Implicit sequence learning of chunking and abstract structures

Qiufang Fu[1, 2] Huiming Sun[3] Zoltán Dienes[4] Xiaolan Fu[1, 2]

[1]State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, China

[2]Department of Psychology, University of Chinese Academy of Sciences, China

[3]Institute of Politics, NDU, PLA, China

[4]School of Psychology and Sackler Centre for Consciousness Science, University of Sussex, UK

Address for correspondence:

Qiufang Fu Ph. D

Institute of Psychology, Chinese Academy of Sciences

16 Liucui Road, Chaoyang District, Beijing 100101, China

Tel: (86 10) 6484-5395

Fax: (86 10) 6487-2070

E-mail: fuqf@psych.ac.cn

Abstract

The current study investigated whether people can simultaneously acquire knowledge about concrete chunks and abstract structures in implicit sequence learning; and whether the degree of abstraction determines the conscious status of the acquired knowledge. We adopted three types of stimuli in a serial reaction time task in three experiments. The RT results indicated that people could simultaneously acquire knowledge about concrete chunks and abstract structures of the temporal sequence. Generation performance revealed that ability to control was mainly based on abstract structures rather than concrete chunks. Moreover, ability to control was not generally accompanied with awareness of knowing or knowledge, as measured by confidence ratings and attribution tests, confirming that people could control the use of unconscious knowledge of abstract structures. The results present a challenge to computational models and theories of implicit learning.

Key words: implicit learning, sequence learning, concrete chunks, abstract structures

Introduction

Although implicit learning has been defined as the acquisition of unconscious complex knowledge (Jiménez, 2003; Reber, 1989; Seger, 1994), it remains controversial whether people acquire knowledge about concrete exemplars (i.e., chunks or fragments) or abstract structures (i.e., other rules or regularities) in implicit learning (e.g., Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998; Goschke & Bolte, 2007). The abstraction of general rules from direct experiences allows for the flexibility and adaptability that are central to intelligent behavior (Wallis, Anderson, & Miller, 2001). The degree of abstraction that can be implicitly learnt has important consequence for computational models and theories of implicit learning.

Reber (1989), one of the founders of implicit learning research, argued that implicit learning is characterized by two critical features: It results in knowledge that is largely (1) unconscious; (2) abstract. The initial empirical evidence in support of this assumption stemmed primarily from transfer effects in artificial grammar learning. For example, in an artificial grammar learning (AGL) task, participants are exposed to a set of letter strings that are generated by a finite-state grammar in the training phase; when participants are presented with novel letter strings or novel tone sequences in the test phase, they can implicitly transfer or apply the grammatical knowledge to test sequences constructed out of the same or indeed a new vocabulary (e.g., Altmann, Dienes, & Goode, 1995; see Reber, 1989 for a review). Despite abundant evidence for people acquiring structural knowledge of artificial grammars, the interpretation of the learning as implicit or abstract has been questioned widely over the last twenty years (e.g., Dulany, 1997; Perruchet & Vinter, 2002; Shanks, 2004). For example, Perruchet and Pacteau (1990) demonstrated that classifying new letter strings as grammatical or ungrammatical may depend on fragmentary knowledge of the bigrams of the training letter strings rather than an unconscious abstract representation of the grammar. Gomez (1997) argued that simple learning, such as learning first-order dependencies (bigrams), could occur without awareness, but more complex learning, such as learning second-order dependencies, was linked to explicit knowledge. The debate regarding what is learned implicitly is far from resolved (contrast Dienes, 2012; Vadillo, Konstantinidis, & Shanks, 2016).

Many recent studies in implicit sequence learning have focused on whether people can implicitly acquire complex knowledge such as second-order conditional (SOC) structure, by adopting SOC sequences in a serial reaction time (SRT) task (e.g., Destrebecqz & Cleeremans, 2001, 2003; Fu, Fu, & Dienes, 2008; Norman, Price, & Duff, 2006; Norman, Price, Duff, & Mentzoni, 2007; Pronk & Visser, 2010; Wilkinson & Shanks, 2004). In the SRT task, participants are asked to respond to the target at one of four locations as accurately and as quickly as possible. Unbeknownst to participants, the stimuli may follow, for example, a SOC sequence. It has been demonstrated that people can implicitly acquire sequence knowledge about the SOC structure when the response-stimulus interval (RSI) is zero (Destrebecqz & Cleeremans, 2001; Fu, Fu, & Dienes, 2008; see Wilkinson & Shanks, 2004 for inconsistent findings). These findings provided important evidence that people can unconsciously acquire complex knowledge such as second-order dependencies. Nonetheless, in the SRT task, the learning effect is mainly defined as shorter reaction times for the training sequence compared to the transfer or random sequence, which depends only on people acquiring concrete chunks or triplets of the training sequence rather than abstract structures. By contrast, in the AGL task, there is evidence of learning not only chunks or associations, but also relations that go beyond chunks or associations, namely patterns of repetitions independent of vocabulary (e.g. Brooks & Vokey, 1991; Tunney & Altmann, 2001) or symmetries (e.g. Ling, Li, Qiao, Guo & Dienes, 2016) or other supra-finite state structures (e.g., Rohrmeier, Fu, & Dienes, 2012).

To address whether people can acquire structure more abstract than memorized chunks in implicit sequence learning, Goschke and Bolte (2007) developed a new serial naming task (SNT), in which participants were asked to name line-drawings of concrete objects from one of four semantic categories. Unbeknownst to participants, the concrete objects were presented in a random order, but the sequence of semantic categories followed a repeating sequence (e.g., furniture–body part–animal–clothing–body part–animal). They found that the reaction times in the SNT were much faster for a repeating category sequence than a random category sequence but performance in a sequence reproduction task was not significantly greater than chance level, which was taken to indicate that people implicitly acquired knowledge about the repeating category sequence. As the acquired knowledge referred to sequential dependencies between semantic categories rather than specific exemplars, it was abstract in this sense (Compare Rebuschat & Williams, 2009, finding implicit learning of the order of grammatical type of word independent of the exact words used in an AGL paradigm). However, Dominey, Lelecov, Ventre-Dominey, & Jeannerod (1998) investigated learning of abstract repetition structure in the SRT task and found that participants in the implicit group did not significantly learn the abstract structure. Nonetheless, in the Experiments 2 and 3 of Dominey et al. (1998), participants in the implicit learning condition also showed significant or marginally significant learning effects of the abstract structures, although it was argued that this abstract learning effect was due to single-item recency effects. Other researchers have also argued that abstract knowledge can be acquired only in explicit learning conditions (Boyer, Destrebecqz, & Cleeremans, 2005; Channon et al., 2002; Cleeremans & Destrebecqz, 2005; Johnston & Shanks, 2001).

Fu, Fu, and Dienes (2008) adopted two second-order conditional (SOC) 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) in the training phase, in which one SOC sequence is the training sequence and its triplets occurred with a large probability and the other SOC sequence is the transfer sequence and its triplets occurred with a small probability. After the training, people were asked to complete two free generation tests according to the logic of the Process Dissociation Procedure (PDP, Jacoby, 1991; for bias of the PDP measure see Stahl, Barth, & Haider, 2015): in an inclusion test, participants were asked to generate a sequence that was same as the training sequence; in an exclusion test, participants were asked to generate a sequence that was different from the training sequence. They found that two types of knowledge were expressed in the explicit tests: 1) knowledge relevant to distinguishing training and transfer SOC sequences, i.e., chunking knowledge about concrete chunks or triplets; 2) knowledge concerning properties both training and transfer SOC sequences had in common, i.e., abstract structures about repetition patterns. They also found that the amount of noise and training influenced the conscious status of chunking knowledge and abstract knowledge in a different way, indicating that people can simultaneously acquire chunking and abstract knowledge in implicit sequence learning.

The abstract feature shared by training and transfer SOC sequences in Fu et al. (2008) is reversal frequency (Pronk & Visser, 2010). A reversal refers to a triplet in which the first and the third stimulus were the same, as found in ABA grammars (Marcus, 1999) or n-2 repetition (Koch, Philipp, & Gade, 2006). Reed and Johnson (1994) considered reversals as salient, and each of the SOC sequence (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) has only one reversal. That is, there is one reversal triplet in the training or transfer SOC sequence, while there are ten reversal triplets in the neither SOC sequence. To investigate the effects of reversal frequencies in probabilistic SOC sequence learning, Pronk and Visser (2010) trained one group of participants with the sequence that contained only a single reversal and one group of participants with the sequence that contained four reversals. They found that the reversal frequency in probabilistic SOC sequence learning influenced how people responded to reversals and non-reversals in the SRT task and which type of knowledge became explicit in the explicit test. Tanaka and Watanabe (2013, 2014) also showed that after learning a set of triplets in an SRT task, participants were particularly fast to the triplets with the elements in reverse order compared to other re-orderings, even in participants who claimed not to notice the pattern.

Interestingly, both simulation and experimental work indicate that rule learning in implicit learning is at least largely associatively-driven and extracting the statistical regularities in the sequence play a crucial role (e.g. Cleeremans & Dienes, 2008; Speiegel & McLaren, 2006). Few studies have investigated abstract learning in the SRT task because it is difficult to distinguish rule learning from associative learning. To address this issue, we adopted three types of stimuli that differed in only the sequence location to detect the effects of associative learning and rule learning separately in the training phase in the present study. “Standard” stimuli refer to the stimuli that followed the training SOC sequence, and appeared with a high probability. “Transfer” stimuli refer to the stimuli that followed the transfer SOC sequence, but appeared with a low probability. “Deviant” stimuli refer to the stimuli that followed neither SOC sequence and appeared with a low probability, like “transfer” stimuli. That is, “standard” and “transfer” had similar abstract structure about repetition patterns but differed in the probability of occurrence, while “transfer” and “deviant” both appeared with a similar low probability but differed in the abstract structure. If people acquired only chunking knowledge about the probability of occurrence for each type of stimuli through associative learning, there would be no difference on RTs between “deviant” and “transfer” stimuli as their probabilities were similarly low. Otherwise, if people simultaneously acquire knowledge about chunks or triplets and abstract structures, they would respond faster for “standard” than “transfer” and faster for “transfer” than “deviant”.

Further, to measure the explicit status of the acquired knowledge, we adopted two methods: the process dissociation procedure (PDP), which measures the ability to control the use of the knowledge (Jacoby, 1991; for bias of the PDP measure see Stahl, Barth, & Haider, 2015); and subjective measures, which measure awareness of knowing (Dienes, 2012). The difference between inclusion and exclusion performance reflects the ability to control the use of the acquired knowledge (Jacoby, 1991; Wilkinson & Shanks, 2004). Chunking knowledge was measured by the difference in probability of occurrence between “standard” and “transfer” triplets; thus, control of chunking knowledge was measured as the difference in generation performance between the inclusion and exclusion test for “standard” and “transfer” triplets. Abstract knowledge was measured by the difference in abstract structure between “transfer” and “deviant”; and control of abstract knowledge was measured as the difference in generation performance between the inclusion and exclusion test for “transfer” and “deviant” triplets. Further, confidence ratings were taken for generation performance to measure awareness of knowing. Awareness of knowing often goes together with ability to control the knowledge, but awareness of knowledge and control of that knowledge can dissociate (Fu, Dienes, & Fu, 2010; Wan, Dienes, & Fu, 2008); thus, both types of measures were used for a nuanced assessment of the explicit nature of the acquired knowledge. For example, if one consciously knew that an item was legal one could include or exclude it as instructed, thereby exerting control; but one need not consciously know why it is legal. Thus in Experiment 3 we will specifically measure the conscious status of both the judgment that a triplet is legal as well as the conscious status of the knowledge that enabled that judgment.

Experiment 1

To explore how associative learning dissociated from rule learning, “standard”, “transfer”, and “deviant” stimuli were adopted in the training phase and the probabilities of “standard”, “transfer”, and “deviant” triplets were set as .883, .083, and .083, respectively in Experiment 1. As three types of stimuli may make the sequence learning difficult, all participants were trained with a relatively long training phase.

Method

Participants

Twenty-five undergraduate students (13 male, 12 female) took part in this experiment. None of them had previously taken part in any implicit learning experiment. They were paid for their attendance. This experiment was approved by the committee for the protection of subjects at the Institute of Psychology, Chinese Academy of Sciences, and so were Experiments 2 and 3.

Apparatus and Materials

The experiment was programmed in E-prime 1.2 and run on Pentium-compatible PCs. The display consisted of a red, yellow, blue, or green square in the centre of the computer's screen against a silver gray background. The red, yellow, blue, and green colour squares corresponded to numerals 1, 2, 3, and 4 in the two SOCs (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) that were presented in a circular fashion (see Reed & Johnson, 1994). Each SOC sequence can be broken down into 12 sequential chunks of three colours, or triplets. For example, 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 the triplets 3-4-1, 4-1-2, 1-2-4, and so on. To generate the probabilistic sequences, the stimuli were followed by the triplets from a training SOC sequence (i.e., "standard" stimuli) with a high probability of .833, followed by the triplets from the other SOC sequence (i.e., "transfer" stimuli) with a low probabilities of .083, and followed by the triplets from neither SOC sequence (i.e., "deviant" stimuli) with a low probabilities of .083 in each block. Figure 1 shows the exemplars of the probabilistic sequences in the training phase.

----------- Insert Figure 1 about here -----------

Procedure

Training phase. Participants were asked to complete a serial four-choice reaction time task. On each trial, a colour square appeared in the center of the screen and covered visual angle of approximately 1°. Participants were instructed to respond as quickly and as accurately as possible by pressing a corresponding key on the keyboard. Keys D, F, J, and K corresponded to red, yellow, blue, and green colour squares. Participants were required to respond to Keys D and F with the middle and index fingers of their left hand, and to respond to Keys J and K with the index and middle fingers of their right hand. The target was removed as soon as a correct key had been pressed, and the next stimulus appeared after 500 ms, i.e. the response stimulus interval is 500 ms. Response latencies were measured from the onset of the target to the completion of a correct response and errors were recorded (see Figure 2). If an incorrect key was pressed, the stimulus would appear again until the correct key was pressed. Unbeknownst to participants, the colour squares followed a probabilistic sequence that consisted of 146 trials in each block. Thirty-second rest breaks occurred between any two experimental blocks. There were 20 training blocks, for a total of 2920 trials. For counter balancing purposes, half of the participants in each condition were trained on SOC1 and half on SOC2.

Test phase. The test phase included two trial-by-trial generation tests: an inclusion test and an exclusion test. At the beginning of each test, all participants were informed that the colour squares had followed a regular sequence, in which most colour squares were determined by the previous two. On each test trial, participants were first instructed to respond to a short sequence of two movements as in the training. Then, a black square appeared and they were required to generate next colour square by pressing a corresponding key. In the inclusion test, they were required to generate the colour square that appeared most frequently after the previous two in the training; and in the exclusion test, they were required to generate the colour square that appeared seldom after the previous two in the training. The same colour square never immediately repeated in training and so the chance level for a correct response is 1/3 every time for each sequence. After each generation, participants reported the confidence level of their judgment by inputting one of: 5, 6, 7, 8, 9, and 0 that corresponded to 50%, 60%, 70%, 80%, 90%, and 100%, in which 50% meant complete guessing and 100% meant certainty. Participants had to take one hand off the keys every trial to input the number. To reduce the influence of the confidence judgment on next trial, a screen stating “Are you ready? Please press the space key to continue” was presented after the confidence judgment. Thus, the next test trial would appear only when participants were ready. In each test, 12 different test trials were presented in a random order and repeated 12 times to make a total of 144 test trials.

Inferential strategy

For all tests *p* values are reported; in addition, for all *t* tests, Bayes factors, *B*, are reported. Bayes factors (B) were used to assess strength of evidence (Wagenmakers et al., in press). A *B* of above 3 indicates substantial evidence for the alternative hypothesis (H1) and below 1/3 substantial evidence for the null hypothesis (H0); “substantial” in the sense of just worth taking note of. *B*s between 3 and 1/3 indicate data insensitivity (see Dienes, 2014; cf Jeffreys, 1939). Thus we will report that there was no effect only when B < 1/3. Here, $B_{\mathrm{N}\,(0,\,\mathrm{x})}$ refers to a Bayes factor in which the predictions of H1 were modeled as normal distribution with an SD of x (see Dienes, 2014), where x scales the size of effect that could be expected.

In Experiment 2 of Fu et al. (2008), RT differences of about 20 ms were found between probable (i.e., standard) and improbable (i.e., transfer) stimuli. Thus, for RTs, we used a half-normal with SD = 20 ms to model H1; for simplicity, we did this for all tests of RT effects. Similarly, in Experiment 2 of Fu et al. (2008), differences in error proportions of about .03 were found between standard and transfer. Thus, we report $B_{N(0, .03)}$ for all contrasts of error proportions. In Experiment 2 of Fu et al. (2008), in the generation test, differences of around .10 were found. Thus, we report $B_{N(0, .10)}$ for contrasts involving the generation test. In Experiment 1 of Fu et al. (2010), an average slope of .89 was found for the I-E difference of standard. Thus, we report $B_{N(0, .89)}$ for the slope. With these assumptions for modeling H1, as it happened, where an effect yielded a *p* value of about .05, the Bayes factor was about 3, though there is no guarantee of such a correspondence between *B* and *p* values (Lindley, 1957). We will interpret all effects with respect to the Bayes factors.

Results

Training data. Trials with RTs greater than 1,500 milliseconds were dropped. these amounted to 0.77% of the trials.

----------- Insert Figure 2 about here -----------

Figure 2 shows mean RTs obtained over the training phase in Experiment 1. The RT advantage of standard over transfer indicates associative learning of chunking knowledge and the RT advantage of transfer over deviant indicates rule learning of abstract structure. To examine whether two types of learning can occurs simultaneously, an ANOVA on RTs with type of stimuli (standard vs. transfer vs.

deviant) and blocks (20 levels) as within-subject variables was used. It revealed a significant effect of type of stimuli (see Table 1), and participants responded to standard more quickly than to transfer stimuli, t (24) = 7.74, $p < .001$, dz = 1.55, $B_{\mathrm{N}\,(0,\,20)} = 9.18\times10^{10}$, and responded to transfer more quickly than to deviant stimuli, t (24) = 10.87, $p < .001$, dz = 2.17, $B_{\mathrm{N}\,(0,\,20)} = 2.49\times10^{16}$. That is, people acquired knowledge of both chunks and the abstract structure of the sequence. The main effect of blocks reached significance, and so did the interaction of type of stimuli by block, indicating that the learning effects were greater later in practice than earlier on.

----------- Insert Table 1 about here -----------

As deviant stimuli included *ten* reversals that differed from the abstract SOC structure of standard and transfer which consisted of *one* reversal, the slower RT effect for deviant may due to the reversal frequency. To examine this possibility, we calculated RTs for each type of stimulus when deviant were reversals or non-reversals. An ANOVA with type of stimulus (standard vs. transfer vs. deviant) and type of deviant (reversal vs. non-reversal) as within-subject variables was used. It revealed a significant effect of type of stimulus, an effect of type of deviant, and a significant type of stimulus by type of deviant interaction (see Table 1). Further analysis revealed that when deviants were reversals, participants responded more quickly to standard than to transfer, t (24) = 7.92, $p < .001$, $dz = 1.58$, $B_{\mathrm{N}\,(0,\,20)} = 8.65\times10^{10}$, more quickly to transfer than to deviant, t (24) = 16.44, $p < .001$, $dz = 3.29$, $B_{\mathrm{N}\,(0,\,20)} = 2.01\times10^{29}$;

but when deviants were non-reversals, they responded more quickly to deviant than to standard or transfer stimuli, *t* (24) = 3.45, *p* < .001, *dz* = .69, $B_{\mathrm{N}\,(0,\,20)}$ = 18.16, *t* (24) = 5.54, *p* < .001, *dz* = 1.11, $B_{\mathrm{N}\,(0,\,20)}$ = 7.50×10^{2}, respectively. In sum, people responded to standard more quickly than to transfer and responded to transfer more quickly than to deviant only when the deviant was a reversal. That is, the slower RTs for deviant stimuli were due to the reversals.

Overall, the mean error proportions for standard, transfer, and deviant stimuli were .05 (*SD* = .03), .09 (*SD* = .05), .12 (*SD* = .06). The error proportion for standard was lower than for transfer stimuli, *t* (24) = 4.99, *p* < .001, *dz* = 1.00, $B_{\mathrm{N}\,(0,\,.03)}$ = 1.31×10^{4}, while the error proportion for transfer were lower than for deviant stimuli, *t* (24) = 4.92, *p* < .001, *dz* = .98, $B_{\mathrm{N}\,(0,\,.03)}$ = 3.73×10^{4}. The RT effects were not compromised by speed-error trade-offs.

Testing Phase. Table 2 shows mean proportions for each type of triplet in Experiment 1. As participants expressed knowledge of concrete triplets and abstract structure only when deviant were reversals, we analyzed only the generation performance for deviants being reversals.

----------- Insert Table 2 about here -----------

More standard generated under inclusion than exclusion, i.e., I > E for standard, was often taken to indicate the acquisition of explicit knowledge. However, if participants mainly explicitly learned chunking knowledge distinguishing standard and transfer triplets, the I-E difference for standard and transfer would be different. To investigate whether people expressed control over chunking knowledge, an ANOVA with instructions (inclusion vs. exclusion) and type of SOC triplet (standard vs. transfer) as the within-subject variables was used. It revealed an instruction effect, a type of SOC triplet effect, and an instruction by type of SOC triplet interaction (see Table 3). Further analysis revealed that there were more standard than transfer under inclusion (i.e., I > B), t (24) = 6.20, p < .001, dz = 1.45, $B_{\mathrm{N}\,(0,\,.10)} = 2.48\times10^{6}$, but there was no differences between standard and transfer under exclusion (i.e., E = B), t (24) = -1.51, p = .15, $B_{\mathrm{N}\,(0,\,.10)}$ = .10. Importantly, participants generated more standard under inclusion than exclusion (i.e., I > E), t (24) = 6.87, p < .001, dz = 1.35, $B_{\mathrm{N}\,(0,\,.10)} = 5.27\times10^{5}$, and more transfer under inclusion than exclusion (i.e., I > E), t (24) = 1.95, p = .063, dz = .39, $B_{\mathrm{N}\,(0,\,.10)}$ = 3.90. The I > E for transfer indicated that people produced more incorrect fragments when attempting to produce the sequence than withhold it. The qualitatively similar I-E pattern for standard and transfer indicated that people lacked control over their chunking knowledge in the sense that they inaccurately represented transfer triplets as high frequency.

----------- Insert Table 3 about here -----------

If participants mainly explicitly learned abstract knowledge distinguishing transfer and different triplets, the I-E difference for transfer and deviant would be different. To test whether people expressed control over the acquired abstract knowledge, an ANOVA with instruction (inclusion vs. exclusion) and type of small-probability triplets (transfer vs. deviant) as the within-subject variables was used. It revealed an instruction effect, a type of small-probability triplet effect, and an instruction by type of small-probability triplet interaction (see Table 3). Further analysis revealed that participants generated more transfer under inclusion than exclusion (i.e., I > E), t (24) = 1.95, p = .063, dz = .39, $B_{\mathrm{N}\ (0,\ .10)}$ = 3.90, but less deviant under inclusion than exclusion (i.e., I < E), t (24) = -7.47, p < .001, dz = 1.06, $B_{\mathrm{N}\ (0,\ .10)} = 4.35\times10^3$. The different I-E pattern between transfer and deviant suggested that people treated transfer and deviant stimuli differently, i.e., they had control over the use of the knowledge of the abstract structures.

----------- Insert Figure 3 about here-----------

After each test trial, participants gave a confidence rating on a 50% to 100% scale. We calculated the regression coefficient of I-E difference against confidence ratings for deviant being reversals separately for each participant (Dienes & Longuet-Higgins, 2004; Fu, Dienes, & Fu, 2010), as participants expressed knowledge of concrete triplets and abstract structure when deviant triplets were reversals. Figure 3 shows mean I-E differences against confidence in Experiment 1. The I-E difference was below zero for deviant stimuli, t (21) = 3.70, p = .001, dz = .79, $B_{\mathrm{N}\ (0,\ .10)}$ = 3.57, when participants gave 50% confidence, but there was no evidence for whether or not the I-E difference was different from chance for standard and transfer stimuli, t (21) = -1.11, p = .28, $B_{\mathrm{N}\ (0,\ .10)}$ = 1.27, t (21) = -1.17, p = .25, $B_{\mathrm{N}\ (0,\ .10)}$ = 1.25, respectively. The slope was above zero for standard and transfer, t (21) = 1.76, p = .09, $B_{\mathrm{N}\ (0,\ .89)}$ = 3.30, t (21) = 2.05, p = .053, dz = .44, $B_{\mathrm{N}\ (0,\ .89)}$ = 5.30, and was below zero for deviant, t (21) = -4.72, p < .001, dz = 1.01, $B_{\mathrm{N}\ (0,\ .89)}$ = 9.14×10^3. The results indicated that the ability to control use of abstract knowledge was associated with awareness of knowing that knowledge.

Discussion

The RT results showed that participants responded faster to standard stimuli than to transfer stimuli and faster to transfer stimuli than to deviant stimuli, providing strong evidence that they simultaneous acquired knowledge about both concrete triplets and abstract structures. Moreover, there was only such a difference among the three types of stimuli when the deviant was a reversal triplet, indicating that the abstract structure people acquired was whether or not the stimulus was a reversal (cf Tanaka & Watanabe, 2013, 2014; also Li, Jiang, Guo, Yang, & Dienes, 2013, for implicit learning of reversals in artificial grammar learning). When the deviants were non-reversals, people responded more quickly to deviant than to standard or transfer stimuli, suggesting that rule learning about the abstract structure of whether being a reversal could overcome associative learning about the probability of occurrence for concrete triplets.

Interestingly, although people were asked to generate more standard under the inclusion test and less transfer and deviant under the exclusion test, they generated more standard and transfer under inclusion than under exclusion, i.e., I > E for both standard and transfer, and less deviant under inclusion than exclusion, i.e., I < E for deviant. The results indicated that people treated standard and transfer stimuli as similar, but transfer and deviant stimuli as different. Standard and transfer stimuli shared the same abstract structure and differed in the probability of occurrence, while transfer and deviant stimuli occurred with small probabilities but differed in abstract structure. The results revealed that the ability to control was mainly based on the abstract structure rather than the triplet as such. Moreover, confidence ratings indicated that only the ability to control the use of abstract knowledge was associated with the awareness of knowing. That is, at least some of the chunking knowledge was unconscious while the abstract knowledge was conscious in terms of knowing the legal status of each triplet. The findings leave open whether the abstract structure distinguishing deviants from transfer stimuli was itself known consciously, a point we return to in Experiment 3.

Experiment 2

Fu, Fu, and Dienes (2008) found that the amount of training influenced the conscious status of knowledge of concrete triplets and abstract structure: participants in the 6-block group acquired unconscious knowledge of abstract structure and concrete triplets, while participants in the 15-block group acquired conscious knowledge of concrete triplets and abstract structure. To further explore whether abstract knowledge can be implicitly acquired early in training, we reduced the training phase from twenty to five blocks in Experiment 2.

Method

Participants

Twenty-four undergraduate students (10 male, 14 female) took part in the experiment. None of them had previously taken part in any implicit learning experiment. They were paid for their attendance.

Apparatus and Materials

Apparatus and Materials were identical to Experiment 1.

Procedure

The procedure was same as Experiment 1 except that the training phase included only five training blocks for each group.

Results

Training data

----------- Insert Figure 4 about here-----------

Trials with RTs greater than 1,500 milliseconds were dropped; these amounted to 1.04% of the trials. Figure 4 shows the mean RTs obtained over the training phase in Experiment 2. To examine whether people can acquire chunking and abstract knowledge that distinguished standard from transfer and transfer from deviant separately, an ANOVA on RTs with type of stimuli (standard vs. transfer vs. deviant) and blocks (5 levels) as within-subject variables was used. It revealed only a significant effect of type of stimuli (see Table 4). Participants responded more quickly to standard than to transfer, t (23) = 2.65, $p < .05$, $dz = .54$, $B_{\mathrm{N}\,(0,\,20)} = 18.76$, and responded more quickly to transfer than to deviant, t (23) = 9.21, $p < .001$, dz = 1.88, $B_{\mathrm{N}\,(0,\,20)} = 1.05\times10^{9}$, indicating two types of learning occurred at the same time.

----------- Insert Table 4 about here-----------

As in Experiment 1, we calculated RTs for each type of triplets when deviant were reversals or non-reversals. An ANOVA with type of stimuli (standard vs. transfer vs. deviant) and type of deviant (reversal vs. non-reversal) as within-subject variables revealed an effect of type of deviant, a significant effect of type of stimuli, and a significant type of stimuli by type of deviant interaction (see Table 4). When deviant were reversals, participants responded more quickly to standard than to transfer stimuli, t (23) = 3.51, $p < .01$, dz = .72, $B_{\mathrm{N}\,(0,\,20)} = 1.39\times10^{2}$, and responded more quickly to transfer than to deviant stimuli, t (23) = 10.62, $p < .001$, dz = 2.17, $B_{\mathrm{N}\,(0,\,20)} = 3.50\times10^{8}$. However, when deviant stimuli were non-reversal triplets, participants responded more quickly to deviant than to standard and transfer, t (23) = 3.19, $p < .01$, dz = .65, $B_{\mathrm{N}\,(0,\,20)}$ = 12.44, t (23) = 5.57, $p < .001$, dz = 1.14, $B_{\mathrm{N}\,(0,\,20)} = 1.21\times10^{2}$, respectively. The results confirmed that people could distinguish transfer from deviant because of the reversals.

Overall, the mean error proportions for standard, transfer, and deviant were .06 (*SD* = .03), .07 (*SD* = .04), and .12 (*SD* = .06). The error proportion for standard was lower than for transfer stimuli, t (23) = -1.76, p = .092, dz = .36, $B_{\mathrm{N}\,(0,\,.03)}$ = 3.15, while the error proportion for transfer were lower than for deviant stimuli, t (24) =

-4.31, $p < .001$, $dz = .88$, $B_{\mathrm{N}(0, .03)} = 4.01\times10^2$. The results indicated that the RT effects were not compromised by speed-error trade-offs.

Test Data

----------- Insert Table 5 about here -----------

Table 5 shows the mean proportion of triplets generated in Experiment 2. As in Experiment 1, we analyzed only the generation performance for deviant stimuli which were reversals. To investigate whether people expressed control over the acquired concrete knowledge, i.e., whether the I-E difference for standard and transfer were different, an ANOVA with instructions (inclusion vs. exclusion) and type of SOC triplets (standard vs. transfer) as within-subject variables was used. It revealed an instruction effect and an interaction of instruction by type of SOC triplets (see Table 6). Further analysis revealed there were more standard than transfer under inclusion (i.e., I > B), $t(23) = 2.36$, $p < .05$, $dz = .48$, $B_{\mathrm{N}(0, .10)} = 8.76$, but there was no differences between standard and transfer under exclusion (i.e., E = B), $t(24) = -.07$, $p = .95$, $B_{\mathrm{N}(0, .10)} = 0.27$. Importantly, there was I > E for standard, $t(23) = 3.01$, $p < .01$, $dz = .62$, $B_{\mathrm{N}(0, .10)} = 32.07$, and for transfer, $t(23) = 2.30$, $p < .05$, $dz = .47$, $B_{\mathrm{N}(0, .10)} = 7.98$. The I > E for transfer confirmed that people lacked control over the use of knowledge of the concrete triplets to some extent.

----------- Insert Table 6 about here -----------

To examine whether people expressed control over the acquired abstract knowledge, i.e., whether the I-E difference for transfer and deviant were different, a comparable ANOVA with instructions (inclusion vs. exclusion) and type of small-probability triplets (transfer vs. deviant) as within-subject variables was used. It revealed an effect of instruction, an effect of type of small-probability triplets, and an interaction of instruction by type of small-probability triplets (see Table 6). Further analysis revealed that there was I > E for transfer, t (23) = 2.30, $p < .05$, $dz = .47$, $B_{\mathrm{N}\,(0,\,10)} = 7.98$, but I < E for deviant, t (23) = -2.98, $p < .01$, $dz = .61$, $B_{\mathrm{N}\,(0,\,.10)} = 12.49$. The different patterns for I-E differences between transfer and deviant confirmed that people had some ability to control the knowledge of abstract structures.

----------- Insert Figure 5 about here-----------

As in Experiment 1, we calculated regression coefficient of I-E difference against confidence ratings separately for each participant when deviant triplets were reversals. Figure 5 shows the mean generation difference between inclusion and exclusion against confidence ratings in Experiment 2. The I-E difference was below zero for standard stimuli when participants gave 50% confidence, t (19) = -2.45, $p < .05$, $dz = .55$, $B_{\mathrm{N}\,(0,\,.10)} = 3.09$, but there was no evidence for whether or not there was an I-E difference for transfer and deviant stimuli, t (19) = -.37, $p = .72$, $B_{\mathrm{N}\,(0,\,.10)} = 1.07$, t (19) = 1.94, $p = .067$, $B_{\mathrm{N}\,(0,\,.10)} = 1.74$, respectively. The slope was above zero for standard, t (19) = 3.16, $p < .01$, $dz = .71$, $B_{\mathrm{N}\,(0,\,.89)} = 69.75$, but there was no evidence for whether the slope was at or above zero for transfer, t (19) = -.04, $p = .97$, $B_{\mathrm{N}\,(0,\,.89)} = 0.41$. The slope was below zero for deviant stimuli, t (19) = -1.76, $p = .094$, $dz = .43$, $B_{\mathrm{N}\,(0,\,.89)} = 3.28$. That is, the ability to control the generation of standard and deviant stimuli was associated with the awareness of knowing the legal status of the stimulus.

Discussion

The training phase was reduced to five blocks in Experiment 2. The RT results showed that participants in both groups simultaneously acquired knowledge of both chunks and abstract structures, when the deviant stimuli were reversal triplets. Importantly, the test performance revealed that there was I > E for both standard and transfer stimuli, and I < E only for deviant stimuli, confirming that people had some ability to control the generation of triplets based on the abstract structure rather than the probability of occurrence. This is inconsistent with the findings of Fu et al. (2008) in which people expressed unconscious knowledge about abstract structure in the free-generation test when the training phase was short. A crucial difference between the free generation tests of Fu et al. and the trial-by-trial generation tests in this experiment is that participants can continuously produce those perhaps few triplets that come to their mind under free generation, while participants are tested on all triplets under trial-by-trial generation (Fu et al., 2010). As knowledge about abstract structures are embedded in all triplets while knowledge about concrete triplets are expressed by specific triplets, the trial-by-trial generation tests might be more sensitive to the conscious status of abstract strictures than the free-generation tests. Moreover, the results of confidence ratings suggested that the ability to control use of abstract knowledge was at least partially associated with awareness of knowing the legal status of each triplet, even with a short training time.

Experiment 3

Dienes and Scott (2005) pointed out that there are two types of knowledge relevant in implicit learning research: judgment knowledge is knowledge of whether a particular stimulus is legal given the context, and 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 (see Rebuschat, 2013). Generation tests and confidence ratings measure the conscious status of judgment knowledge and not structural knowledge. Fu et al. (2010) found that conscious judgment knowledge can be based on unconscious structural knowledge in the SRT task. Thus, to investigate whether people can acquire unconscious structural knowledge about abstract structures with a short training phase, we adopted the same attribution tests in the trial-by-trial tests as used by Fu et al. (2010). For each triplet in the generation task, subjects indicated whether their judgment was a pure guess, based on intuition (they had confidence but no idea why their answer was right), or based on memory of that triplet or knowledge of a rule they could state. When performance was above chance, memory and rule attributions were taken to indicate that both judgment and structural knowledge were conscious; the intuition attribution that judgment knowledge was conscious but structural knowledge unconscious; and the guess attribution that both judgment and structural knowledge were unconscious

Method

Participants

Twenty-four undergraduate students (7 male, 17 female) took part in the experiment. None of them had previously taken part in any implicit learning experiment. They were paid for their attendance.

Procedure

Training phase. The training phase was same as Experiment 1 except that it included six training blocks for each group.

Testing phase. The testing phase was similar to Experiment 1 with an exception that after each generation, participants were required to report the basis of their judgment by ticking one of: random or guess, intuition, rules or memory. Participants were provided with definitions taken from Dienes and Scott (2005). The guess attribution indicated that the judgment had no basis whatsoever, and it was equivalent to flipping a coin to arrive at the judgment. The intuition attribution indicated that the participant knew to some degree the judgment was right, but they had absolutely no idea why it was right. The rules or memory attribution indicated the participant felt they based their answer on some rule or rules acquired from the training phase and which they could state if asked or the person felt that the judgment was based on memory for particular items or parts of items from the training phase. To guarantee all subjects remembered the meanings of guess, intuition, rules and memory, the definitions were presented on each trial.

Inferential strategy

For the proportion of each attribution in the inclusion and exclusion test, with three choices the average proportion was .33, so we report $B_{\mathrm{N}\,(0,\,.33)}$.

Results

Training data

----------- Insert Figure 6 about here-----------

Trials with RTs greater than 1,500 milliseconds were dropped; these amounted to 0.93% of the trials. Figure 6 shows mean RTs obtained over the training phase in Experiment 3. To examine whether people can acquire chunking and abstract knowledge that distinguished standard from transfer and transfer from deviant at the same time, an ANOVA on RTs with type of stimuli (standard vs. transfer vs. deviant) and blocks (6 levels) as within-subject variables was used (see Table 7). It revealed only a significant effect of type of stimuli. Participants responded more quickly to standard than to transfer stimuli, $t\ (23) = 4.70$, $p < .001$, $dz = .96$, $B_{\mathrm{N}\,(0,\,20)} = 2.24\times10^4$, and more quickly to transfer than to deviant stimuli, $t\ (23) = 7.14$, $p < .001$, $dz = 1.46$, $B_{\mathrm{N}\,(0,\,20)} = 8.55\times10^6$, confirming the two types of learning occurred simultaneously.

----------- Insert Table 7 about here -----------

As in Experiments 1 and 2, we calculated RTs for each type of triplet when deviants were reversals or non-reversals. An ANOVA with type of stimuli (standard vs. transfer vs. deviant) and type of deviant (reversal vs. non-reversal) as within-subject variables revealed an effect of type of stimuli, a significant effect of type of deviant, and a significant type of stimuli by type of deviant interaction (see Table 7). Further analysis revealed that when deviants were reversals, participants responded more quickly to standard than to transfer stimuli, t (23) = 3.39, $p < .01$, $dz = .69$, $B_{\mathrm{N}\,(0,\,20)} = 1.43\times10^{2}$, and more quickly to transfer than to deviant stimuli, t (23) = 7.14, $p < .001$, $dz = 1.46$, $B_{\mathrm{N}\,(0,\,20)} = 8.96\times10^{7}$. However, when deviants were non-reversals, all participants responded more quickly to deviant than to standard or transfer stimuli, t (23) = 8.43, $p < .001$, $dz = 1.72$, $B_{\mathrm{N}\,(0,\,20)} = 1.65\times10^{5}$, t (23) = 5.96, $p < .001$, $dz = 1.22$, $B_{\mathrm{N}\,(0,\,20)} = 88.95$, respectively. This confirmed that the abstract structure people acquired was the property of being a reversal.

Overall, the mean error proportions for standard, transfer, and deviant stimuli were.06 ($SD = .05$), .07 ($SD = .04$), and .12 ($SD = .08$). There was no evidence one way or the other for a difference in error proportions between standard and transfer stimuli, t (23) = -1.26, $p = .22$, $dz = .26$, $B_{\mathrm{N}\,(0,\,.03)} = 1.30$, but there was evidence that the the error proportion for transfer was lower than for deviant stimuli, t (23) = -3.56, $p < .01$, $dz = .73$, $B_{\mathrm{N}\,(0,\,.03)} = 46.64$. The results suggested that the RT effects were not compromised by speed-error trade-offs.

Test Data

----------- Insert Table 8 about here -----------

Table 8 shows the mean proportions of triplet generated in Experiment 3. As in Experiments 1 and 2, we analyzed only the generation performance for deviants when they were reversals. To examine whether people expressed control over the acquired chunking knowledge, i.e., whether the I-E difference for standard and transfer were different, an ANOVA with instructions (inclusion vs. exclusion) and type of SOC triplets (standard vs. transfer) as within-subject variables was used. It revealed an instruction effect, a type of SOC triplets effect, and an interaction of instruction by type of SOC triplets (see Table 9). Further analysis revealed that there were more standard than transfer under inclusion (i.e., I > B), $t\ (23) = 3.40$, $p < .01$, $dz = .69$, $B_{\mathrm{N}\ (0,\ .10)} = 1.51\times10^2$, but there was no differences between standard and transfer under exclusion (i.e., E = B), $t\ (24) = .04$, $p = .97$, $B_{\mathrm{N}\ (0,\ .10)} = 0.31$. Importantly, there was I > E for standard, $t\ (23) = 4.86$, $p < .001$, $dz = .81$, $B_{\mathrm{N}\ (0,\ .10)} = 8.29\times10^3$, and there was no evidence one way or the other for the I-E difference for transfer, $t\ (23) = 1.52$, $p = .14$, $B_{\mathrm{N}\ (0,\ .10)} = 2.19$. That is, people lacked control over the concrete triplets to some extent.

----------- Insert Table 9 about here -----------

To test whether people expressed control over the acquired abstract knowledge, i.e., whether the I-E difference for transfer and deviant were different, a comparable

ANOVA with instructions (inclusion vs. exclusion) and type of low probability triplets (transfer vs. deviant) as within-subject variables was used. It revealed an effect of instruction, an effect of type of low probability triplets, and an interaction of instruction by type of low probability triplets (see Table 9). Further analysis revealed that there was no evidence one way or the other for the I-E difference for transfer, t (47) = 1.52, p = .14, $B_{\mathrm{N}\,(0,\,.10)}$ = 2.19, but there was evidence that I < E for deviant stimuli, t (23) = 3.76, p = .001, dz = .77, $B_{\mathrm{N}\,(0,\,.10)}$ = 69.71. The results indicated that people could express some control over abstract structure.

----------- Insert Figure 7 about here-----------

Figure 7 shows proportion and accuracy of each attribution in Experiment 3. An ANOVA on proportions with attributions (guess vs. intuition vs. rules or memory) and instructions (inclusion vs. exclusion) as within-subject variables revealed a significant attribution effect, F (1.40, 32.14) = 27.46, p < .001, η_p^2 = .54, which was qualified by a significant attribution by instruction interaction, F (1.30, 29.90) = 13.17, p < .001, η_p^2 = .36. Further analysis revealed that there were less guess attributions during inclusion than exclusion, t (23) = -3.76, p = .001, dz = .77, $B_{\mathrm{N}\,(0,\,.33)} = 4.82\times10^2$, but more intuition attributions during inclusion than exclusion, t (23) = 4.86, p < .001, dz = .99, $B_{\mathrm{N}\,(0,\,.33)} = 2.37\times10^4$, and no evidence one way or the other for the difference in rules or memory attribution between inclusion and exclusion, t (23) = 1.52, p = .14, $B_{\mathrm{N}\,(0,\,.33)}$ = .69, respectively.

Guess attributions indicate that the participant is unaware of both judgment and structural knowledge; intuition attributions indicate that the participant is aware of judgment knowledge but not structural knowledge; and rules and memory indicate the participants were aware of both judgment and structural knowledge (Dienes & Scott, 2005; Fu et al., 2010). To examine the conscious status of the acquired structural knowledge, we calculate the I-E difference for each type of triplet when people gave guess, intuition, and rules or memory attributions. The one-sample *t* tests revealed that when people gave guess attribution, there was no evidence one way or the other for an I-E difference for standard or transfer stimuli, $t\ (18) = .98$, $p = .34$, $B_{\mathrm{N}\ (0,\ .10)} = 1.44$, $t\ (18) = .44$, $p = .66$, $B_{\mathrm{N}\ (0,\ .10)} = 0.97$, but there was evidence for I < E for deviant stimuli, $t\ (18) = -2.14$, $p < .05$, $dz = .48$, $B_{\mathrm{N}\ (0,\ .10)} = 5.42$. When people gave intuition attributions, there was evidence for I > E for standard stimuli, $t\ (20) = 1.78$, $p = .091$, $dz = .39$, $B_{\mathrm{N}\ (0,\ .10)} = 3.38$, and no evidence for whether or not there was an I-E difference for transfer stimuli, $t\ (20) = .89$, $p = .38$, $B_{\mathrm{N}\ (0,\ .10)} = 1.33$, and evidence for I < E for deviant stimuli, $t\ (20) = -2.98$, $p < .01$, $dz = .65$, $B_{\mathrm{N}\ (0,\ .10)} = 19.65$. When people gave a rules or memory attribution, there was insensitive evidence for I-E differences for standard, transfer, or deviant, $t\ (15) = 1.94$, $p = .072$, $B_{\mathrm{N}\ (0,\ .10)} = 2.93$, $t\ (15) =.00$, $p = 1.00$, $B_{\mathrm{N}\ (0,\ .10)} = .79$, $t\ (15) = -1.87$, $p = .081$, $B_{\mathrm{N}\ (0,\ .10)} = 2.79$. The results provide evidence that the ability to control can be based on unconscious structural knowledge.

Discussion

As in Experiments 1 and 2, the RT results suggested that participants acquired knowledge about both concrete triplets and abstract structures when the deviant stimuli were reversal triplets, and the generation performance confirmed that the ability to control was mainly based on the abstract structure rather than knowledge of chunks. Importantly, we found that the ability to control was expressed when people gave guess and intuition attributions, suggesting that people acquired unconscious structural knowledge about abstract structures with a short training phase.

General discussion

The aim of the current study was to investigate whether people can implicitly acquire abstract structures rather than just memorized chunks in the SRT. To dissociate abstract or rule learning from associative learning, we adopted three types of stimuli in the training phase, in which standard and transfer shared the same abstract structure but differed in the probability of occurrence, while transfer and deviant stimuli occurred with the same low probabilities but differed in abstract structure. To measure the conscious status of the acquired knowledge, we adopted both ability to control as revealed by the PDP method, and awareness of knowing as revealed by the subjective measures as the key methods. The RT results showed that only when deviants were reversals did people respond more quickly to standard than to transfer, and more quickly to transfer than to deviant. People simultaneously acquired knowledge of chunks and the abstract structure of being a reversal. The generation performance showed that I > E for both standard and transfer stimuli but I < E for deviant stimuli; that is, the ability to control generation was mainly based on knowledge of abstract structures rather than concrete triplets. Moreover, generation performance for each attribution in Experiment 3 indicated that when the training phase was short the ability to control generation of abstract structures could be based on unconscious structural knowledge.

Can people acquire abstract knowledge in implicit sequence learning?

Although Goschke and Bolte (2007) found that people could implicitly acquire knowledge about abstract sequence structures in implicit sequence learning, other researchers have argued that abstract knowledge can be acquired only in explicit learning conditions (Boyer et al., 2005; Cleeremans & Destrebecqz, 2005; Shanks & St John, 1994). On this view, implicit sequence learning is associatively based rather than rule based (see Speiegel & McLaren, 2006). Of course, devices based on principles of association can have rules emerge in their representations (Dienes, 1992); they can become graded finite state devices (Cleeremans, 1993) or, if representationally rich enough, even graded supra-finite state devices (Rodriguez, Wiles, & Elman, 1999). Nonetheless, models based purely on chunking would have difficulties accounting for differences between transfer and deviant stimuli in the training and test phase in our experiments. The RT results in the three experiments showed that all participants responded faster to standard than to transfer and faster to transfer than to deviant stimuli regardless of the amount of training in all three experiments, indicating that people acquired knowledge about concrete chunks and, above and beyond chunks, about abstract structures. Importantly, the RTs were faster to standard than to transfer and faster to transfer than to deviant stimuli only when the deviants were reversals, suggesting that the abstract structure was being a reversal, as in ABA grammars (Marcus, 1999) or n-2 repetition structures (Koch, Philipp, & Gade, 2006). Forming associations between colours, or forming simply chunks of colours, would not allow this learning. An associative device that uses hidden units to capture abstract structure, then uses variable mappings onto this abstract structure in order to transfer the relations between different colours, could in principle learn (Altmann & Dienes, 1999; Dienes, Altmann, & Gao, 1999).

Whether the sort of network used by Dienes et al (1999) could simulate the results is an open question. Such a network would find chunks easier than abstract structure. Yet, we found that people expressed knowledge about abstract structures in the early training phase. Consistently, Pronk and Visser (2010) found that people responded faster to non-reversals than to reversals early in training in the one-reversal condition, indicating an early learning effect of the abstract structure. This might be partially because the abstract structure, an n-2 dependency, is a relatively easy repetition pattern. Moreover, there was significant interaction of type of stimuli by block in Experiments 1, indicating that the abstract learning effect was larger later than early on. While pure chunking models are ruled out, as well as associative models with colours as inputs and without hidden layers, further work is needed to discover which models of implicit learning could account for the present results.

Moreover, as the color sequence corresponded to the motor sequence in the present study, the acquired knowledge could be learned via either perceptual or motor sequence learning. Future work is needed to explore whether people can acquire abstract structures through pure perceptual or motor sequence learning.

Can the degree of abstraction determine the conscious status of knowledge?

We found that people generated more standard and transfer triplets under inclusion than exclusion (I > E) but less deviant triplets under inclusion than exclusion (I < E), regardless of the amount of training. As standard and transfer triplets had the same abstract structure but differed in the probability of occurrence (.83 vs. .083), while transfer and deviant triplets shared the same low probability of occurrences but differed in the abstract structure (reversal vs. non-reversal), the results revealed that people had control over the use of knowledge of the abstract structure but lacked control over the use of knowledge of concrete triplets in the sense that they inaccurately represented transfer triplets as high frequency. However, in some previous studies (e.g., Wilkinson & Shanks, 2004), the results that people generated more standard triplets under inclusion than exclusion (i.e., I > E for standard) and similar levels of standard under exclusion as baseline (i.e., E = B under exclusion), had been taken to suggest that people can express control over the use of knowledge of concrete triplets. Indeed, we also found I > E for standard and E = B under exclusion in most of conditions even when we analyzed only test trials when deviants were reversals. But, more importantly, our results further revealed that there was also I > E for transfer but I < E for deviant, suggesting that people treated standard and transfer triplets similarly but transfer and deviant triplets differently. That is, their ability to control was mainly based on abstract structures rather than concrete triplets.

Control over accepting a test triplet may plausibly go with knowing that one knows the triplet is legal or not. The relation between control and awareness of knowing is not necessary; one could have the experience of guessing and still exert control (e.g. Norman, Scott, Price, & Dienes, 2016). Still, control and awareness of knowing tend to go together, and consistently we found that confidence ratings predicted control (Experiments 1 and 2). Both confidence ratings and control assess awareness of “judgment knowledge”, i.e. the knowledge that an item is legal (Dienes, 2012). The judgment that a triplet was legal or not, when its legality was defined by being a reversal, appears to be conscious. However, the most interesting type of implicit knowledge may be structural knowledge, i.e. the knowledge of structural relations that enabled the judgment. Just because the judgment knowledge of reversals was conscious, that does not mean subjects knew that it was being a reversal that made the triplet legal. That is, structural knowledge can be unconscious when judgment knowledge is conscious. Experiment 3 explored the conscious status of both judgment and structural knowledge and found evidence for unconscious judgment knowledge of each of chunks and abstract structure, co-existing with conscious judgment knowledge of each. Importantly, there was unconscious structural knowledge of each.

To sum up, the current study demonstrated that people can simultaneously acquire knowledge about concrete chunks or abstract structures in implicit sequence learning. Moreover, our result also revealed that the ability to control was mainly based on knowledge of abstract structures rather than knowledge of chunks, which was not generally accompanied with the awareness of knowing measured by confidence and attribution tests. The results confirmed that people can acquire unconscious knowledge about abstract structures which presents a challenge to computational models and theories of implicit learning.

References

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

Altmann, G., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 899–912.

Boyer, M., Destrebecqz, A., & Cleeremans, A. (2005). Processing abstract sequence structure: learning without knowing, or knowing without learning? *Psychological Research, 69,* 383–398.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*(3), 316–323.

Channon, S., Shanks, D., Johnstone, T., Vakili, K., Chin, J., & Sinclair, E. (2002). Is implicit learning spared in amnesia? Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia, 40,* 2185–2197.

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

Cleeremans, A., & Destrebecqz, A. (2005). Real rules are conscious: Open peer commentary to Pothos. *Behavioral and Brain Sciences, 28,* 19–20.

Cleeremans, A., & Dienes, Z. (2008). Computational models of implicit learning. In R. Sun (Ed.), *Cambridge Handbook of Computational Psychology* (pp. 396–421). Cambridge University Press.

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.

Dienes, Z. & Longuet-Higgins, H. C. (2004). Can musical transformations be implicitly learned? *Cognitive Science, 28*, 531–558.

Dienes, Z. (1992). Connectionist and memory-array models of artificial grammar learning. *Cognitive Science, 16*, 41-79.

Dienes, Z. (2012). Conscious versus unconscious learning of structure. In P. Rebuschat & J. Williams (Eds), *Statistical Learning and Language Acquisition* (pp. 337 - 364). Mouton de Gruyter Publishers.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781. doi: 10.3389/fpsyg.2014.00781.

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

Dienes, Z., Altmann, G., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of transfer of implicit knowledge. *Cognitive Science, 23*(1):53–82.

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

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.

Fu, Q., Dienes, Z., & Fu, X. (2010). Can unconscious knowledge allow control in sequence learning? *Consciousness and Cognition, 19,* 462–474.

Fu, Q., Fu, X., & Dienes, Z. (2008). Implicit sequence learning and conscious awareness. *Consciousness and Cognition, 17*, 185–202.

Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology, 33*, 154–207.

Goschke ,T., & Bolte, A. (2007). Implicit learning of semantic category sequences: response-independent acquisition of abstract sequential regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 394–406.

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

Jeffreys, H. (1939). *The theory of probability*. Oxford, England: Oxford University Press.

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

Johnstonem, T., & Shanks, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology, 42*, 61–112.

Koch, I., Philipp, A. M., & Gade, M. (2006). Chunking in task sequences modulates task inhibition. *Psychological Science, 17*(4), 346–350.

Li, F., Jiang, S., Guo, X., Yang, Z., & Dienes, Z. (2013). The nature of the memory buffer in implicit learning: Learning Chinese tonal symmetries. *Consciousness & Cognition*, *22* (3), 920-930.

Lindley, D. V. (1957). A statistical paradox. *Biometrika, 44*(1–2), 187–192.

Ling, X., Li, F., Qiao, F., Guo, X., & Dienes, Z. (2016). Fluency Expresses Implicit Knowledge of Tonal Symmetry. *Frontiers in Psycholology, 7*, 57, doi: 10.3389/fpsyg.2016.00057

Marcus, G. F. (1999). Connectionism: with or without rules? Response to J.L. McClelland and D.C. Plaut. *Trends in Cognitive Sciences*, *3*(5), 168–170.

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.

Norman, E., Price, M. C., Duff, S. C., & Mentzoni, R. A. (2007). Gradations of awareness in a modified sequence learning task. *Consciousness and Cognition*, *16*, 809–837.

Norman, E., Scott, R. B., Price, M. C., & Dienes, Z. (2016). The relationship between strategic control and conscious structural knowledge in artificial grammar learning. *Consciousness and Cognition, 42*, 229–236.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General, 119,* 264–275.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *TRENDS in Cognitive Science, 10(5),* 233–238.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25,* 297–330.

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

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning, 63*(3), 595–626.

Rebuschat, P., & Williams, J. N. (2009). Implicit learning of word order. In N. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society*. (pp. 425-430). Austin, TX: Cognitive Science Society.

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.

Rodriguez, P., Wiles, J., & Elman, J. L. (1999). A recurrent neural network that learns to count. *Connection Science, 11*(1), 5-40.

Rohrmeier, M., Fu, Q., & Dienes, Z. (2012).Implicit learning of recursive, hierarchical grammatical structures. *PLoS ONE 7*(10): e45885. doi:10.1371/journal.pone.0045885

Sager, C. A. (1994). Implicit learning. *Psychological Bulletin, 115*(2), 163–196.

Shanks, D. R. (2004). Implicit learning. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of cognition*. London: Sage.

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

Speiegel, R., & McLaren, I. P. L. (2006). Associative sequence learning in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 32(2),* 150–163.

Stahl, C., Barth, M., & Haider, H. (2015). Distorted estimates of implicit and explicit learning in applications of the process-dissociation procedure to the SRT task. *Consciousness & Cognition*, *37*, 27-43. doi: 10.1016/j.concog.2015.08.003

Tanaka, K., & Watanabe, K. (2013). Implicit transfer of reversed temporal structure in visuomotor sequence learning. *Cognitive Science*, *38*, 565–579.

Tanaka, K., & Watanabe, K. (2014). Implicit transfer of spatial structure in visuomotor sequence learning. *Acta Psychologica,153*, 1–12.

Tunney, R. J., & Altmann, G. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 614

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic bulletin & review, 23*(1), 87-102.

Wagenmakers, E.-J., Verhagen, A. J., Ly, A., Matzke, D., Steingroever, H., Rouder, J.

N. et al. (in press). *The need for Bayesian hypothesis testing in psychological science*. In Lilienfeld, S. O., & Waldman, I. (Eds.), Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions. John Wiley and Sons.

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature, 411*, 958–956.

Wan, L. L., Dienes, Z., & Fu, X. L. (2008) Intentional control based on familiarity in artificial grammar learning.. *Consciousness and Cognition, 17*, 1209-1218.

Werheid, K., Zysset, S., Müller, A., Reuter, M., & von Cramon, D. Y. (2003). Rule learning in a serial reaction time task: An fMRI study on patients with early Parkinson’s disease. *Cognitive Brain Research, 16*, 273–284.

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

Acknowledgements

This research was supported by the National Natural Science Foundation of China and the German Research Foundation (NSFC 61621136008/DFG TRR-169) and the grant from the Natural Science Foundation of China (31270024).

Table 1. Significant results of the ANOVA on RTs with type of stimuli and blocks or type of stimuli type of deviants as within-subject variables in Experiment 1.

	Type of stimuli * Blocks		Type of stimuli * Type of deviants	
	F	η_p^2	*F*	η_p^2
Type of stimuli	144.50***	.86	9.69***	.29
Blocks/Type of deviants	2.52*	.10	9.27***	.28
Two-way interaction	1.79*	.07	76.42***	.76

Note: In each ANOVA, we report F values with significance and η_p^2. * $p < .05$; **$p < .01$; ***$p < .001$.

Table 2. Mean proportion of three types of triplets generated under inclusion and exclusion tests in Experiment 1.

	Reversals-deviant			Non-reversals-deviant		
	Standard	Transfer	Deviant	Standard	Transfer	Deviant
Inclusion	.52 (.01)	.37 (.01)	.11 (.01)	.33 (.07)	.38 (.05)	.29 (.05)
Exclusion	.29 (.03)	.32 (.02)	.39 (.05)	.26 (.06)	.24 (.05)	.51 (.07)

Table 3. Significant results of the ANOVA on RTs with instructions and type of SOC triplets or type of small-probability triplets as the within-subject variables in Experiment 1.

	Type of stimuli * Type of SOC triplets		Type of stimuli * Type of small-probability triplets	
	F	$\eta_p{}^2$	F	$\eta_p{}^2$
Type of stimuli	35.57***	.60	47.24***	.66
Blocks/Type of deviants	26.02***	.52	5.60*	.19
Two-way interaction	22.23***	.35	23.37***	.49

Note: In each ANOVA, we report F values with significance and $\eta_p{}^2$. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 4. Significant results of the ANOVA on RTs with type of stimuli and blocks or type of deviants as within-subject variables in Experiment 2.

	Type of stimuli * Blocks		Type of stimuli * Type of deviants	
	F	η_p^2	*F*	η_p^2
Type of stimuli	64.60***	.74	6.81*	.23
Blocks/Type of deviants			5.00*	.18
Two-way interaction			51.23***	.69

Note: In each ANOVA, we report *F* values with significance and η_p^2. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 5. Mean proportion of three types of triplets generated under inclusion and exclusion tests in Experiment 2.

	Reversals-deviant			Non-reversals-deviant		
	Standard	Transfer	Deviant	Standard	Transfer	Deviant
Inclusion	.46 (.02)	.41 (.02)	.13 (.03)	.16 (.04)	.35 (.05)	.49 (.06)
Exclusion	.35 (.03)	.35 (.03)	.30 (.06)	.24 (.07)	.45 (.07)	.31 (.06)

Table 6. Significant results of the ANOVA on RTs with instructions and type of SOC triplets or type of small-probability triplets as the within-subject variables in Experiment 2.

	Instructions * Type of SOC triplets		Instructions * Type of small-probability triplets	
	F	η_p^2	*F*	η_p^2
Instructions	8.86**	.28	9.08***	.28
Blocks/Type of deviants	4.12*	.15	9.98**	.30
Two-way interaction			8.13**	.26

Note: In each ANOVA, we report *F* values with significance and η_p^2. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 7. Significant results of the ANOVA on RTs with type of stimuli and blocks or type of deviants as within-subject variables in Experiment 3.

	Type of stimuli * Blocks		Type of stimuli * Type of deviants	
	F	$\eta_p{}^2$	F	$\eta_p{}^2$
Type of stimuli	73.45***	.76	8.80**	.23
Blocks/Type of deviants			3.61*	.14
Two-way interaction			55.86***	.71

Note: In each ANOVA, we report F values with significance and $\eta_p{}^2$. * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 8. Mean proportion of three types of triplets generated under inclusion and exclusion tests in Experiment 3.

	Reversals-deviant			Non-reversals-deviant		
	Standard	Transfer	Deviant	Standard	Transfer	Deviant
Inclusion	.49 (.01)	.40 (.01)	.11 (.01)	.19 (.06)	.41 (.07)	40 (.06)
Exclusion	.36 (.03)	.35 (.03)	.29 (.05)	.26 (.07)	.32 (.06)	.42 (.07)

Table 9. Significant results of the ANOVA on RTs with instructions and type of SOC triplets or type of small-probability triplets as the within-subject variables in Experiment 3.

	Instructions * Type of SOC triplets		Instructions * Type of small-probability triplets	
	F	$\eta_p{}^2$	*F*	$\eta_p{}^2$
Instructions	14.12***	.38	23.56***	.56
Type of SOC triplets / type of small-probability tripets	5.65*	.20	25.24***	.52
Two-way interaction	7.07*	.24	9.02**	.28

Note: In each ANOVA, we report *F* values with significance and $\eta_p{}^2$. * $p < .05$; **$p < .01$; ***$p < .001$.

Figure captions

Figure 1. A) The exemplars of the probabilistic sequences in the training phase, in which numerals 1, 2, 3, and 4 corresponded to red, yellow, blue, and green colour squares, respectively; B) Trial by trial frame of the procedure in the test phase; C) Trial by trial frame of the procedure in the test phase.

Figure 2. Mean RTs in Experiment 1. Error bars depict standard errors.

Figure 3. Mean differences between inclusion and exclusion of standard, transfer, and deviant generated in the test phase for each confidence rating in Experiment 1. Error bars depict standard errors.

Figure 4. Mean RTs in Experiment 2. Error bars depict standard errors.

Figure 5. Mean differences between inclusion and exclusion of standard, transfer, and deviant generated in the test phase for each confidence rating in Experiment 2. Error bars depict standard errors.

Figure 6. Mean RTs in Experiment 3. Error bars depict standard errors.

Figure 7. Proportions for different attributions and I-E differences for different type of

triplets generated in Experiment 3. Error bars depict standard errors.