Learning*, *7*, 161–193.

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

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

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science*, *16*, 1–40.

Stadler, M. A. (1992). Statistical structure and implicit serial learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 318–327.

Taylor, T. L., & Klein, R. M. (1998). On the causes and effects of inhibition of return. *Psychonomic Bulletin & Review*, *5*, 625–643.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of the literature. *Psychological Bulletin*, *77*, 65–72.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. Nature, 427, 352–355.

Wegner, D. M. (2002). The illusion of conscious will. Cambridge, MA: Bradford Books/MIT Press.