Psychol Res (1997) 60: 4–13

ORIGINAL ARTICLE

Pierre Perruchet · Emmanuel Bigand ·
Fabienne Benoit-Gonin

The emergence of explicit knowledge during the early phase of learning in sequential reaction time tasks

Received: 3 June 1996 / Accepted: 28 October 1996

Abstract Five experiments investigated the formation of explicit knowledge of a repeating sequence in a sequential reaction time task. Reliable explicit knowledge was obtained even though various conditions prevented the selective improvement of RTs (Exps. 1–4). This knowledge emerged early during training. Participants were able to recognize segments of the sequence (Exps. 3 and 4) or correctly assess the probabilities of transition of the target between successive locations (Exp. 5) after only two blocks of training trials. These findings rule out an interpretation of sequence learning that posits that explicit knowledge emerges from implicit knowledge during the course of training. Although these findings are compatible with a framework centered around the notion of dissociation between implicit and explicit knowledge, they are also consonant with a questioning of the usefulness of the concept of implicit knowledge.

Introduction

According to the majority of authors, implicit learning elicits behavioral modifications that testify to the elaboration of a knowledge base which is unavailable to explicit thought. This feature is considered as so important that it constitutes the first of the criteria put forward by several contributors to define implicit learning. The main conclusion from Reber's (1989) extensive review of the literature is that "implicit learning produces a tacit knowledge base" (p. 219). It should be noted that the notion of implicit knowledge is invoked by those authors who subscribe to an abstractionist view of implicit learning (e.g., Seger, 1994), as well as by those who support a memory-based or episodic account (Higham & Vokey, 1994; Neal & Hesketh, in press).

In sharp contrast to the widespread use of the concept of implicit knowledge, supporting experimental evidence is ambiguous at best. In the context of sequential reaction time tasks, which will be our only concern here, Cohen, Ivry, and Keele (1990), Nissen and Bullemer (1987), and Willingham, Nissen, and Bullemer (1989) have claimed that the reaction times of subjects decrease more quickly when the same sequence is continuously repeated than when the series are randomly generated, even though subjects are not aware of the repeating sequence. However, these alleged demonstrations of the independence of performance and awareness were flawed by a series of biases (see Shanks & St. John, 1994, for a review). For instance, Perruchet and Amorim (1992) have argued that the claim of independence was primarily based on the use of inadequate tests of conscious knowledge. In particular, most of the earlier evidence relied on a generation task in which subjects were *not* told to generate the repeating sequence that was displayed during the training phase. Perruchet and Amorim used both a new version of the generation task (termed "free generation test") and a recognition test as their measures of explicit memory. In both cases, participants were explicitly told to retrieve their memories of the training sequence. Using these tests, they showed the existence of a close parallelism between conscious knowledge of fragments of the repeating sequence and performance in the very same experimental settings that were previously used to demonstrate dissociation.

More recently, Reed and Johnson (1994) presented new evidence for implicit knowledge. After training, participants were submitted to a cued generation test in which they were asked to respond to two targets as in the training phase and then to generate the locations which they thought had followed these two cues in the prior phase. Participants trained with a repeating sequence performed no better in this task than control subjects previously exposed to randomly generated sequences. In contrast, other participants trained with the same repeating sequence exhibited a significant increase in RTs

P. Perruchet (✉) · E. Bigand · F. Benoit-Gonin
LEAD, Faculté des Sciences, Université de Bourgogne,
6 Bd. Gabriel, 21000 Dijon, France;
e-mail: perruche@satie.u-bourgogne.fr

when the sequence was changed, hence providing evidence of implicit learning. Such a result seemingly testifies to some knowledge of the structural constraints of the sequence unavailable to explicit thought. However, Shanks and Johnstone (in press) have argued that this result was due to a flaw in Reed and Johnson's design. In brief, the sequences used to train the participants who acted as controls for the cued-generation test were more similar to the repeating sequence than were the new sequences used to demonstrate negative transfer with RT measures. In three experiments, Shanks and Johnstone revealed participants' explicit knowledge of the repeating sequence with cued prediction and free generation tests, thus invalidating Reed and Johnson's demonstration of implicit knowledge using their own paradigm.

However, it is worth noting that most studies of sequential learning involve extensive training. For instance, Shanks and Johnstone (in press) trained their subjects over 16 blocks of trials, thus presenting the repeating sequence 128 times. Because explicit knowledge is always assessed after completion of training, it may be argued that explicit knowledge emerges late during the session and follows RT improvement. Perruchet and Amorim (1992) provided evidence running counter to this argument. However, given the theoretical implications of the issue under discussion, notably with regard to the theories positing that explicit knowledge emerges from early implicit knowledge (e.g., Pascual-Leone, Grafman, & Hallett, 1994), it seemed necessary to provide additional support for Perruchet and Amorim's conclusion that explicit knowledge emerges during the early phase of practice. In order to reinforce the generality of this conclusion, the present series of experiments involved different stimuli (spatial location of a target or tones of different frequencies), various lengths of sequences (10, 12, or 16 trials), different experimental designs (e.g., between- or within-subject control), and different measures of explicit knowledge (a conventional recognition task or a new probability assessment task).

Experiment 1

Method

Participants. The participants consisted of 24 students at the René-Descartes University. They included 15 women and 9 men, with a mean age of 22 years.

Materials. The target stimuli were asterisks, whose four possible positions on the computer screen were indicated by a permanent mark at the very bottom of the monitor. Participants responded by pressing the *V*, *N*, *semi-colon*, and *equal-sign* keys on an azerty keyboard. The keyboard was positioned so that the keys, which were covered with blue patches, were aligned approximately with the target locations.

Procedure. Most of the characteristics of this experiment were borrowed from the Willingham, Greeley, and Bardone (1993) study. The participants were instructed to press the key that was below the position where the asterisk had appeared as fast as possible. Once the correct key was pressed, the target was removed and, after a 250-ms delay, the next stimulus appeared. All the participants completed five blocks of 84 trials, separated by a break of about 1 min.

Half of the participants were presented with the repeating sequence. They saw four random asterisks at the start of each trial block, and then a 16-position repeating sequence was repeated five times within a trial block. The location of the target was determined randomly, except that the four positions appeared four times, and two asterisks could not appear in the same position consecutively. Each subject saw a different repeating sequence.

The other half of participants were tested in a random condition. The 84-trial sequence of a block was generated using the same algorithm as for the structured group. However, the 16-trial sequences were different within a block and between the two blocks.

Immediately after this RT task, participants from the structured group performed a recognition task. They were shown 32 four-trial sequences. These sequences comprised the 16 different four-trial sequences that can be formed with the cycled repeating sequence. The other four-trial sequences were generated at random with the constraint that each sequence had to differ from the old sequences by at least one position. In addition, the following restrictions were imposed: No successive trials should occupy the same location and, in the 16 new sequences, the four possible locations appeared the same number of times (as in the 16 old sequences). Old and new sequences were intermixed and displayed in random order.

The four-trial sequences for recognition were displayed as a series of digits ranging from *1* to *4* and corresponding to asterisk positions from left to right. Participants were asked to indicate on a five-point scale if they believed that the four-trial sequence displayed in the coded format was part of the series they had seen in the previous phase. The scale was displayed on the screen and its endpoints were clearly labeled "not seen, sure" at the left and "seen, sure" at the right. Participants moved a cursor to below a point of the scale using the arrow keys of the keyboard and then validated their choice by pressing the *zero* key.

Results and discussion

As shown in Fig. 1, the mean RTs of correct responses for participants exposed to the structured sequences were lower that those of participants exposed to the random sequences for each of the five training blocks. The main effect of groups was marginally significant, $F(1,22) = 3.18$, $p = .088$. There was some evidence that the selective effect of repetitions emerged early during training. Planned comparisons showed that the differences in RT were not significant for Block 1, $F < 1$, marginally significant for Block 2, $F(1, 22) = 3.08$, $p = .093$, and fully significant for Block 3, $F(1,22) = 4.46$, $p = .046$. However, similar comparisons for Blocks 4 and 5 yielded non-significant results, $F(1,22) = 2.84$, $p = .106$ and $F(1,22) = 2.94$, $p = .100$, respectively. Overall, the interaction between Groups and Blocks was not significant, $F(4,88) < 1$.

The responses on the five-point scale of recognition were scored from 1 (unrecognized) to 5 (recognized). The mean score on old sequences, $M = 3.42$, reliably exceeded the mean score on new sequences, $M = 3.07$, $F(1,11) = 26.48$, $p < .0001$. This latter result replicates Willingham et al.'s (1993) finding using a different measure of explicit knowledge. It may be recalled that Willingham et al. assessed explicit knowledge through free reports, on the one hand, and by means of a test in

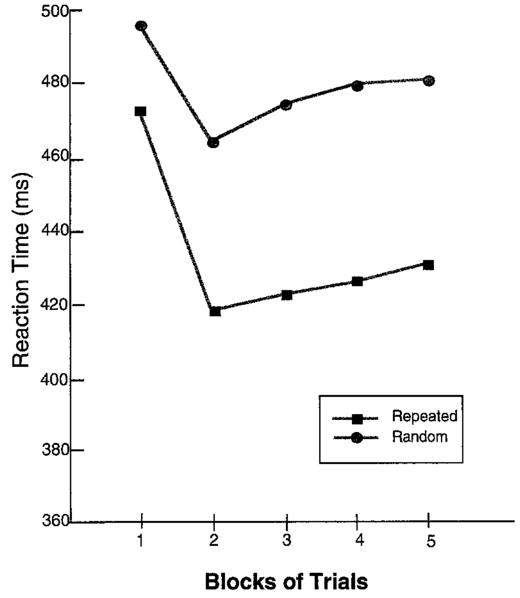

Fig. 1 Exp. 1: mean reaction times across 5 blocks of trials for a group of subjects trained with a repeated sequence and for another group of subjects trained with random sequences

which participants were asked to recognize the whole repeating sequence, on the other. Willingham et al. observed reliable explicit knowledge of the repeating sequence with both sets of measures.

However, Willingham et al. (1993) reported a straightforward indication of learning in the RT data. Given that our procedure was as similar as possible to that used by Willingham et al., we can currently offer no explanation to account for our difficulty to replicate their results. Arguably, the discrepancy may be due to the fact that Willingham et al. used a larger number of subjects than we did (45 vs. our 12). However, even if this (somewhat uninteresting) explanation turned out to be the correct one, it is still worthwhile to consider that, in some conditions, recognition can be fully reliable while RT measures reveal only moderate evidence of learning. This contrast was investigated further in Exp. 2.

Experiment 2

This experiment was designed to investigate whether recognition would persist above random response level in experimental conditions even less conducive to the selective improvement of RT measures. All the prior attempts to reveal dissociations between performance and explicit knowledge have had the opposite aim. For instance, many experiments have involved a dual task paradigm, the idea being that the attention devoted to the secondary task will affect explicit knowledge while keeping RT improvement intact. Our objective here was to manipulate a variable whose detrimental effect is presumably limited to motor performance. Selecting this variable was not a simple matter, given that, although there is overwhelming evidence for RT improvement in serial reaction time tasks, the mechanisms underlying this improvement are still poorly understood. We reasoned that the improvement in RT in repeated sequence paradigms is presumably due to some sensori-motor anticipation processes taking place between the response to the preceding stimulus and the appearance of the next target. If this is the case, suppressing the response stimulus interval (RSI) would prevent the action of these preparatory processes, hence eliminating the selective improvement in RT for the repeating sequence. Our hypothesis was that this manipulation would not affect the measure of explicit knowledge.

Method

Participants. The participants consisted of 24 new subjects from the same pool as for Exp. 1. They included 18 women and 6 men with a mean age of 25 years.

Material and procedure. These were the same as in Exp. 1, with the following exceptions. First, the RSI, which was 250 ms in Exp. 2, was omitted. (Below, we consider that the RSI = 0. In fact, a small number of milliseconds elapsed due to the time taken to run the program instructions for response analysis and stimulus display.) Secondly, participants had to rate the likelihood that each four-trial set had been displayed during training on a continuous scale instead of on a five-point scale.

Results and discussion

RTs for repeated sequences were in fact longer than RTs for random sequences, although not significantly so, $F(1,22) = 1.06$. RTs decreased across the blocks, $F(4,88) = 9.16$, $p < .0001$, but this improvement was the same for the two groups of participants, as attested by the non-significant interaction between Groups and Blocks, $F(4,88) = 1.51$, $p = .206$.

The recognition scores were digitized on a 100-point scale. Participants scored reliably higher for the segments belonging to the repeating sequence, $M = 61.86$, than for random segments, $M = 51.68$, $F(1,11) = 6.63$, $p = .026$. Thus, the suppression of the response stimulus interval, which prevents the expression of learning with RT measures, does not prevent participants from recognizing the repeating sequence.

Experiment 3

In the first two experiments, participants were shown the repeated sequence during five training blocks. The main objective of Exp. 3 was to investigate whether recognition of sequence fragments persists above random response level when the number of repetitions is decreased even further. This experiment was designed in the same way as the first one, but training was reduced to two blocks of trials.

Method

Participants. The participants consisted of 24 new subjects from the same pool as for Exps. 1 and 2. They included 17 women and 7 men with a mean age of 24 years.

Material and procedure. These were the same as in Exp. 2, with the following two exceptions. First, there were two blocks of training instead of five. Because each block included five 16-trial sequences, the experimental group was shown the repeating sequence ten times, whereas the control group was shown ten different sequences. Secondly, the RSI was set to 500 msec.

In addition, all the participants performed the recognition test, whereas only the participants from the repeating sequence group performed this test in the prior experiments. For the random group, one of the ten different training sequences was randomly selected, and the old sub-sequences for recognition comprised the 16 different four-trial sequences that can be formed with this sequence. Recognition was assessed on a five-point scale.

Results and discussion

Figure 2 shows the mean RT of correct responses for participants exposed to structured and random sequences for each of the two blocks. (The four initial random trials of each block were discarded for the purposes of analysis in both groups of participants.) An ANOVA performed with Group as a between-subjects factor and Blocks as a repeated measures factor indicated a significant effect for Blocks, $F(1,22) = 14.23$, $p = .001$, but although RTs were shorter for structured subjects than for random subjects, this effect was only marginally significant, $F(1,22) = 3.29$, $p = .083$. In addition, there was no interaction between Groups and Blocks of trials, $F < 1$.

For the recognition task, the responses on the five-point scale were scored from 1 (unrecognized) to 5 (recognized). For the structured group, the mean score on old sequences, $M = 3.53$, reliably exceeded the mean score on new sequences, $M = 3.12$, $F(1,11) = 6.84$, $p = .024$. For the control group, the mean score on the sequences built from a 16-trial sequence displayed once, $M = 3.36$, did not differ from the mean score on new sequences, $M = 3.39$, $F < 1$. The interaction between Groups and Status of items was significant, $F(1,22) = 6.13$, $p = .022$. Thus, this experiment indicated that participants were able to recognize the components of a repeated sequence after only two blocks of training, although they failed to recognize the components of a sequence displayed one time among other different sequences.

Experiment 4

In most sequential reaction time tasks, including those used above, the target stimuli are locations on a computer screen. In order to extend the generality of the above conclusions to different experimental conditions, Exp. 4 involved pure tones of different frequencies. Note that this change introduces deep modifications in the way in which the task is introspectively apprehended by the participants. This is primarily due to the fact that response stimulus mapping, which is direct in the standard procedure, requires some controlled intermediary processes in the case of tones.

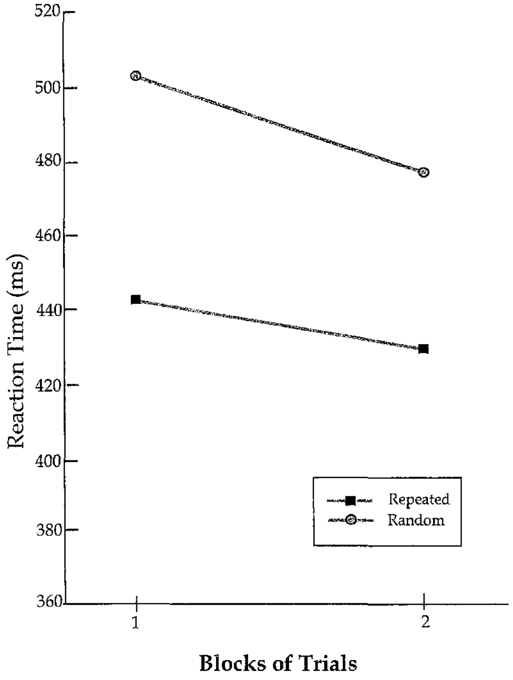

Fig. 2 Exp. 3: mean reaction times across 2 blocks of trials for a group of subjects trained with a repeated sequence and for another group of subjects trained with random sequences

This initial modification made certain other changes necessary. Indeed, pilot experiments revealed that participants experienced great difficulty in responding selectively to four different tones. (This characteristic may be due partly to the fact that we used the built-in loudspeakers of the computer which provided only moderate acoustic quality.) As a consequence, the number of different tones was reduced to three. With only three different stimuli, the number of possible sequences of a given length is substantially reduced, whereas the proportion of potentially salient events (such as trills and runs) is substantially increased. In order to minimize the risk of biases, we constructed a fixed, well-controlled repeated sequence instead of generating these sequences at random as in the prior experiments. In Exp. 4, participants were exposed to two training blocks, each comprising eight presentations of a fixed 12-trial sequence. They were then asked to recognize the 4-trial series constituting the repeated sequence, as in the preceding experiments.

Method

Participants. The participants consisted of 24 students at the University of Bourgogne. They included 18 women and 6 men with a mean age of 22 years.

Material. The target stimuli consisted of three tones of 300, 1500, and 3500 Hz delivered through the built-in loudspeakers of the computer. Participants responded by pressing the keys *1*, *2*, and *3* on the numeric keypad of the keyboard.

Procedure. The participants were instructed to press a key as quickly as possible after hearing a tone. They had to press *1*, *2*, and *3* for the lower, intermediate, and higher pitch, respectively. The tone ended as soon as the correct key was pressed. After a 500-ms delay, the next tone was delivered. All participants completed two blocks of 100 trials, separated by a break of about 1 min.

Half of the participants listened to the repeating sequence. The first three and the last three tones of each block were randomly selected among the three pitches, with the restriction that no repetition could occur. For the other trials, a 12-position sequence was repeated eight times. This sequence was *213231312321*, in which *1*, *2*, and *3* designated tones of increasing frequency. This sequence had notable characteristics. Each tone was presented four times and, when cycled, each of the 6 (3×2) pairs of tones was presented twice and each of the 12 ($3 \times 2 \times 2$) triplets of tones was presented once. This implies that the simple frequencies of occurrence, as well as the first-order and the second-order probabilities of transition, provide no information. This further implies that the number of perceptually salient pairs or triplets of tones (such as trills and runs) included in the sequence corresponds exactly to a random distribution. The only structural constraints were the third-order probabilities of transition.

The other half of the participants were tested in a random condition. The first three and the last three tones of each block were randomly selected among the three pitches, as in the structured group. The other trials were also generated randomly, except that, as in the structured group, the tones had equal frequencies of occurrence and two tones of the same pitch were never presented consecutively. The random sequences were different between the two blocks for a subject and between subjects.

Immediately after this RT task, participants from the structured group performed a recognition task. They were presented with the 12 four-trial sequences that can be formed from the cycled repeating sequence and with the other 12 possible four-trial sequences. (Indeed, with three stimuli it is possible to generate only 24 four-trial sequences without immediate repetition.) Although the 24 sequences were intermixed and displayed in random order, Old and New four-trial sequences were yoked, so that the sequences from any pair differed only in their last trial. It should be noted that, as a by-product of the nature of the repeating sequence, Old and New four-trial sequences comprised the same pairs and triplets of tones, including, of course, the salient ones. (In fact, both Old and New sequences comprised all the possible pairs and triplets, each possible pair being repeated six times, and each possible triplet being repeated twice.)

The four-trial sequences for recognition were delivered through the loudspeakers at the rate of one tone per second. Participants could hear the sequence as often as they wished by pressing the space bar. They were asked to indicate on a five-point scale if they believed that the sequence was part of the series they had heard in the previous phase. The scale was displayed on the screen and had clear endpoint labels: "not heard, sure" at the left and "heard, sure" at the right. Participants moved a cursor to below a point of the scale using the arrow keys of the keyboard, and then confirmed their choice by pressing the *zero* key (which was covered with a green patch).

Results and discussion

There was no indication of learning in the RT task. Participants from the structured group made fewer errors than participants from the random group (9.58% vs. 11.81%), but the difference was not significant, $F(1,22) = 1.02$. It can be noted that the number of errors was larger than in comparable tasks using spatial location as stimuli, in which the rate of errors rarely exceeds 5%. An ANOVA performed on the RTs of correct responses with Groups as a between-subjects factor and Blocks as a repeated measures factor indicated a significant effect for Blocks, $F(1,22) = 22.01$, $p = .0001$, but no effect of Groups, $F = (1,22) = 1.28$, $p = .270$, and no interaction between Groups and Blocks, $F < 1$. In fact, random subjects tended to respond more quickly than structured subjects.

For the recognition task, the responses on the five-point scale of the participants from the structured group were scored from 1 (unrecognized) to 5 (recognized). The mean score on old sequences, $M = 3.55$, reliably exceeded the mean score on new sequences, $M = 3.11$, $F(1,11) = 7.17$, $p = .021$.

On the suggestion of the Editors of this special issue, the recognition scores were analyzed in greater detail. The to-be-recognized sequences were ordered as a function of the mean recognition scores, irrespective of their actual status. A quite remarkable pattern emerged from the analysis of the results: the sequences which included only two different tones (such as *2121* or *3232*) scored far lower that the sequences including the three tones, $M = 2.32$ vs. $M = 3.67$, $F(1.11) = 34.5$, $p < .0001$. There was no difference in the scores for the two-tone sequences as a function of their actual status (mean for old: 2.33; mean for new: 2.31). When these sequences were deleted from the analysis, the residual difference between old and new sequences (3.79 vs. 3.51) was no longer significant, $F(1,11) = 1.86$, $p = .19$. This particular result pattern accounts for positive recognition when all the items are considered, because New sequences included more sequences with only two tones than Old sequences (4 vs. 2, respectively).

In fact, this characteristic of the recognition sequences reflected a structural feature of the repeating sequence. Indeed, the repeating sequence included only two 4-trial sub-sequences comprising only two tones (*3131* and *2121*), whereas randomization would be expected to produce three subsequences of this kind. Thus, a reasonable interpretation of the results is that participants become sensitive to the relative rarity of this kind of sequence during training and tend not to accept them as old during the recognition test. The specific way in which these sequences were subjectively coded remains obscure. Indeed, a sequence such as *3131* may be represented alternatively as the repetition of the pair *31*, the overlapping of two trills (*313* and *131*), or even a sequence in which there is no *2*. Whatever the subjective coding, however, it turns out that subjects became sensitive to a relatively abstract property of the repeating sequence and were able to discriminate between the sequences on this basis during the recognition test.

It is somewhat puzzling that RTs were not sensitive to the same constraint. It is possible that the difficulty inherent in the task, as the high rate of errors testifies, prevented selective RT improvement. First, the perceptive discrimination of the tones was somewhat difficult for some participants (a characteristic that may be due to the mediocre acoustic quality of listening provided by the built-in loudspeakers of the computer). Secondly, as mentioned above, key pressing involved a somewhat arbitrary mapping between pitches and spatial locations (although the order of pitches along a frequency dimension matched the order of keys along a spatial dimension). Whatever the actual reasons, however, it is worth considering again that conditions hampering the expression of learning through RT measures do not prevent recognition of the repeating sequence.

Experiment 5

In the first four experiments, recognition scores were above random response level irrespective of whether participants were trained during five (Exps. 1–2) or two (Exps. 3–4) blocks of trials. Somewhat surprisingly, the level of recognition did not appear to be related to the number of training trials. For instance, if we consider only the experiments using a five-point scale of recognition for the sake of comparison, the differences between the scores for old and new sequences are .35 after five training blocks (Exp. 1) and .42 after two training blocks (Exps. 3–4). However, these data are indicative only because these experiments differ on a number of parameters. Experiment 5 was primarily designed to permit the direct comparison of recognition scores after five and two blocks of training within a single experiment.

In order to test whether the existence of early explicit knowledge may be generalized to other conditions, we implemented several other procedural changes. First, we used a control procedure in which random trials were intermixed with the repeated sequence instead of being presented to another group of subjects. Stadler (1993) showed that learning occurs in these conditions, and it turns out that this procedure provides a control that possesses several advantages over the more traditional arrangements (see the discussion by Curran, this issue).

A characteristic of the prior experiments is that evidence for learning in RT measures is always weak, if not completely absent. Even when RTs were shorter for structured subjects than for random subjects, the Block by Groups interaction was never reliable. For Exps. 1–3, this difficulty may be due to the fact that the repeated sequences were randomly generated 16-trial sequences. These sequences presumably had weak statistical regularity (Stadler, 1992). Likewise, the sequence used in Exp. 4, although shorter, included only third-order sequential dependency rules. Presumably, the within-subject control used here should make the repeated sequence still less salient than in the standard procedure. In order to compensate for the increased difficulty due to the mix of repeated sequence and random trials, Exp. 5 involved a 10-trial sequence with relatively strong statistical regularities.

A final change related to the test of explicit knowledge. The validity of the recognition measure used in the preceding experiments is sometimes questioned, essentially because recognition may be mediated by perceptual or motor fluency. According to Willingham et al. (1993), the digital coding used in the preceding experiments ensured that performance was not due to this factor alone. In addition, our negative findings concerning motor performance run counter to this interpretation, at least with regard to motor fluency. However, in order to extend the generality of our conclusions regarding explicit knowledge, we used a new test in Exp. 5 devised to assess the subjective probabilities of target transition between successive locations.

Method

Participants. The participants consisted of 36 students at the University of Bourgogne. They included 34 women and 2 men with a mean age of 20 years.

Material and procedure. The stimuli were the same as those in Exps. 1–3. All the participants were presented with the repeating sequence, intermixed with random trials. Each block comprised 84 trials. After 4 initial random trials, there were 5 presentations of a 16-trial sequence composed of a 10-trial repeating component followed by a 6-trial random component. The repeating sequence was *1324324132*. The location of the target in the random sequence was determined randomly, with the restriction that two asterisks could not appear in the same position consecutively. In addition, in each block, the 4 positions appeared in the same proportions as in the repeating sequence (i.e., *1* and *4* twice, and *2* and *3* three times). Half of the participants were exposed to 2 training blocks, and the other half to 5 blocks.

Immediately after this RT task, all the participants performed a probability assessment task. In each trial, participants saw a cue composed of 1, 2, or 3 successive targets. They had to press a key in response to each of these targets, as in the training task. The participants then had to assess the probability that the next target would be each of the three possible successors in the sequence which they had previously seen. First, the four possible locations and the 12(4 × 3) possible pairwise transitions were displayed in random order. Then the eight 3-trial sequences which had really been presented during the training phase were shown. The presentation of all the combinations of two successive locations was intended to minimize the possibility of subjects learning from the cues during the test phase, while the restriction of the 3-trial locations to those really presented was intended to keep the length of the procedure to manageable proportions.

Participants expressed their assessment of the probability of the different possible successors by means of a special device. The computer screen displayed a bar-chart with four bars, each of which corresponded to a location. The y-axis was graduated from 0 to 100 and, at the start of each trial, three bars were set to 33 and the fourth (the bar corresponding to the location of the last trial of the cue) to zero. Participants had to press a key corresponding to a location in order to increment the bar for this location. The same keys were used as during the training phase. When a key was pressed, the values for the two other possible locations were automatically decremented, so that the total of the three values automatically equalled 100. Participants were instructed to shape the bar-chart so that it mapped the subjective probabilities (in per cent) for the trial following the cue to be at each of the locations. (Note that in order to facilitate the handling of the device subjects could not increment the bar corresponding to a repetition. This procedure is insensitive to the knowledge that there was no repetition. However, there is much evidence that this piece of knowledge is acquired by all of the participants.)

Results and discussion

Because repeating and random sequences were alternated, subjects were unable to anticipate the location of the target at the beginning of the repeating sequence on the basis of the prior events. In the following analysis, we pooled the first two trials of the repeating sequence with the random sequence, so that in each 16-trial sequence, 8 trials were considered as repeated and 8 trials were considered as random. (Note that the status of the second trial of the repeating sequence is questionable. However, had learning occurred on this trial, this would have improved the performance on the "random" sequence and hence decreased the likelihood that learning would be detected.)

As shown in Fig. 3, there was a clear indication of learning with RT measures. For the group exposed to two training blocks, an ANOVA performed with Blocks and Type of sequence (repeated vs. random) as repeated-measure factors indicated effects for Blocks, $F(1,17) = 4.88$, $p = .041$, Type of sequence, $F(1,17) = 4.008$, $p = .061$, and interaction between Blocks and Type of sequence, $F(1,17) = 4.40$, $p = .051$, all of these effects being around the conventional significance threshold. For the group exposed to five training blocks, the same analysis demonstrated significant effects of Blocks, $F(4,68) = 2.55$, $p = .047$, Types of sequence, $F(1,17) = 13.12$, $p = .002$, and interaction between Blocks and Type of sequence, $F(4,68) = 5.57$, $p = .0006$. The effect of the type of sequences was reliable at an early stage of training. Planned comparisons showed that the differences in RT for repeated and random trials were not significant for Block 1 ($F < 1$) but reached significance for each of the following blocks, $Fs(1,17) > 7.44$, $ps < .014$.

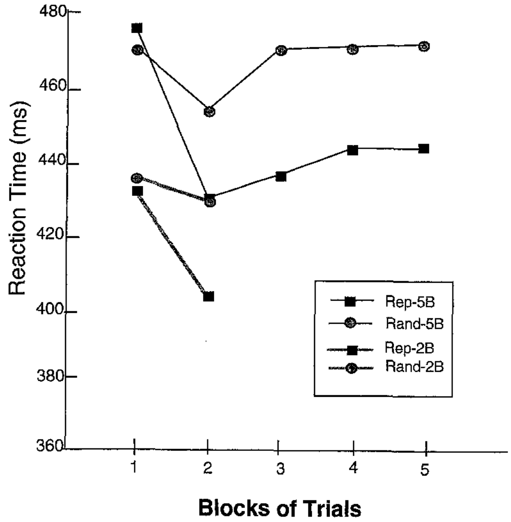

Fig. 3 Exp. 5: mean reaction times across 2 blocks of trials for one group of subjects and across 5 blocks of trials for the other group of subjects. Within each group, RTs are given for the repeated part and for the random part of each block

Performance in the probability assessment test was scored as the value attributed to the correct completions. The overall percentage attributed to the correct outcome was 37.29 for the group exposed to two training blocks, and 38.13 for the group exposed to five training blocks. This performance significantly exceeded the 33% random response level: $t(17) = 2.77$, $p = .0065$; $t(17) = 3.42$, $p = .0016$, respectively. An ANOVA was performed with the Groups (2 blocks vs. 5 blocks of training) as a between-subject factor, and the Length of the cue (1, 2, or 3) as a repeated-measure factor. There were no main effects or interaction (all $Fs < 1$).

The fact that the correct response level was independent of the amount of training supports our hypothesis, which stemmed from a post hoc comparison of the results of Exp. 1 and Exps. 3 and 4. The question of interest is whether this feature is specific to the test of explicit knowledge. In fact, the present data provide no evidence for a positive response because, as shown in Fig. 3, selective RT improvement was also limited to the early phase of training. Presumably, the fact that there was no indication of an increase in performance after the second training blocks, irrespectively of whether RT or explicit predictions are considered, is due to certain procedural features. However, we can advance no particular hypothesis to account for this idiosyncratic pattern of results.

The independence of the level of correct responses from the length of the cue could be anticipated, given that, due to the nature of the repeated sequence, 3-trial cues did not allow better predictions than single-trial cues. Our reasoning was that an effect could emerge, nevertheless, given that it seems easier to respond "*L*" when asked to identify the alphabetic successor of *IJK* than when asked to name the successor of *K* alone, even though *IJ* actually provides no supplementary information. No such effect was obtained, thus indicating that participants were able to assess the relative probability of the possible events following a single, given position.

General discussion

There was significant evidence for explicit memory of the repeating sequence in all the five reported experiments. In support of our main hypothesis, evidence for explicit memory was obtained after very limited training. A significant explicit knowledge score was regularly obtained after two blocks of trials, that is to say after only 10 (Exps. 3 and 5) or 16 (Exp. 4) presentations of the repeating sequence. This result was observed in a variety of situations: (a) with spatial location of a target or tones of different frequencies as stimuli, (b) with sequences of different lengths (10, 12, or 16 trials) and different structural constraints, going from predominantly first-order (Exp. 5) to third-order (Exp. 4) dependency rules,

(c) with different experimental designs, involving either between-subjects or within-subjects control, and (d) with two tests of explicit knowledge (a conventional recognition task and a new probability assessment task).

Still more interestingly, reliable explicit memory of the repeating sequence was observed, whereas there was no (Exps. 2 and 4) or only marginally significant (Exps. 1 and 3) evidence for learning from RT measures. This pattern of results has been previously reported (e.g., Perruchet & Amorim, 1992; Shanks Johnstone, in press). For instance, Shanks and Johnstone (Exp. 2) examined the knowledge of individual triplets constituting a 12-trial repeating sequence using a transfer task (implicit), on the one hand, and a generation task (explicit) on the other. Implicit and explicit measures revealed knowledge of four and eight locations, respectively. The authors concluded that free generation, a measure of explicit knowledge, "is if anything more sensitive to sequence knowledge than the magnitude of the negative transfer effect" (this latter measure is based on RTs).

The causes for the absence of the selective effects of repetition of RT in many of our conditions were not entirely clear. The only factor that was intentionally manipulated to disrupt the effect of repetition on RT was the suppression of the interval between response and stimulus in Exp. 2. A comparison of Exps. 1 and 2 suggests that this manipulation was indeed effective. A rough comparison between the other experiments suggests that other factors were also involved, such as the use of arbitrary stimulus-response mapping in Exp. 4. Interestingly, none of these factors seemed to influence the formation of explicit knowledge. These findings are reminiscent of the results obtained by Frensch, Buchner, and Lin (1994). These authors showed that both the use of first- or second-order dependency rules and the use of a response-stimulus interval of 500 or 1500 ms affected the RT performances, although they had no or only minor effect on free recall and generation tests (at least when only the effect of incidental learning was considered – that is to say, when analyses were limited to the first trial of the generation test).

A number of authors have suggested that a recognition test, which was used in most of the present experiments, might be sensitive to the increasing perceptual or motor fluency involved in dealing with the training task and might therefore overestimate the amount of genuine explicit knowledge about the task structure. Although this interpretation may be valid in some contexts, it is especially unlikely in the present studies. First, the recognition tasks were designed to minimize the possible "contamination" of performance by the effects of perceptual or motor fluency. Indeed, we were careful here to represent the sequences for recognition through a digital, arbitrary coding which was also used for the participants' responses. Secondly, the results obtained using recognition tasks in Exps. 1–4 were replicated in Exp. 5 using a probability assessment task, which can hardly be accused of being susceptible to indirect influences. Indeed, participants were found to be able to assess the relative probability of the possible events following a single, given position of the target. Thirdly, the results themselves do not support this kind of interpretation, insofar as explicit knowledge was apparent, whereas there was no reliable motor facilitation, at least as indicated by chronometric measures.

Our results are compatible with several theoretical frameworks, but not with all of them. Most contributors to the implicit learning field agree that performances in the RT task and in the explicit memory tasks are indicators of implicit and explicit knowledge, respectively. Interpretations differ with regard to the kind of relations posited between the two forms of knowledge. Some contributors posit that explicit knowledge emerges from implicit knowledge. For instance, the Pascual-Leone et al. (1994) study was aimed at investigating the neural modifications associated with the transfer of knowledge from an implicit to an explicit state in a sequential reaction time task. This framework supposes that modifications in RTs anticipate the appearance of explicit knowledge. Maybe the clearest implication of our results is the refutal of this conception, at least when it is framed as a general, ubiquitous proposal. It should be noted that such a conception, although infrequently advocated in the implicit learning context, is common in the developmental area. For instance, Karmiloff-Smith (e.g., 1992) argues in favor of a general iterative process of redescription of implicit knowledge into explicit knowledge. To the extent that laboratory results in adults are relevant to developmental issues in children, the present findings put forward new arguments against such a conception of development (see Perruchet & Vinter, in press, for a more general discussion).

In contrast, many contributors to the implicit learning and memory literature conceive of implicit and explicit forms of knowledge as independent. Most of the data put forward in support of this view consists of reports of improvement in performance without concomitant explicit knowledge. Some of our results illustrate the reverse dissociation, namely, the occurrence of explicit knowledge unaccompanied by a selective improvement in RTs. These results are obviously compatible with the hypothesis of a dissociation between implicit and explicit knowledge. However, we believe that they may be better understood within a framework resting on quite different postulates, as we will argue in the remainder of this paper.

Perruchet and Gallego (in press; see also Perruchet, Vinter, & Gallego, in press; Perruchet & Vinter, in press; Vinter & Perruchet, in press) recently proposed an account of implicit learning in which the notion of implicit knowledge has no place. We claimed that subjects learn to process the material by parsing it into small and disjunctive (i.e., not overlapping) units, as a mandatory consequence of the attentional processing of the input data. Units emerge from the automatic association of the primitive features that are processed conjointly in the attentional focus. This initial contention is close to other proposals, such as Stadler's (e.g., 1995) model of sequence learning (although in this latter model, units are not conceived as disjunctive). However, in complete contrast to current theories, these units, we claim, directly affect the way the external world is *consciously* perceived and processed. In this view, *implicit learning shapes the phenomenal experience of the world.* The phenomenon is observable in any natural situation in which implicit learning is assumed to operate, whether this situation concerns the acquisition of first or second language, natural categories, reading and writing abilities, or even sensitivity to musical structure. It is difficult to assert that our subjective experience of that part of the environment with which we interact in each of these cases remains unchanged while training progresses. Our argument is that these changes in the way we consciously perceive and interact with the environment are at the core of implicit learning. Note that, within this perspective, neither the conscious nor the unconscious aspects of mind are negated. However, they characterize different elements of mental life: associative processes and mechanisms responsible for the emergence of knowledge are intrinsically unconscious, and the resulting mental representations are intrinsically conscious. No other components are needed and, most importantly, the notion of implicit knowledge is objectless.

In keeping with this perspective, we suggest that all the measures obtained from an experimental implicit learning setting reveal the unconscious formation of conscious percepts and representations, whether these measures are collected during the training task (the so-called "performance") or in subsequent "explicit" tests. At first glance, this account leads us to anticipate a very close parallelism between performance changes and explicit knowledge about the structure of the situation. However, it must be considered that the content of the conscious knowledge base tapped by the two categories of measure is partly different. Let us suppose that a subject is shown a sequence including a salient fragment such as *ABC*. This fragment will presumably become a subjective perceptual and representational unit. This (explicit) knowledge may be detected through the fact that the RT to *C* after *AB* is shorter than the RT to *C* embedded in a random sequence. Likewise, this knowledge may be detected by the classification of *ABC* as old in a subsequent recognition test. Indeed, both phenomena imply the formation of *ABC* as a subjective unit. However, the recognition test also implies that the subject is able to recollect the spatio-temporal context of the initial encoding. The so-called explicit tests involve some memory about the spatio-temporal context in which knowledge has been acquired, whereas this is not needed in the reaction time task. The difference between the two ways of measuring the effects of the exposure to the material does not lie in the implicit versus explicit status of the underlying knowledge, but instead in the specific pieces of explicit knowledge which are required in order to perform the task.

This framework makes it possible to account for the observation of a parallelism between motor performance and explicit memory in normal subjects, such as was obtained, for instance, by Perruchet and Amorim (1992). Indeed, both kinds of measure imply the formation of the same units of coding. However, this framework also makes it possible to account for the observation of improved performance in RT tasks unaccompanied by the memory of the spatio-temporal context in which the new ability was acquired, as observed in studies with amnesic patients (e.g., Nissen & Bullemer, 1987). To put it simply, these dissociations are no longer accounted for in terms of the manifestation of implicit knowledge, which would be intact in amnesics. They are due to the fact that the motor task does not require certain elements of explicit knowledge that are required in the explicit memory task. Improved performance in the absence of the explicit recollection of having viewed the material during the training phase is not indicative of implicit knowledge, but of impoverished explicit knowledge.

To illustrate by means of a comparison, each of us has explicit knowledge about the fact that London is the capital of the U.K., and, presumably, some of us are able to specify the spatio-temporal context in which they read or heard this information, while others are not. This does not mean that the representation of London as the U.K.'s capital is explicit for the former and implicit for the latter: it is fully conscious for everyone. Simply, the context in which this explicit knowledge was acquired is either still available in memory, or forgotten. This is exactly the same, trivial phenomenon that is observed in what is commonly presented as a dissociation between performance (or implicit knowledge) and explicit knowledge.

More important for our current concern is the fact that our framework also accounts for performance above random response level in explicit memory tasks without concomitant RT improvement. Indeed, the improvement in RTs cannot be thought of as a ubiquitous consequence of the changed phenomenal experience of the training situation. The expression of learning through RT measures may depend on specific parameters. As a case in point, let us consider the effect of varying the response-stimulus interval between, say, 0, 500, and 1500 ms. Taken together, Frensch et al. (1994)'s and our findings suggest that a repeating sequence improves RT when the RSI is 500 ms but has only minor or no influence at all at the other two intervals. In contrast, explicit knowledge appears to a roughly similar extent in all the conditions. Chronometric measures, we suggest, need certain specific conditions to reveal an effect. To summarize, our account predicts parallelism between explicit memory of the training situation and performance improvement, because both depend on the change in the subjective coding of the displayed material. However, neither of these aspects is a direct expression of the change in phenomenal experience. The formation of explicit memories implies that the phenomenal change is attributed to its actual cause, and improvement in performance depends on the occurrence of certain parametrical conditions.

The findings of the present studies run counter to a specific model of cognition in which explicit knowledge emerges from implicit knowledge. More generally, they enable us to refute a common belief concerning the antecedence of performance changes over explicit knowledge in implicit learning settings. However, these findings are still compatible with interpretations starting from strikingly different postulates about the genuine existence of implicit knowledge.

Acknowledegments This research was supported by the University of Bourgogne and the Centre National de la Recherche Scientifique (ESA CNRS 5022). We are indebted to Bertrand Caceres for his help with data collection for Exps. 1–3. We would like to thank the Editors of this special issue, and also Michael Stadler and an anonymous reviewer, for their careful reviews and helpful suggestions on a previous draft of this article.

References

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

Frensch, P. A., Buchner, A., & Lin, J. (1994). Implicit learning of unique and ambiguous serial transitions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 567–584.

Higham, P. A., & Vokey, J. R. (1994). Recourse to stored exemplars is not necessarily explicit: A comment on Knowlton, Ramus, and Squire (1992). *Psychological Science*, *5*, 59–60.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: Bradford/ MIT press.

Neal, A., & Hesketh, B. (in press). Episodic knowledge and implicit learning. *Psychonomic Bulletin and Review*.

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

Pascual-Leone, A., Grafman, J., & Hallet, M. (1994). Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science*, *263*, 1287.

Perruchet, P., & Amorim, M. A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 785–800.

Perruchet, P., & Gallego, J. (in press). A subjective unit formation account of implicit learning. In D. Berry (Ed.). *How implicit is implicit learning*. Oxford: Oxford University Press.

Perruchet, P., & Vinter, A. (in press) Learning and development: The implicit knowledge assumption reconsidered. In M.A. Stadler & P. A. Frensch (Eds), *Handbook of implicit learning*. Thousand Oaks, CA: Sage Publications.

Perruchet, P., Vinter, A, & Gallego, J. (in press). Implicit learning shapes new conscious percepts and representations. *Psychonomic Bulletin and Review*.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.

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.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, 115, 163–196.

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

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

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

Stadler, M. A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memory and Cognition*, *21*, 819–827.

Stadler, M.A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 674–685.

Vinter, A., & Perruchet, P. (in press). Relational problems are not fully solved by a temporal sequence of statistical learning episodes. *Behavioral and Brain Sciences*.

Willingham, D. B., Greeley, T., & Bardone, A. M. (1993). Dissociation in a serial reaction time task using a recognition measure: Comment on Perruchet and Amorim (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1424–1430.

Willingham, D. B., Nissen, M. J. & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1047–1060.