Available online at www.sciencedirect.com
ScienceDirect
Consciousness and Cognition xxx (2007) xxx–xxx
Consciousness and Cognition
www.elsevier.com/locate/concog

Implicit sequence learning and conscious awareness

Qiufang Fu [a], Xiaolan Fu [a,*], Zoltán Dienes [b]

[a] *State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, 10A Datun Road, Chaoyang District, Beijing 100101, China*
[b] *Department of Psychology, University of Sussex, UK*

Received 11 October 2006

Abstract

This paper uses the Process Dissociation Procedure to explore whether people can acquire unconscious knowledge in the serial reaction time task [Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the Process Dissociation Procedure. *Psychonomic Bulletin & Review, 8*, 343–350; Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 354–369]. Experiment 1 showed that people generated legal sequences above baseline levels under exclusion instructions. Reward moved exclusion performance towards baseline, indicating that the extent of motivation in the test phase influenced the expression of unconscious knowledge. Experiments 2 and 3 revealed that even with reward, adding noise to the sequences or shortening training led to above-baseline exclusion performance, suggesting that task difficulty and the amount of training also affected the expression of unconscious knowledge. The results help resolve some current debates about the role of conscious awareness in sequence learning.

Keywords: Implicit learning; Sequence learning; Subjective measures; Objective measures; The process dissociation procedure; Unconscious knowledge; Reward; Task difficulty; Amount of training

1. Introduction

An issue that continues to divide psychologists, despite decades of research, is whether people can acquire and use unconscious knowledge. The two factions of psychologists can often be separated by whether they use objective or subjective measures of conscious knowledge. Researchers who measure the conscious status of mental states purely by objective measures, i.e. the ability to discriminate features of the *world* (worldly discrimination), tend to be skeptics concerning the existence of unconscious knowledge states (e.g. Dulany, 1997; Perruchet & Viner, 2002; Shanks, 2005). With objective measures, the ability to determine whether an item or part of an item has certain objective properties (for example, whether it occurred previously) is taken to indicate not just knowledge but also conscious knowledge. On the other hand, people who measure the conscious

* Corresponding author. Fax: +86 10 6487 2070.
E-mail address: fuxl@psych.ac.cn (X. Fu).

1053-8100/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.concog.2007.01.007

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

status of mental states by subjective measures, i.e. the ability to report or discriminate *mental states*, tend to accept the existence of unconscious knowledge (e.g. Dienes & Berry, 1997; Willingham & Goedert-Eschmann, 1999). If a participant says they know nothing, or cannot discriminate the cases where they know something from when they are guessing, any knowledge as revealed in worldly discrimination is taken to be unconscious knowledge.

The subjective methodology is based on the assumption that a conscious mental state is a mental state of which one is conscious (Rosenthal, 2002). Merely being able to discriminate features in the world shows only that one has knowledge; it clearly does not show that one is conscious of having knowledge. It is only when there is evidence that a person knows that they know that there is evidence for conscious knowledge. While the rationale of subjective measures is intuitively appealing, there has been resistance to adopting subjective measures because of the problem of bias: How do we know that a subject who *says* they are guessing might not actually *think* they know something (for a discussion of bias, see Dienes, 2004; and for other assumptions of subjective measures, see Dienes & Perner, 2004)?

Jacoby (1991) suggested one could bypass the whole objective–subjective measure issue by the use of control. Conscious knowledge is knowledge one can control the use of. In the Process Dissociation Procedure (PDP), a person is asked to use the same knowledge to do opposite things: For example, in a subliminal perception experiment, to complete a stem or refrain from completing a stem with a word just briefly displayed (e.g. Debner & Jacoby, 1994). If the knowledge is conscious, people should be able to use the knowledge to perform whatever task is requested; if the knowledge is unconscious it will have the same consequence whatever one's intentions (i.e. to complete the stem with the just displayed word even when told to refrain from doing so). Jacoby's insight spawned a vast literature using the PDP (e.g. Buchner, Steffens, Erdfelder, & Rothkegel, 1997; Buchner, Steffens, & Rothkegel, 1998; Dienes, Altmann, Kwan, & Goode, 1995; Goschke, 1998; Kane, Picton, Moscovitch, & Winocur, 2000; McBride & Dosher, 1999; Reingold, 1995).

Recently, Destrebecqz and Cleeremans (2001, 2003) used PDP methodology to look at the conscious status of the knowledge acquired in the serial reaction time (SRT) task. The SRT task was introduced by Nissen and Bullermer (1987) and is now one of the most widely used implicit learning tasks (e.g. Cleeremans, Destrebecqz, & Boyer, 1998; Destrebecqz et al., 2005; Hoffman, Sebald, & Stöcker, 2001; Jiménez, 2003; Wilkinson & Shanks, 2004; Willingham, Wells, & Farrell, 2000). It is a choice reaction time task in which the sequence of buttons to be pressed is structured. The participant is told which button to press by a corresponding location on a screen being indicated. The participant can thus in principle follow instructions without being aware that a structured sequence exists. People come to respond faster when the sequence is maintained rather than switched, indicating structure has been learned. On the one hand, such learning occurs when people deny there is a sequence or cannot freely report it (subjective measures indicate the knowledge is not conscious; e.g. Willingham, Nissen, & Bullemer, 1989; Willingham et al., 2000; Ziessler, 1998); on the other hand, people can generally recognize the sequence or generate it, when such tests have established power (objective measures indicate knowledge is conscious; e.g. Cleeremans & McClelland, 1991; Perruchet & Amorim, 1992; Shanks, Wilkinson, & Channon, 2003).

Subjective and objective measures give opposite answers; can Jacoby's PDP shed any light on the conscious status of the knowledge? Can participants control their use of the knowledge? Destrebecqz and Cleeremans (2001, 2003) used PDP in the SRT task by having an *inclusion* and *exclusion* test after a training phase. In the inclusion test, participants were asked to generate the sequence by remembering it, or else guessing. In the exclusion test, participants were asked to freely generate a sequence that was different from the one they were trained on. Destrebecqz and Cleeremans also manipulated the response stimulus interval (RSI) during training, on the grounds that a long RSI would give conscious knowledge a greater chance both to form and to be applied. Consistently, they found that when RSI was zero, the number of chunks from their trained sequence generated in the inclusion task (I) was not significantly different from the number generated in the exclusion task (E), i.e. they found $I = E$. Further, E was greater than the number of chunks generated of a control sequence (the baseline, B), i.e. they found $E > B$. These results indicated that when RSI was zero, participants compulsively generated the sequence even when they were trying to refrain from doing so. They lacked control over the use of the knowledge. On the other hand, when RSI was 250 ms, participants had control over the use of the knowledge: They found that $I > E$, and also that participants could refrain from producing legal sequences above baseline levels, i.e. $E = B$. In sum, there was evidence of unconscious knowledge when RSI was zero, and evidence of conscious knowledge when RSI was 250 ms, according to the logic of PDP. The results for an RSI of zero, initially obtained in 2001 and replicated in 2003, provided new and crucial evidence that a simple laboratory task could be used to explore the properties of unconscious knowledge.

Wilkinson and Shanks (2004) and Norman, Price, and Duff (2006) did not replicate the crucial findings for an RSI of zero. They found I > E, indicating participants had some ability to control the use of their knowledge, and argued that Destrebecqz and Cleeremans obtained a null I = E probably because the latter lacked power. According to the logic of PDP, I > E indicates the presence of some conscious knowledge, but it does not rule out the possibility of there being some unconscious knowledge as well. Both Wilkinson and Shanks and Norman, Price and Duff found E = B. Wilkinson and Shanks calculated they had a power of about .80 to pick up the size of effect (E > B) found by Destrebecqz and Cleeremans, which is not impressively large, and Norman et al. had a power of only .75. Further no other explanation was offered by either author for the failure to directly replicate Destrebecqz and Cleeremans' crucial result. Wilkinson and Shanks (Experiment 3) used a trial-by-trial generation procedure which they argued should be more sensitive, but still could not replicate, nor explain the failure to replicate.

Below we show how the findings of both Destrebecqz and Cleeremans (2001, 2003), on the one hand, and Wilkinson and Shanks (2004) and Norman et al. (2006), on the other, can be replicated. Experiment 1 finds above-baseline exclusion performance, like Destrebecqz and Cleeremans—but only for people not offered a reward for good performance. Reward reduces exclusion performance to baseline, as found by Wilkinson and Shanks. Experiment 1 used a deterministic sequence; Experiment 2 shows that even with reward, when the sequence-appropriate stimulus occurs only 75% of the time in training, people generate above baseline. Finally, Experiment 3 shows even when the sequence-appropriate stimulus occurs commonly in training (87.5% of the time), unconscious knowledge emerges early in training, before detectable conscious knowledge emerges.

2. Experiment 1

One difference between Destrebecqz and Cleeremans (2001, 2003) and Wilkinson and Shanks (2004) is that the latter rewarded participants for good performance on the inclusion and exclusion tasks, whereas the former did not (Destrebecqz, personal communication, June 2, 2006). Visser and Merikle (1999) showed that rewards enabled people to exclude more successfully on a perception task (but not a memory task). The aim of Experiment 1 was to explore whether reward influences performance on the SRT generation tests. Given reward is an important factor then exclusion performance in the no-reward condition will be above chance, replicating Destrebecqz and Cleeremans' (2001, 2003), and performance in the reward condition will be at chance-level, replicating Wilkinson and Shanks' (2004).

2.1. Method

2.1.1. Participants

Fifty-six undergraduate students (24 male, 32 female) took part in the experiment. None of them had previously taken part in any implicit learning experiment. They were randomly assigned to the no-reward or reward groups ($n = 28$, per group), and were paid a ¥ 15 attendance fee.

2.1.2. Apparatus

The experiment was programmed in Virtual C++ 6.0 and run on Pentium-compatible PCs. The display consisted of a yellow background and four blue target-location bars, which corresponded to the key D, F, J, and K on the computer's keyboard from left to right. The bars were arranged in a horizontal line on the computer's screen and separated by intervals of 3 cm. The stimulus was a blue circle 1 cm in diameter that appeared above one of the four positions on each trial.

2.1.3. Materials

Two second-order conditional sequences (SOC1 = 3-4-2-3-1-2-1-4-3-2-4-1; SOC2 = 3-4-1-2-4-3-1-4-2-1-3-2) were used in the target-location task, in which every location is completely determined by the previous

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

two locations. The sequences were balanced for location frequency (each location occurred three times), transition frequency (each possible transition from one location to a different one occurred once), reversal (e.g. 1-2-1) frequency (one in each sequence), repetitions (no repetitions in either sequence), and rate of full coverage (see Reed & Johnson, 1994). The only difference between the sequences is in their second-order conditional structure. For example, 3-4 was followed only by a 2 in SOC1 but only by a 1 in SOC2.

2.1.4. Procedure

2.1.4.1. Training phase. The no-reward and reward groups received identical training phases, composed of 15 training blocks during which participants were exposed to a serial four-choice RT task. Each block consisted of 98 trials, for a total of 1470 trials. On each trial, a stimulus appeared at one of the four possible screen locations. Participants were instructed to respond as quickly and as accurately as possible by pressing the corresponding key. Keys D, F, J, and K corresponded to Locations 1-4, respectively. Participants were required to respond to Locations 1 and 2 with the middle and index finger, respectively, of their left hand and to Locations 3 and 4 with the index and middle finger, respectively, of their right hand.

Each block of target location trials began at a random point in one of the two sequences. The target was removed as soon as a correct key had been pressed, and the next stimulus appeared immediately (i.e., RSI = 0 ms). Response latencies were measured from the onset of the target to the completion of a correct response. Error responses were signaled to participants by means of a tone, and errors were recorded. One-minute rest breaks occurred between any two experimental blocks. For counter balancing purposes, half of the participants in each condition (10 in each group) were trained on SOC1 during the first 12 blocks and Blocks 14 and 15; and on SOC2 during Block 13. This design was reversed for the other half of the participants.

2.1.4.2. Test phase. The test phase involved two free-generation tests for the training sequence. After the training blocks, participants were informed that the targets had followed a regular repeating sequence, in which every location is completely determined by the previous two locations, and that they would have to perform a slightly different task now. At the beginning of each test, a stimulus would sequentially appear at two locations, and participants had to respond to them as in the training phase. After the completion of the second correct response the stimulus disappeared and participants were required to generate a sequence of 96 trials by pressing corresponding keys on the keyboard. Targets appeared when participants made their key press and remained on the screen until they made their next response. Participants under inclusion instructions were required to generate a sequence that resembled the training sequence as much as possible. Conversely, participants under exclusion instructions were required to generate another sequence but they had to try to *avoid* reproducing the sequential regularities of the training sequences. For example, if they thought the stimuli would appear in Location 2 after appearing in Locations 3 and 4, then they should press Location 1 or 3. Moreover, participants were forbidden from generating the same location twice or more in a row. Half of the participants first did the inclusion test, and then did the exclusion test; and half first did the exclusion test, and then did the inclusion test. The only difference between the two groups was that participants in the no-reward condition were not informed they would receive additional money for good performance, while participants in the reward condition were informed after training and before the generation test that they would receive an additional ¥ 50 for good performance (defined as a combined inclusion and exclusion performance of more than 56%) in order to encourage participants to do their best. Five of the 28 participants received the reward.

2.2. Results

2.2.1. Training phase

Participants trained with SOC1 and SOC2 were combined in all analyses. RT analyses were conducted for correct responses across 15 blocks. RTs for the first two targets of each block were excluded, because their locations could not be predicted. Trials with RTs greater than 1000 ms were dropped; these amounted to only 0.26% and 0.37% of the trials in the reward and no-reward groups, respectively. Fig. 1 shows the mean RTs obtained over the training phase. RTs for the two groups decreased from Blocks 1 to 12, increased dramat-

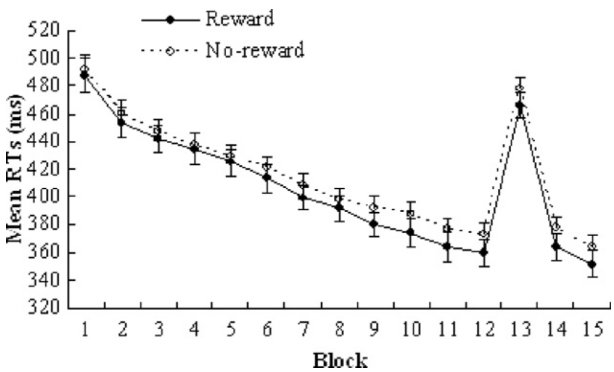

Fig. 1. Mean reaction times (RTs) across training blocks in Experiment 1, which used a deterministic sequence. Block 13 was a transfer block. Error bars depict standard errors. Ms, milliseconds.

ically on Block 13, and returned to lower level on Blocks 14 and 15. An ANOVA on RTs with incentive (rewards vs. no-rewards) as a between-subjects variable and blocks (15 levels) as a within-subject variable revealed only a significant effect of block, $F(14,756) = 154.96$, $MSE = 621.08$, $p < .001$. That is, there was no detectable difference between the two incentive groups in the training phase. To study the transfer effects, an ANOVA on RTs with incentive as a between-subjects variable and blocks (transfer block 13 vs. the average of blocks 12 and 14) as a within-subject variable revealed only a block effect, $F(1,54) = 404.87$, $MSE = 743.83$, $p < .001$. That is, participants of the two incentive groups learned equally about the sequential structure of the materials, as would be expected as participants were only informed about reward in the test phase.

Table 1 shows the means for errors. The pattern of error data did not indicate a possible speed-error trade-off for any of the above effects.

2.2.2. Test phase

Participants generated sequences of 96 key presses based on what they had learned about the training sequence. We computed the number of generated chunks of three elements that were part of the training or transfer sequence in both inclusion and exclusion tasks. An "own" triplet is a triplet that was part of the training sequence; an "other" triplet is a triplet that was part of the transfer sequence; a "neither" triplet is a triplet that was neither "own" nor "other." The maximum number of own or other triplets was 96. Fig. 2 shows the mean number of triplets generated for each group in Experiment 1. If participants had learned the training sequence explicitly, then they should be able to control the expression of their sequential knowledge. We would expect they would produce more own than other triplets in the inclusion task, and would not in the exclusion task. Thus, evidence of implicit knowledge would come from either the finding I = E or E > B, where I and E refer to the number of own triplets generated in the inclusion and exclusion tasks, and B refers to an appropriate baseline level of generation. We take the number of other triplets as the baseline because own and other triplets are counterbalanced across participants (Wilkinson & Shanks, 2004).

We first compared the number of own triplets (from the training sequence) generated under inclusion and exclusion instructions in the two conditions. An ANOVA with incentive (no-reward vs. reward) as a

Table 1
Mean error proportions across training blocks under the reward and no-reward groups in Experiment 1

Block	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Reward															
M	.020	.020	.025	.027	.029	.029	.029	.031	.030	.030	.023	.028	.053	.022	.031
SE	.004	.005	.005	.004	.006	.005	.007	.006	.006	.006	.004	.005	.008	.004	.006
No-reward															
M	.020	.021	.025	.027	.024	.022	.027	.020	.020	.026	.023	.028	.055	.019	.021
SE	.004	.004	.004	.004	.004	.002	.004	.004	.003	.004	.003	.004	.007	.003	.003

Note. Block 13 was a transfer block.

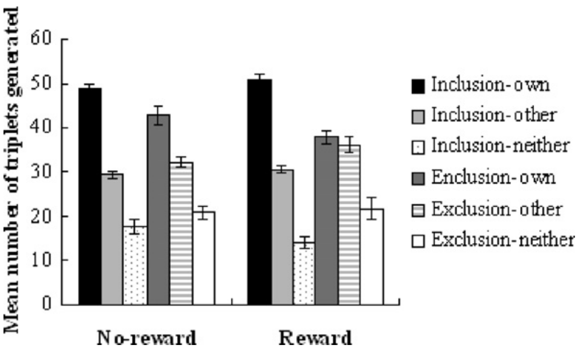

Fig. 2. Mean number of triplets generated by the reward and no-reward groups and tested under inclusion and exclusion instructions in Experiment 1. Own, number of SOC triplets generated from the training sequence; other, number of triplets from the alternate, untrained sequence; neither, number of triplets from neither the training nor the untrained sequence. Error bars depict standard errors.

between-subjects variable and instructions (inclusion vs. exclusion) as a within-subject variable revealed a significant instruction effect, $F(1,54) = 26.12$, $MSE = 98.22$, $p < .001$, and a marginal instruction by incentive interaction, $F(1,54) = 3.49$, $MSE = 98.22$, $p = .067$. Simple effects of instruction for each incentive condition showed that there was a significant instruction effect for the reward condition, $F(1,54) = 24.35$, $p < .001$, and for the no-reward condition, $F(1,54) = 5.25$, $p < .05$, indicating that overall the number of own triplets was greater in inclusion than exclusion (i.e. I > E).

In order to further explore whether participants can have some level of control over the expression of their knowledge, it is important to determine whether the number of own triplets was above or below that of other triplets (the baseline) in inclusion and exclusion tests. For the inclusion tests, it is clear that the number of own triplets was greater than that of others. This observation was confirmed by a two-way ANOVA with incentive (no-reward vs. reward) as a between-subjects variable and type of triplets (own vs. other) as a within-subject variable. This revealed only a significant triplet type effect, $F(1,54) = 52.43$, $MSE = 212.08$, $p < .001$, suggesting that participants were quite able to demonstrate above-baseline sequence knowledge under inclusion instructions (i.e. I > B). For the exclusion tests, a comparable ANOVA revealed a significant main effect of triplet type, $F(1,54) = 11.09$, $MSE = 94.76$, $p < .01$, and a triplet type by incentive interaction, $F(1,54) = 5.56$, $MSE = 94.76$, $p < .05$. Simple effects of triplet type for each incentive condition showed that there was an effect of triplet type for the no-reward condition, $F(1,\ 54) = 16.18$, $p < .001$, but not for the reward condition, $F(1,54) = .47$, $p = .50$, indicating that participants in the reward group were able to withhold their responses in the exclusion task (i.e. E = B), but participants in the no-reward group could not (i.e. E > B). Simple effects of incentive condition for each triplet type showed that there was a marginally significant effect of incentive for own triplets, $F(1,54) = 3.52$, $MSE = 95.33$, $p = .066$, and for other triplets, $F(1,54) = 3.30$, $MSE = 60.74$, $p = .075$, suggesting that reward made participants generate less own but more other triplets in exclusion.

The above analysis indicated control people had over knowledge that distinguishes the two SOC sequences (own vs. other). Participants may in addition have some level of control over knowledge concerning properties the training and transfer sequences shared. If the latter knowledge was conscious, people would produce greater "neither" triplets under exclusion than inclusion. We computed the number of "neither" triplets by adding up own triplets plus other triplets and subtracting from 96. An ANOVA on "neither" triplets with incentive (no-reward vs. reward) as a between-subjects variable and instructions (inclusion vs. exclusion) as a within-subject variable only revealed a significant instruction effect, $F(1,54) = 8.80$, $MSE = 90.71$, $p < .01$, indicating that in both groups participants generated greater "neither" triplets under exclusion than inclusion.

2.3. Discussion

During the training phase, RTs for the two groups were faster for the regular than transfer blocks indicating that participants learned the second-order conditional structure. Importantly, during the test phase, reward significantly affected exclusion performance. The number of trained triplets generated in the no-reward condition exceeded the baseline level (E > B) under exclusion instructions, conceptually replicating Destrebecqz and Cleeremans' (2001, 2003), while the number of triplets generated for the reward condition did not (E = B), replicating Wilkinson and Shanks' (2004). The number of "neither" triplets was greater under exclusion than inclusion in both groups, indicating that participants consciously learned properties that own and other triplets shared although only participants in the reward group expressed conscious knowledge of the difference between own and other triplets.

The results indicated that reward influences performance on exclusion generation tests. There are two explanations: One is that reward changed the amount of conscious and unconscious knowledge of whether a location can occur next in a context, and the other is that reward changed only the *measurement* of the amount of conscious and unconscious knowledge. According to the former, with no reward, knowledge was unconscious at the time it was applied on the exclusion task but could be made conscious with effort. According to the latter, the knowledge was always conscious, and with no reward, participants simply did not try very hard especially in the difficult exclusion tests, which made the measures insensitive. That is, according to the former explanation, when a person produced an own triplet in exclusion when not rewarded, they were not aware of knowing this was an own triplet; according to the latter explanation, the person was aware of knowing it was an own triplet but did not wish to put effort into producing another triplet.

Consistent with both these explanations, reward significantly affected exclusion performance but not inclusion performance, indicating the difficulty of the exclusion task. If effort is needed to make some knowledge conscious, it would not affect inclusion, which is influenced in the same direction by conscious and unconscious knowledge; so on the explanation that reward affects the *amount* of conscious knowledge (by decreasing unconscious knowledge), the effect of reward on only exclusion is as expected. And, similarly, if reward affects not the amount but only the effort which one puts into acting on the conscious knowledge, reward should affect exclusion rather than inclusion.

An argument for the explanation that reward changed only the report but not the amount of conscious knowledge, is that if people have the competence to control behavior (as demonstrated when they are motivated) then they have awareness (even when unmotivated). In contrast, we view awareness as an actual state rather than a potential one (compare Rosenthal, 2002, 2005). Saying a person has the competence to see a chair by turning their head and looking at it does not mean they actually see it before they turn their head. In turn, we view knowledge as conscious when we are actually conscious of it and not just able to become conscious of it with effort (Rosenthal, 2005). If a mental state's being conscious were just a potential to be conscious of it then it would be meaningless to even ask whether an unconscious state can be made conscious with effort; but clearly it is not meaningless to ask this. In sum, both explanations—reward produces a change in the amount or only in the measurement of conscious knowledge—remain logically possible.

In terms of the former explanation, i.e. that reward induced a change in the actual amount of conscious and unconscious knowledge, there are two possible versions. One is that the same knowledge applied whether or not there was reward, but with reward people were motivated to monitor their knowledge more closely and became more aware of the knowledge they had. The other possible explanation is that there were separate conscious and unconscious knowledge bases of the structure of the sequence and reward induced the person to use more conscious rather than unconscious knowledge so more judgments involved consciously knowing an appropriate next location (see Twyman & Dienes, in press).

If there were knowledge that was actually unconscious at the time it was applied, it should be possible to create conditions whereby even with effort the knowledge would not be made conscious. Making the sequences harder to detect consciously may allow unconscious knowledge to be applied in the exclusion task even with reward. Experiment 2 addressed this issue by manipulating amount of noise.

3. Experiment 2

Like Destrebecqz and Cleeremans' (2001, 2003), Wilkinson and Shanks (2004, experiments 1 and 2), and Norman et al. (2006), Experiment 1 employed a deterministic sequence. That is, each SOC sequence repeated exactly throughout each block. Jiménez (e.g. Cleeremans & Jiménez, 1998; Jiménez, 2003; see also Schvaneveldt & Gomez, 1998) has argued that learning is more likely to be implicit if the sequence is noisy, for example if on only 80% of trials does the sequence-appropriate location come up. Wilkinson and Shanks (2004, experiments 1 and 3) introduced a probabilistic sequence whereby on 85% of trials the appropriate SOC location appeared. Whether the sequence was deterministic or probabilistic, Wilkinson and Shanks did not obtain evidence for exclusion exceeding baseline.

The aim of Experiment 2 was an attempt to investigate whether different amounts of noise would influence performance on the generation tests. A probabilistic sequence was used in the experiment, in which the sequence-appropriate location appeared on either 87.5% or 75% of trials, depending on the condition. All participants were rewarded for good performance as in the reward condition in Experiment 1.

3.1. Method

3.1.1. Participants

Thirty-four undergraduate students (17 male, 17 female) took part in the experiment. None of them had previously taken part in any implicit learning experiment. They were randomly assigned to two groups (.875-probability, $n = 16$; .75-probability, $n = 18$). They were told before the generation tests that in addition to an attendance fee of ¥ 15, they would receive an additional ¥ 50 for good generating performance.

Data from 2 participants, both in the .75-probability group, were excluded because they did not follow the instructions for inclusion or exclusion tests and chose to adopt the single-response strategy 2-3-4 or 1-2-4 for all 96 trials of a generation test. This strategy ensures that performance is unrelated to any conscious or unconscious knowledge the participant has.

3.1.2. Materials

A deterministic SOC sequence can be broken down into 12 sequential chunks of three locations, or triplets (e.g. SOC1 can be broken down into the triplets 3-4-2, 4-2-3, 2-3-1, and so on; and SOC2 can be broken down into 3-4-1, 4-1-2, 1-2-4, and so on). In each triplet, the third location was completely determined by the previous two locations. To generate the probabilistic sequences, we arranged for the corresponding probability to be less than 1.0. That is, every two locations were followed by the corresponding location from the training sequence SOC1 or SOC2 with a probability of .875 or .75, and they were followed by the corresponding location from the other sequence SOC2 or SOC1 with a probability of .125 or .25.

3.1.3. Procedure

3.1.3.1. Training phase. The procedure was identical to that of Experiment 1 except that participants in both groups were trained on 12 training blocks of 98 trails instead of 15 blocks in the training phase.

3.1.3.2. Test phase. The test phase was identical to the reward condition in Experiment 1. One of the 32 participants received the reward; she belonged to the .875-probability group.

3.2. Results

3.2.1. Training phase

Trials with RTs greater than 1000 ms were dropped; these amounted to 0.44% and 0.64% of the trials in the .875- and .75-probability groups. Fig. 3 shows the mean RTs obtained over the training phase. An ANOVA on RTs with probability (probable vs. improbable) and blocks (12 levels) as within-subject variables and amount of noise (.875-probability vs. .75-probability) as a between-subjects variable revealed a significant effect of probability, $F(1,30) = 77.49$, $MSE = 850.78$, $p < .001$, indicating that participants responded to probable locations more quickly than to improbable locations. The probability by noise interaction was also significant, $F(1,30) = 10.01$, $MSE = 850.78$, $p < .01$, indicating a greater probability effect in the .875-probability group than in the .75-probability group. The main effect of block reached significance, $F(11,330) = 3.46$, $MSE = 1182.83$, $p < .001$, and the block by probability interaction was also significant, $F(11,330) = 6.16$, $MSE = 354.35$, $p < .001$, revealing a greater probability effect later in practice than earlier on. The Probability × Noise × Block interaction reached significance, $F(11,330) = 3.51$, $MSE = 354.35$, $p < .001$. The three-way interaction was further analyzed by probability by block partial two-way interactions separately for each

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

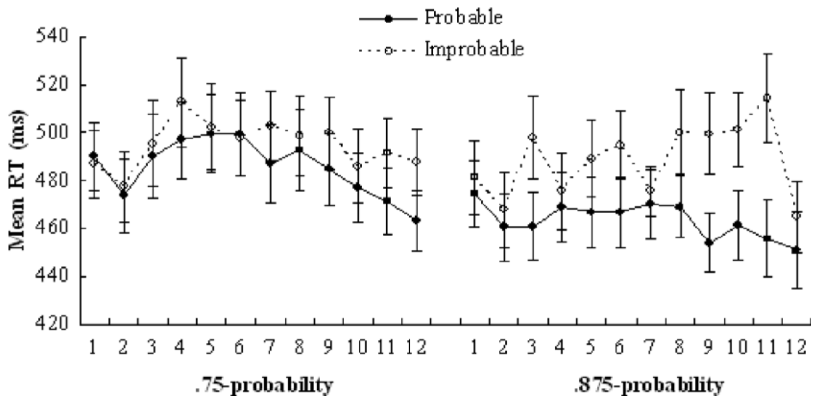

Fig. 3. Mean reaction times (RTs) across training blocks in Experiment 2. The probabilistic data were broken down into probable targets, which were consistent with the training sequence and improbable targets, which were not. Ms, milliseconds. Error bars depict standard errors.

noise condition. The partial interaction was significant for the .875-probability group, $F(11, 165) = 5.05$, $MSE = 497.70$, $p < .001$, and for the .75-probability group, $F(11, 165) = 4.33$, $MSE = 210.99$, $p < .001$. In the .875-probability group, there was no block effect for the probable locations, $F(11, 165) = 1.63$, $MSE = 515.87$, $p = .096$, but there was for the improbable locations, $F(11, 165) = 2.91$, $MSE = 1302.53$, $p < .01$; in the .75-probability group, there was a block effect for both the probable locations, $F(11, 165) = 5.72$, $MSE = 520.23$, $p < .001$, and the improbable, $F(11, 165) = 2.61$, $MSE = 735.72$, $p < .01$. In sum, the .875-probability group compared to the .75-probability group showed a greater difference between probable and improbable items as practice progressed.

Table 2 shows the means for errors. None of the significant RT effects were compromised by possible speed-error trade-offs.

3.2.2. Test data

Fig. 4 shows the mean number of triplets generated for each group in Experiment 2. An ANOVA on own triplets with amount of noise (.875-probability vs. .75-probability) as a between-subjects variable and instructions (inclusion vs. exclusion) as a within-subject variable revealed only a significant instruction effect, $F(1, 30) = 5.78$, $MSE = 101.70$, $p < .05$, indicating that overall the number of own triplets was significantly greater in inclusion than exclusion (i.e. I > E).

Table 2
Mean error proportions for probable and improbable locations across training blocks under the .875- and .75-probability groups in Experiment 2

Block	.875-probability				.75-probabiltiy			
	Probable	*SE*	Improbable	*SE*	Probable	*SE*	Improbable	*SE*
1	.028	.007	.036	.019	.013	.004	.021	.009
2	.027	.004	.021	.012	.023	.007	.034	.012
3	.027	.004	.063	.019	.026	.006	.042	.008
4	.037	.008	.068	.023	.030	.006	.042	.011
5	.043	.007	.073	.020	.028	.006	.034	.008
6	.032	.007	.083	.022	.030	.006	.042	.011
7	.028	.006	.027	.010	.031	.004	.047	.013
8	.037	.008	.073	.023	.033	.006	.047	.009
9	.038	.008	.104	.026	.028	.004	.052	.011
10	.037	.010	.078	.018	.030	.006	.044	.008
11	.029	.007	.135	.024	.030	.006	.052	.013
12	.036	.008	.099	.025	.030	.005	.031	.010

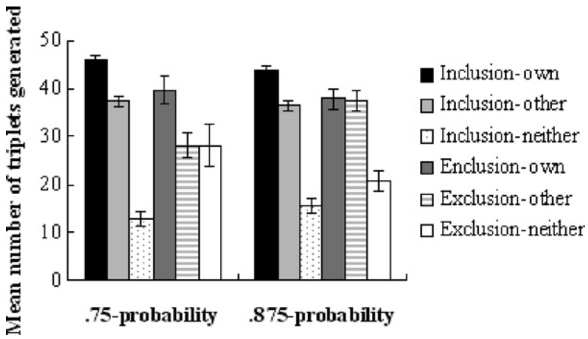

Fig. 4. Mean number of triplets generated by the .875- and .75-probability groups and tested under inclusion and exclusion instructions in Experiment 2. Own, number of SOC triplets generated from the training sequence; other, number of triplets from the untrained sequence; neither, number of triplets from neither the training nor the untrained sequence. Error bars depict standard errors.

For the inclusion tests, an ANOVA with amount of noise (.875-probability vs. .75-probability) as a between-subjects variable and type of triplets (own vs. other) as a within-subject variable only revealed a significant triplet type effect, $F(1,30) = 7.86$, $MSE = 129.23$, $p < .01$, i.e. participants demonstrated above-baseline sequence knowledge during the inclusion task (i.e. I > B). For the exclusion tests, a comparable ANOVA revealed a significant main effect of triplet type, $F(1,30) = 6.49$, $MSE = 89.66$, $p < .05$. The noise by triplet type interaction also reached significance, $F(1,30) = 5.58$, $MSE = 89.66$, $p < .05$. Simple effects of triplet type for each noise condition showed that there was an effect of triplet type for the .75-probability condition, $F(1,30) = 12.06$, $p < .01$, but not for the .875-probability condition, $F(1,30) = .02$, $p = .897$, indicating that participants in the .75-probability group generated above-baseline performance in the exclusion task (i.e. E > B) but participants in the .875-probability group did not (i.e. E = B). Simple effects of noise condition for each triplet type revealed that there was an effect of noise for other triplets, $F(1,30) = 8.30$, $MSE = 83.59$, $p < .01$, but not for own triplets, $F(1,30) = .27$, $MSE = 105.42$, $p = .61$, suggesting that adding more noise made participants generate less other triplets in exclusion.

An ANOVA on "neither" triplets with noise condition (.875-probability vs. .75-probability) as a between-subjects variable and instructions (inclusion vs. exclusion) as a within-subject variable revealed a significant instruction effect, $F(1,30) = 17.05$, $MSE = 97.41$, $p < .001$. The instruction by noise interaction also reached significance, $F(1,30) = 4.31$, $MSE = 97.41$, $p < .05$. Simple effects of instruction for each noise condition showed that there was an effect of instruction for the .75-probability condition, $F(1,30) = 19.26$, $p < .001$, but not for the .875-probability condition, $F(1,30) = 2.10$, $p = .16$, i.e., participants in the .75-probability group generated greater "neither" triplets under exclusion than inclusion but participants in the .875-probability group did not. That is, participants in the .75-probability group relied on conscious knowledge of the common properties of own and other; participants in the .875-probability group relied on conscious knowledge of the distinguishing properties of own and other. Further, there was no main effect of noise for the exclusion tests, $F(1,30) = 2.23$, $MSE = 198.71$, $p = .15$; that is decreasing noise did not increase the amount of conscious knowledge of common properties.

3.3. Discussion

For the two probabilistic groups, RTs were faster for probable than improbable trials, indicating that both groups learned the sequence although the .875-probability group learned more than the .75-probability group. Importantly, the number of triplets generated in exclusion exceeded the baseline level (i.e. E > B) in the .75-probability condition, conceptually replicating the findings of Destrebecqz and Cleeremans (2001, 2003), but did not in the .875-probability condition (i.e. E = B), which was consistent with the findings of Wilkinson and Shanks (2004) and Norman et al. (2006). The .875-probability condition presented only 1 in 8 improbable targets and therefore was in practice very similar to the deterministic sequence of Experiment 1; replication of the results of Experiment 1 with this group is thus reassuring. In contrast, the .75-probability condition involved 2 in 8 improbable targets, which was sufficient noise to deter the formation of conscious knowledge. Participants in the .75-probability group could avoid both own and other triplets together, but participants in the .875 had acquired more refined conscious knowledge, being able to suppress own triplets relative to other.

Making the sequences harder to detect consciously led to unconscious knowledge being demonstrated, despite the fact participants were rewarded for good performance. While reward can allow more conscious knowledge to be expressed with a deterministic sequence, when a pattern is sufficiently non-salient perhaps no amount of focusing attention on one's knowledge states induces them to become conscious. This is also surely true more generally; no amount of intensive introspection is likely to make conscious sensorimotor knowledge (e.g. to reveal one uses tau to determine time-to-contact, Lee, 1980) or linguistic knowledge (or else linguists would have been out of a job a long time ago). The difficulty is setting up such situations in the laboratory, and Experiment 2 suggests that probabilistic SRT tasks may be a useful paradigm for this purpose.

There are two ways of viewing why greater noise resulted in greater expressed unconscious knowledge. There may be separate implicit and explicit learning systems and noise makes it harder for the explicit system to induce reliable knowledge. As a result, there is less conscious knowledge to mask the expression of the unconscious knowledge induced by the implicit system (e.g. Berry & Dienes, 1993; Curran & Keele, 1993; Jiménez & Vázquez, 2005; Willingham & Goedert-Eschmann, 1999). Alternatively, there may be one learning system (learning is just learning) but knowledge that is weak or of poor quality cannot support awareness of itself even though it is strong enough to influence behavior (e.g. Cleeremans & Jiménez, 2002; Destrebecqz & Cleeremans, 2001, 2003; Whittlesea & Dorken, 1993). A natural consequence of the latter position is that unconscious knowledge should be especially detectable early rather than later in training. Experiment 3 was designed to explore this argument.

4. Experiment 3

Unconscious knowledge might emerge only late in training, because a long training period might allow participants the opportunity to develop more automatic knowledge (e.g. Anderson, 1983; Shanks, Rowland, & Ranger, 2005; Wilkinson & Shanks, 2004). Relatedly, Perruchet, Bigand, and Benoit-Gonin (1997) argued that conscious knowledge emerged early in training on the SRT task; however, they used objective measures of conscious knowledge and so there is little reason to believe they were measuring unconscious knowledge, rather than simply knowledge. A second view is that unconscious knowledge is characterized by weak, poor-quality representations, likely to arise early in training (Cleeremans & Jiménez, 2002). A third view is that conscious and unconscious knowledge develops in parallel (Willingham et al., 1989). Experiment 3 was an attempt to explore the amount of unconscious knowledge that can be detected after either 15 blocks or only 6 blocks of training (contrast the 12 blocks of the previous experiment) with the .875-probability sequence. All participants were rewarded for good performance as in the reward condition in Experiment 1.

4.1. Method

4.1.1. Participants

Forty-eight undergraduate students (22 male, 26 female) took part in the experiment. None of them had previously taken part in any implicit sequence learning experiment. They were randomly assigned to two groups (6-block, $n = 24$; 15-block, $n = 24$). They were told before the generation tests that in addition to an attendance fee of ¥ 15, they would receive an additional ¥ 50 for good generation performance.

4.1.2. Procedure

4.1.2.1. Training phase. The procedure was identical to the .875-probability group of Experiment 2 except that each condition was composed of 6 or 15 training blocks, respectively (cf. 12 in Experiment 2). Each block consisted of 98 trials, for a total of 588 and 1470 trials depending on the condition.

4.1.2.2. Test phase. The test phase was identical to the reward condition in Experiment 1. Three of the 48 participants received the reward, all belonging to the 15-block group.

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

4.2. Results

4.2.1. Training data

Trials with RTs greater than 1000 ms were dropped; these amounted to 0.57% and 0.96% of the trials in the 6- and 15-block groups, respectively. Fig. 5 shows the mean RTs obtained over the training phase. For the 6-block group, an ANOVA on RTs with probability (probable vs. improbable) and blocks (6 levels) as within-subject variables revealed a significant effect of probability, $F(1,23) = 28.01$, $MSE = 659.41$, $p < .001$, suggesting that participants responded to the probable locations more quickly than to the improbable. The main effect of block was also significant, $F(5,115) = 3.46$, $MSE = 907.15$, $p < .01$, indicating that participants responded the targets more quickly later in practice than earlier on. For the 15-block group, a comparable ANOVA revealed a significant effect of probability, $F(1,23) = 68.10$, $MSE = 1871.57$, $p < .001$, suggesting that participants responded to the probable locations more quickly than to the improbable. The main effect of block was significant, $F(14,322) = 7.24$, $MSE = 1275.99$, $p < .001$, and the probability by block interaction also reached significance, $F(14,322) = 4.14$, $MSE = 587.82$, $p < .001$, indicating a greater probability effect later in practice than earlier on.

In order to compare the two training conditions, an ANOVA on the last three blocks of each training condition with training (6-block vs. 15-block) as a between-subjects variable, probability (probable vs. improbable) and blocks (3 levels) as within-subject variables was used. This revealed a significant probability effect, $F(1,46) = 96.52$, $MSE = 729.52$, $p < .001$, and a probability by training interaction, $F(1,46) = 10.51$, $MSE = 729.52$, $p < .01$, suggesting that a greater probability effect in the 15-block group than in the 6-block group. The pattern of error data mirrored that of the RT data, as shown in Table 3.

4.2.2. Test data

Fig. 6 shows the mean number of triplets generated for each group in Experiment 3. An ANOVA on own triplets with training (6 blocks vs. 15 blocks) as a between-subjects variable and instructions (inclusion vs. exclusion) as a within-subject variable revealed a significant instruction effect, $F(1,46) = 18.78$, $MSE = 87.85$, $p < .001$, and the training by instruction interaction was also significant, $F(1,46) = 12.30$, $MSE = 87.85$, $p = .001$. Simple effects of instruction for each training condition showed there was effect of instruction for the 15-block group, $F(1,46) = 30.74$, $p < .001$, but not for the 6-block group, $F(1,46) = .34$, $p = .56$. The results revealed that participants of 15-block groups could generate more own triplets in inclusion than exclusion (i.e. I > E), but participants of 6-block groups could not (i.e. I = E).

For the inclusion tests, an ANOVA with training (6 blocks vs. 15 blocks) as a between-subjects variable and type of triplets (own vs. other) as a within-subject variable revealed only a significant triplet type effect, $F(1,46) = 17.76$, $MSE = 200.75$, $p < .001$, suggesting that participants were quite able to demonstrate above-baseline sequence knowledge during the inclusion instructions (i.e. I > B). For the exclusion tests, a

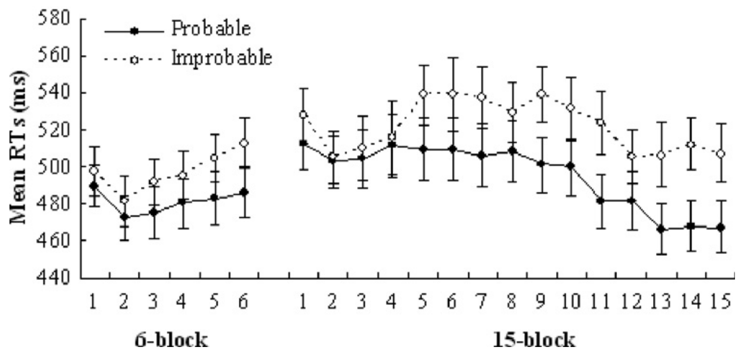

Fig. 5. Mean reaction times (RTs) across training blocks in Experiment 3. The probabilistic data were broken down into probable targets, which were consistent with the training sequence and improbable targets, which were not. Ms, milliseconds. Error bars depict standard errors.

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

Table 3
Mean error proportions for probable and improbable locations across training blocks under the 15- and 6- block groups in Experiment 3

Block	15-block				6-block			
	Probable	*SE*	Improbable	*SE*	Probable	*SE*	Improbable	*SE*
1	.022	.005	.028	.010	.016	.005	.028	.011
2	.025	.006	.038	.012	.028	.006	.021	.008
3	.028	.007	.052	.014	.026	.005	.045	.012
4	.026	.005	.073	.015	.024	.005	.038	.013
5	.027	.005	.042	.015	.032	.006	.059	.015
6	.029	.006	.080	.018	.029	.006	.052	.014
7	.029	.006	.069	.017				
8	.031	.006	.094	.020				
9	.030	.005	.056	.016				
10	.034	.005	.076	.017				
11	.021	.005	.080	.020				
12	.026	.005	.097	.020				
13	.028	.005	.101	.020				
14	.030	.006	.101	.019				
15	.028	.005	.069	.018				

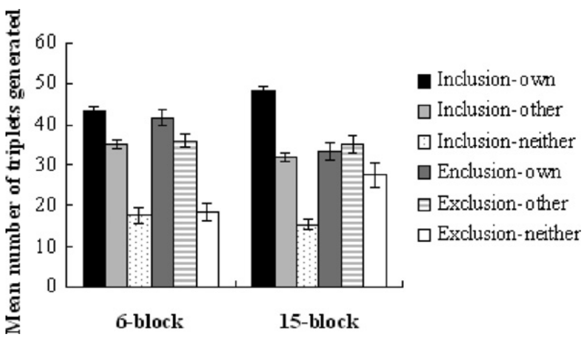

Fig. 6. Mean number of triplets generated by the 15- and 6-block groups and tested under inclusion and exclusion instructions in Experiment 3. Own, number of SOC triplets generated from the training sequence; other, number of triplets from the untrained sequence; neither, number of triplets from neither the training nor the untrained sequence. Error bars depict standard errors.

comparable ANOVA revealed a significant main effect of training, $F(1,46) = 6.41$, $MSE = 79.43$, $p < .05$, and a triplet type by training interaction, $F(1,46) = 4.21$, $MSE = 81.01$, $p < .05$. Simple effects of type for each training condition revealed that there was an effect of type for the 6-block group, $F(1,46) = 4.90$, $p < .05$, but not for the 15-block group, $F(1,46) = .48$, $p = .49$, indicating that participants of the 15-block group were quite able to withhold their responses under exclusion instructions (i.e. E = B), but participants of the 6-block group could not (i.e. E > B). Simple effects of training condition for each triplet type revealed a significant training effect for own triplets, $F(1,46) = 9.15$, $MSE = 92.01$, $p < .01$, but not for other triplets, $F(1,46) = .12$, $MSE = 68.43$, $p = .73$, suggesting that training enabled participants to withhold more own triplets in exclusion.

An ANOVA on "neither" triplets with training condition (6-block vs. 15-block) as a between-subjects variable and instructions (inclusion vs. exclusion) as a within-subject variable revealed a significant instruction effect, $F(1,46) = 11.42$, $MSE = 85.43$, $p = .001$. The instruction by training interaction also reached significance, $F(1,46) = 9.02$, $MSE = 85.43$, $p < .01$. Simple effects of instruction for each training condition showed that there was an effect of instruction for the 15-block group, $F(1,46) = 20.37$, $p < .001$, but not for the 6-block group, $F(1,46) = .07$, $p = .79$, suggesting that participants in the 15-block group generated greater "neither" triplets under exclusion than inclusion but participants in the 6-block group did not. Further, training increased the amount of "neither" triplets in exclusion, $F(1,46) = 6.41$, $MSE = 158.85$, $p < .05$, indicating training increased the amount of conscious relative to unconscious knowledge of common properties.

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

4.3. Discussion

For both the 15- and 6-block groups, RTs were faster for probable than improbable target locations, confirming that participants acquired the structure of the sequence. After training, the 15-block group generated more own under inclusion than exclusion (i.e. $I > E$), and generated no more own than other triplets under exclusion (i.e. $E = B$), indicating that participants acquired at least some conscious knowledge. However, the 6-block group could not generate more own triplets in inclusion than exclusion (i.e. $I = E$), and they did generate more own than other triplets under exclusion (i.e. $E > B$), suggesting that what participants acquired was unconscious. The 6-block group also did not generate greater "neither" triplets under exclusion than inclusion, revealing that they did not acquire conscious knowledge of even the common properties of own and other triplets. The results indicated that unconscious knowledge was detectable given a shorter rather than longer period of training.

This is as predicted by the view of Cleeremans and Jiménez (2002), that unconscious knowledge has weak, poor-quality representations that are already capable of influencing performance but they are too weak for the system as a whole to exert control over them. Cleeremans and Jiménez argue that implicit knowledge starts in the context of a single trial, or at an early stage of development of skill acquisition. The results are also consistent with multiple systems views (e.g. Berry & Dienes, 1993; Curran & Keele, 1993; Jiménez & Vázquez, 2005; Willingham & Goedert-Eschmann, 1999), but are not specifically predicted by them. The multiple systems theory claims there are implicit and explicit learning system but it does not explain whether implicit knowledge appears specifically early or late in training. The results are inconsistent with the view that implicit knowledge emerges only late in training (Anderson, 1983; Shanks et al., 2005; Wilkinson & Shanks, 2004).

Notice a slight difference in pattern between Experiments 2 and 3: while both noise and training affect the difference between own and other triplets generated in exclusion, noise appears to have decreased the amount of other triplets generated in exclusion but training seems to have reduced the number of own triplets and increased the number of "neither" triplets generated in exclusion (compare Figs. 4 and 6). A person's knowledge can be partitioned into (a) knowledge relevant to distinguishing own and other triplets and (b) knowledge concerning properties both SOC sequences have in common. Consequently, the effect of noise or training can be similarly partitioned into two: the difference between own and other; and the sum of own and other. In exclusion, unconscious knowledge of (a) increases the difference; similarly unconscious knowledge of (b) increases the sum of own and other (and conscious knowledge decreases it). The difference changes in both experiments between the groups in much the same way; the sum changes differently. In Experiment 2, the sum of own and other changes little between the probability groups; in Experiment 3, training reduces the sum. Therefore, training increases the balance of conscious over unconscious knowledge of the SOC structure both SOC sequence share in common; less noise does not. Both increase the conscious knowledge of the difference between own and other.

5. General discussion

The aim of the experiments presented here was to investigate whether participants could acquire implicit knowledge in sequence learning. Evidence of implicit knowledge comes from either finding $I = E$ or $E > B$, where I and E refer to the number of chunks from their trained sequence generated in the inclusion and exclusion tasks, and B refers to an appropriate baseline level of generation. Our results provided evidence for implicit knowledge in that after training people could generate legal sequences above-baseline levels under exclusion instructions. Above-baseline exclusion performance was affected by reward, with reward moving performance towards baseline. Even with reward, adding noise to the sequences led to above-baseline exclusion performance. Further, such above-baseline performance was more likely after short rather than long training, which can be predicted by theories that postulate unconscious knowledge can arise when underlying representations are of sufficiently poor quality they cannot support awareness of them, but they are sufficiently good to have some behavioral consequences (Cleeremans & Jiménez, 2002).

5.1. PDP and subjective measures

Before PDP was used with the SRT task, investigators had primarily used objective measures of conscious awareness, for example the ability to generate a training sequence. Destrebecqz and Cleeremans (2001, 2003), and our replication of them, highlight that generation per se can be based on unconscious knowledge, a fact particularly evident when generation happens under exclusion instructions.

The logic of PDP is based on the assumption that conscious knowledge allows control. Whereas Jacoby takes control to be definitional of whether knowledge is conscious or not (compare Higham, Vokey, & Pritchard, 2000), PDP can be taken as a type of subjective measure (Jack & Shallice, 2001; see also Snodgrass, 2002), where control is relevant only when it is based on awareness of knowing rather than guessing. For example, on this account, the ability to refrain from completing a stem in a subliminal perception experiment reflects conscious knowledge only to the extent that the person refrains from completing the stem because *they thought they saw the word*. Refraining from completing the stem because a word was—it appeared to the subject—randomly guessed, would not count as evidence of having consciously seen the word, regardless of the fact control was objectively exerted. In fact, it would seem strange to count such control as an indication of conscious knowledge if the person sincerely denied seeing the word at all. Exclusion without conscious awareness may be rare but it can happen; for example, Dienes and Perner (2007) argue that a key feature of hypnosis is performance of executive function tasks, like exclusion, without conscious awareness of what is excluded.

Now consider the cases where we found above-baseline performance in the SRT task in the exclusion condition. When participants think of a response, their procedural knowledge will sometimes tend to make them think of legal responses. If they feel they are sure the response is legal they would withhold that response. But if they feel they are guessing about whether it is legal or not (they don't know that they know it is legal) they press the key for that response. Thus, where $E > B$, we can infer participants had knowledge they did not know they had.

This analysis indicates why the PDP results are consistent with previous results using subjective rather than objective measures (e.g.Willingham & Goedert-Eschmann, 1999): PDP is a type of subjective measure, and hence must engage with the same conceptual issues raised by any subjective measure. Deciding whether to withhold a response depends on a person's assessment of whether she knows, and thus on how the person defines knowing. We see this not as a shortcoming of PDP but as a strength: It must confront the issues at the core of the problem of conscious awareness (Rosenthal, 2005).

Dienes and Scott (2005) pointed out that there are two types of knowledge relevant in the SRT and other implicit learning tasks: structural knowledge and judgment knowledge (for a related theoretical framework, see Norman et al., 2006). For example, in natural language a person may not have conscious knowledge of the structure of the language but may consciously know that a particular sentence or word in a sentence is grammatical or not. In terms of the SRT task, structural knowledge is knowledge about the structure of the sequences in the training phase, and in principle may consist of knowledge of fragments, the whole sequence, of abstract patterns, conditional probabilities, and so on. Judgment knowledge is knowledge of whether a particular location is legal given the context. The judgment is based on the use of structural knowledge in the presented context. If structural knowledge is unconscious, judgment knowledge can be either conscious or unconscious. However, if structural knowledge is conscious, judgment knowledge will be conscious. PDP measures the conscious status of judgment knowledge, not structural knowledge. In generating a sequence under exclusion, a person need only know that this location can or cannot occur at this point; there is no need for them to have any idea why that should be so. The above-baseline exclusion performance we have found so far is evidence for unconscious judgment knowledge. Thus, a person generating above baseline would of course not know that they were doing so. When exclusion performance is at (or below) baseline, it is completely open whether structural knowledge is conscious or unconscious. For example, a native speaker of a language can refrain from producing a grammatical completion to a sentence precisely because it is their judgment knowledge that is conscious; they need not have conscious structural knowledge (and typically do not; hence the continued viability of computational linguistics as a discipline). Future research needs to explore the conscious and unconscious status of structural knowledge in the SRT task when judgment knowledge is conscious (as shown by successful exclusion, as in, for example, the reward condition of experiment 1).

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007

5.2. Single vs. dual systems

Our results help reconcile the apparently contradictory empirical findings of Wilkinson and Shanks (2004) with those of Destrebecqz and Cleeremans (2001, 2003). Further, despite the rhetoric, their theoretical views can be reconciled as well, as we now indicate. Both Destrebecqz and Cleeremans (2001, 2003) and Shanks et al. (2003) provided a single system account of learning on the SRT task. Destrebecqz and Cleeremans adopt a connectionist approach. They argue as learning progresses the representations formed become more stable and of better quality. There is a period when the representations are good enough to influence performance but not to become clearly conscious. So, forced-choice recognition performance could be above baseline, because participants can choose even when they believe they are guessing. The latter point was not argued for by Destrebecqz and Cleeremans, though above-chance recognition was empirically confirmed by Shanks et al. (2003) and Norman et al. (2006).

In fact, Shanks et al. (2003) provide an entirely complimentary set of ideas. They suggest there is one source of knowledge, familiarity, and different measures access this one source with independent errors. For example, reaction time is a function of familiarity plus some random error. Judgments about whether a sequence is legal or old are also based on familiarity plus some random error, independent of the RT error term. Because errors are independent, on some trials a legal continuation will be generated based on the knowledge but not consciously judged as legal. These trials will contribute to above-baseline exclusion performance and prima facie constitute cases of unconscious knowledge. The trials where conscious judgment is accurate and consistent with generation will contribute to below chance exclusion performance. Whether performance on the exclusion task is overall above baseline or not will depend simply on the relative size of the error terms for generation and judgment.

We could regard the familiarity of Shanks et al. as corresponding with the quality of representations invoked by Destrebecqz and Cleeremans. In that case, when familiarity is small, the error for judgments needs to be large relative to the error for generation for Shanks et al's account to match that of Destrebecqz and Cleeremans. Then the probability of making accurate conscious judgments is small, but knowledge can be reliably expressed in generation, as in the early stages of learning, consistent with Experiment 3.

On the Destrebecqz and Cleeremans account, low quality representations cannot support awareness of their contents; so judgment knowledge is unconscious. Similarly, on the Shanks et al. account familiarity can lead to accurate generations without reliable conscious judgment knowledge, depending on the relative error terms. On neither account is there any reason to suppose structural knowledge is ever conscious. For Destrebecqz and Cleeremans, structural knowledge is embedded in the weights of a connectionist network, and these representations remain unconscious. Similarly, Shanks et al. do not explicitly claim that the means by which familiarity is determined need be conscious. Familiarity is just a unidimensional variable. Shanks (e.g. Shanks, Johnstone, & Kinder, 2002) has worked on connectionist models of other implicit learning paradigms, and consistently the model of Shanks et al. (2003) is perfectly compatible with structural knowledge being embedded in weights. It is a natural step to conclude such structural knowledge is unconscious. Similarly, Norman et al. (2006) postulated that a sense of familiarity or rightness (which they called "fringe consciousness") can guide generation or recognition without people consciously being aware of the underlying structural knowledge.

In the accounts of both Destrebecqz and Cleeremans (2001) and of Shanks et al. (2003), the type of learning involved can be contrasted with hypothesis testing and episodic memory. Here, structural knowledge is conscious. Learning relies on different principles when structural knowledge is conscious: We can *consciously* consider and test any hypothesis about structure that can be imagined. The types of connectionist networks used by Cleeremans, Destrebecqz, and Shanks (e.g. Cleeremans & McClelland, 1991; Destrebecqz & Cleeremans, 2003; Kinder & Shanks, 2001, 2003; Shanks et al., 2003) are more constrained in what they learn, and they don't consider *possibilities*—they just try to match how the environment IS, as best as their architectural presuppositions allow (this also applies to the chunk models of e.g. Boucher & Dienes, 2003; Perruchet & Peereman, 2004). This is a fundamental difference between two types of learning (Dienes & Perner, 1999). It is difficult to see how pure single system theories can capture the different representational resources needed for simply updating one's model of the one real world on the one hand, and considering possible and counterfactual worlds on the other (Evans & Over, 1996; Perner, 1991; see also Hazeltine & Ivry, 2003, for evidence there are different neural regions supporting learning in different contexts encouraging implicit or explicit learning). Nonetheless, as a methodological heuristic, it is useful to consider how far a device like a connectionist network can take us in explaining learning and the generation of both conscious and unconscious judgment knowledge.

Acknowledgments

This research was supported in part by grants from 973 Program of Chinese Ministry of Science and Technology (Grant # 2006CB303101) and the National Natural Science Foundation of China (Grant # 0433030). The research was in partial fulfillment of the requirements of a PhD for Qiufang Fu.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

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

Boucher, L., & Dienes, Z. (2003). Two ways of learning associations. *Cognitive Science, 27*, 807–842.

Buchner, A., Steffens, M. C., Erdfelder, E., & Rothkegel, R. (1997). A multinomial model to assess fluency and recollection in a sequence learning task. *The Quarterly Journal of Experimental Psychology, 50A*, 631–663.

Buchner, A., Steffens, M. C., & Rothkegel, R. (1998). On the role of fragment knowledge in a sequence learning task. *The Quarterly Journal of Experimental Psychology, 51A*, 251–281.

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

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

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 189–202.

Debner, J. A., & Jacoby, L. L. (1994). Unconscious perception: attention, awareness, and control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 304–317.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the Process Dissociation Procedure. *Psychonomic Bulletin & Review, 8*, 343–350.

Destrebecqz, A., & Cleeremans, A. (2003). Temporal effects in sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 181–213). Amsterdam: John Benjamins Publishing Company.

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Fiore, G. D., Aerts, J., et al. (2005). The neural correlates of implicit and explicit sequence learning: interacting networks revealed by the process dissociation procedure. *Learning and Memory, 12*, 480–490.

Dienes, Z. (2004). Assumptions of subjective measures of unconscious mental states: higher order thoughts and bias. *Journal of Consciousness Studies, 11*, 25–45.

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1322–1338.

Dienes, Z., & Berry, D. (1997). Implicit learning: below the subjective threshold. *Psychonomic Bulletin & Review, 4*, 3–23.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioural and Brain Sciences, 22*, 735–755.

Dienes, Z., & Perner, J. (2004). Assumptions of a subjective measure of consciousness: three mappings. In R. Gennaro (Ed.), *Higher order theories of consciousness* (pp. 173–199). Amsterdam: John Benjamins Publishers.

Dienes, Z., & Perner, J. (2007). The cold control theory of hypnosis. In G. Jamieson (Ed.), *Hypnosis and conscious states: The cognitive neuroscience perspective* (pp. 293–314). Oxford University Press.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: distinguishing structural knowledge and judgment knowledge. *Psychological Research, 69*, 338–351.

Dulany, D. E. (1997). Consciousness in the explicit (deliberative) and implicit (evocative). In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 179–211). Mahwah, NJ: Erlbaum.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove: Psychology Press.

Goschke, T. (1998). Implicit learning of perceptual and motor sequences: evidence for independent learning systems. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 401–444). Thousand Oaks, CA: Sage.

Hazeltine, E., & Ivry, R. B. (2003). Neural structures that support implicit sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 71–107). Amsterdam: John Benjamins Publishing Company.

Higham, P. A., Vokey, J. R., & Pritchard, J. L. (2000). Beyond dissociation logic: evidence for controlled and automatic influences in artificial grammar learning. *Journal of Experimental Psychology: General, 129*, 457–470.

Hoffman, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 470–482.

Jack, A. I., & Shallice, T. (2001). Introspective physicalism as an approach to the science of consciousness. *Cognition, 79*, 161–196.

Jacoby, L. L. (1991). A process dissociation framework: separating automatic from intentional use of memory. *Journal of Memory and Language, 30*, 513–541.

Jiménez, L. (2003). Intention, attention, and consciousness in probabilistic sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 43–68). Amsterdam: John Benjamins Publishing Company.

Jiménez, L., & Vázquez, G. A. (2005). Sequence learning under dual-task conditions: alternatives to a resource-based account. *Psychological Research, 69*, 352–368.

Kane, K. A., Picton, T. W., Moscovitch, M., & Winocur, G. (2000). Event-related potentials during conscious and automatic memory retrieval. *Cognitive Brain Research, 10*, 19–35.

Kinder, A., & Shanks, D. R. (2001). Amnesia and the declarative/procedural distinction: a recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience, 13*, 648–669.

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: a single-system connectionist account. *Psychological Review, 110*, 728–744.

Lee, D. (1980). A theory of visual control of braking based on information about time-to collision. *Perception, 5*, 437–459.

McBride, D. M., & Dosher, B. A. (1999). Forgetting rates are comparable in conscious and automatic memory: a process-dissociation study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 583–607.

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

Norman, E., Price, M. C., & Duff, S. C. (2006). Fringe consciousness in sequence learning: the evidence of individual differences. *Consciousness and Cognition, 15*, 723–760.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

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., Bigand, E., & Benoit-Gonin, F. (1997). The emergence of explicit knowledge during the early phase of learning in serial reaction time tasks. *Psychological Research, 60*, 4–13.

Perruchet, P., & Peereman, R. (2004). The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics, 17*, 97–119.

Perruchet, P., & Viner, A. (2002). The self-organizing consciousness: a framework for implicit learning. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness* (pp. 41–67). Hove, UK: Psychology Press.

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.

Reingold, E. M. (1995). Facilitation and interference in indirect/implicit memory tests and in the process dissociation paradigm: The letter insertion and the letter deletion tasks. *Consciousness and Cognition, 4*, 459–482.

Rosenthal, D. M. (2002). *Consciousness and higher order thought*. *Encyclopedia of cognitive science*. Hampshire: Macmillan, pp. 717-726.

Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford: Oxford University Press.

Schvaneveldt, R. W., & Gomez, R. L. (1998). Attention and probabilistic sequence learning. *Psychological Research, 61*, 175–190.

Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. Goldstone (Eds.), *Handbook of cognition* (pp. 202–220). London: Sage.

Shanks, D. R., Johnstone, T., & Kinder, A. (2002). Modularity and artificial grammar learning. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness* (pp. 93–120). Hove, UK: Psychology Press.

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research, 69*, 369–382.

Shanks, D. R., Wilkinson, L., & Channon, S. (2003). Relationship between priming and recognition in deterministic and probabilistic sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 248–261.

Snodgrass, M. (2002). Disambiguating conscious and unconscious influences: do exclusion paradigms demonstrate unconscious perception? *American Journal of Psychology, 115*, 545–579.

Twyman, M., & Dienes, Z. (in press). Are subjective measures biased? *Consciousness and Cognition*.

Visser, T., & Merikle, P. M. (1999). Conscious and unconscious processes: the effects of motivation. *Consciousness and Cognition, 8*, 94–113.

Whittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: an episodic processing account of implicit learning. *Journal of Experimental Psychology: General, 122*, 227–248.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 354–369.

Willingham, D. B., & Goedert-Eschmann, K. (1999). The relation between implicit and explicit learning: evidence for parallel development. *Psychological Science, 10*, 531–534.

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

Willingham, D. B., Wells, L. A., & Farrell, J. M. (2000). Implicit motor sequence learning is represented in response locations. *Memory & Cognition, 28*, 366–375.

Ziessler, M. (1998). Response-effect learning as a major component of implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 962–978.

Please cite this article in press as: Fu, Q. et al., Implicit sequence learning and conscious awareness, *Consciousness and Cognition* (2007), doi:10.1016/j.concog.2007.01.007