THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2008, 61 (3), 400–424

Psychology Press
Taylor & Francis Group

Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies

Gilbert Remillard
Morehead State University, Morehead, KY, USA

Serial reaction time (SRT) task studies have established that people can implicitly learn first- and second-order adjacent dependencies. Sequential confounds have made it impossible to draw conclusions regarding learning of nonadjacent dependencies and learning of third- and fourth-order adjacent dependencies. Addressing the confounds, the present study shows that people can implicitly learn second-, third-, and fourth-order adjacent and nonadjacent dependencies embedded in probabilistic sequences of target locations.

Implicit sequence learning is sequence learning that is not the result of conscious, intentional processes and has been studied using the serial reaction time (SRT) task. On each trial, a target appears at one of a number of locations on a monitor, and the key corresponding to the location of the target is pressed. In many cases, the sequence of target locations is deterministic. Sequence learning occurs when the repeating sequence of target locations elicits shorter reaction times (RTs) than does a random sequence of target locations. In other cases, the sequence of target locations is probabilistic. Sequence learning occurs when, given previous target locations, more probable succeeding locations elicit shorter RTs than do less probable succeeding locations.

Most SRT task studies establish implicit sequence learning by assessing awareness of the sequence of target locations. Sequence learning that is explicit (i.e., the result of conscious processes) would presumably lead to an awareness of the sequence of target locations. Thus, a lack of awareness of the sequence of target locations would suggest that sequence learning was implicit. In many studies, RTs reveal learning of the sequence of target locations, and free recall, cued recall, or recognition tasks reveal no awareness of the sequence (e.g., Cleeremans & McClelland, 1991; Curran & Keele, 1993; Lewicki, Hill, & Bizot, 1988; McDowall, Lustig, & Parkin, 1995; Reed & Johnson, 1994; Stadler, 1989, 1993, 1995; Willingham, Nissen, & Bullemer, 1989).

When exposed to a sequence of events, there are various types of information that people potentially could learn implicitly. Table 1 lists some of the possibilities. *N*th-order probability information means that information on trial $t - n$ is required to differentially predict events on trial t. *N*th-order information can be subdivided into adjacent and nonadjacent dependencies. Adjacent dependencies involve information from consecutive trials (e.g., lag 1, lag 2–1, lag 3–2–1, and

Correspondence should be addressed to Gilbert Remillard, Department of Psychology, 601 Ginger Hall, Morehead State University, Morehead, KY 40351, USA. E-mail: g.remillard@moreheadstate.edu
This research was supported by a Morehead State University Research and Creative Productions Committee grant to the author.

http://www.psypress.com/qjep
DOI:10.1080/17470210701210999

Table 1. *Types of information*

Probability information	*Dependency*	*Specific probability*	*Symbolically*
First-order	Adjacent	Lag 1	$P(E\|A_1)$
Second-order	Adjacent	Lag 2−1	$P(E\|A_2{-}A_1)$
	Nonadjacent	Lag 2−x	$P(E\|A_2{-}x)$
Third-order	Adjacent	Lag 3−2−1	$P(E\|A_3{-}A_2{-}A_1)$
	Nonadjacent	Lag 3−2−x	$P(E\|A_3{-}A_2{-}x)$
		Lag 3−x−1	$P(E\|A_3{-}x{-}A_1)$
		Lag 3−x−x	$P(E\|A_3{-}x{-}x)$
Fourth-order	Adjacent	Lag 4−3−2−1	$P(E\|A_4{-}A_3{-}A_2{-}A_1)$
	Nonadjacent	Lag 4−3−2−x	$P(E\|A_4{-}A_3{-}A_2{-}x)$
		Lag 4−3−x−1	$P(E\|A_4{-}A_3{-}x{-}A_1)$
		Lag 4−x−2−1	$P(E\|A_4{-}x{-}A_2{-}A_1)$
		Lag 4−3−x−x	$P(E\|A_4{-}A_3{-}x{-}x)$
		Lag 4−x−2−x	$P(E\|A_4{-}x{-}A_2{-}x)$
		Lag 4−x−x−1	$P(E\|A_4{-}x{-}x{-}A_1)$
		Lag 4−x−x−x	$P(E\|A_4{-}x{-}x{-}x)$

Note: In the last column, E refers to an event on trial t, A_n refers to an event on trial $t-n$, and the letter x is a placeholder. For example, the lag 3−2−1 probability $P(E|A_3{-}A_2{-}A_1)$ is read as "the probability of E occurring on trial t given the occurrence of A_3, A_2, and A_1 on trials $t-3$, $t-2$, and $t-1$, respectively". The lag 4−x−2−x probability $P(E|A_4{-}x{-}A_2{-}x)$ is read as "the probability of E occurring on trial t given the occurrence of A_4 and A_2 on trials $t-4$ and $t-2$, respectively".

lag 4−3−2−1 probabilities). Nonadjacent dependencies involve information that skips over at least one trial (e.g., lag 2−x, lag 3−x−1, and lag 4−x−2−x probabilities). The latter dependencies could be further subdivided according to the number of trials that are skipped and the locations of the skipped trials.

People can implicitly learn lag 1 and lag 2−1 probabilities (e.g., Remillard & Clark, 2001). Implicit learning of lag 2−x probabilities, however, is an open issue. This is because SRT task studies have (a) confounded lag 2−1 and lag 2−x probabilities (Cleeremans, 1993, pp. 144–167; D. V. Howard, Howard, Japikse et al., 2004; J. H. Howard & Howard, 1997; Schvaneveldt & Gomez, 1998; Shanks, Wilkinson, & Channon, 2003; Wilkinson & Shanks, 2004) or (b) varied lag 2−1 probabilities while keeping lag 2−x probabilities constant (Jimenez & Mendez, 1999; Remillard & Clark, 2001).[1]

Lag 2−x probability learning has been the focus of word segmentation studies. In these studies, participants listen to a sequence of trisyllabic nonsense words (e.g., pagute, bepodu), six-tone words (e.g., $T_1{-}T_2{-}T_3{-}T_4{-}T_5{-}T_6$, $U_1{-}U_2{-}U_3{-}U_4{-}U_5{-}U_6$), or three-word phrases (e.g., pel−wadim−rud, vot−wadim−jic) with lag 2−x probabilities of 1.0 between the first and third syllables in a word, the first and third words in a phrase, or between every second letter or tone in a word. The results of these studies suggest that people are capable of learning the lag 2−x probabilities when there are short pauses between the words or phrases in a sequence (Gomez, 2002; Perruchet, Tyler, Galland, & Peereman, 2004). However, if a sequence of words is presented in a continuous fashion without pauses between words then people appear incapable of learning the lag 2−x probabilities unless odd and even elements in a word come from different domains (Bonatti, Pena, Nespor, & Mehler, 2005; Creel, Newport, & Aslin, 2004; Newport & Aslin, 2004; Onnis, Monaghan, Richmond, & Chater, 2005). For example, people will learn the lag 2−x probabilities between the tones T_1 and T_3, and T_3 and T_5 in the tone word $T_1{-}T_2{-}T_3{-}T_4{-}T_5{-}T_6$ only if the odd

[1] Most studies do not report exact lag 2−1 or lag 2−x probabilities. For those studies, I generated sequences as described in the articles and calculated lag 2−1 and lag 2−x probabilities.

tones T_1, T_3, and T_5 are of one pitch, and the even tones T_2, T_4, and T_6 are of a different pitch. This result is similar to the results of SRT task studies showing that people can simultaneously learn two, interleaved sequences from different domains (Frensch, 1998; Hoffmann, Sebald, & Stocker, 2001; Mayr, 1996; Shin & Ivry, 2002). However, it is not clear in these studies that participants were learning lag 2−*x* probabilities. There is the possibility that the cognitive system was segregating odd and even elements into two streams and learning adjacent dependencies (e.g., lag 1 probabilities) within each stream (Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003; Newport & Aslin, 2004, pp. 156−157).

There is good evidence from SRT task studies that people can implicitly learn third-order information (Cleeremans, 1993, pp. 144−167; Cleeremans & McClelland, 1991; D. V. Howard et al., 2004; Remillard & Clark, 2001; but see Jimenez, Mendez, & Cleeremans, 1996). However, in these studies, lag 3−2−1 probabilities were confounded with lag 3−*x*−*x*, lag 3−*x*−1, or lag 3−2−*x* probabilities. Thus it is not clear whether adjacent or nonadjacent dependencies were learned.

Only two SRT task studies have examined learning of fourth-order information. Cleeremans and McClelland (1991, Exp. 2) compared last-element RTs of the runs AJCM−A, AMLJ−A, XJCM−X, and XMLJ−X with last-element RTs of the runs AJCM−X, AMLJ−X, XJCM−A, and XMLJ−A (where letters represent target locations). Although fourth-order probabilities were higher in the former set of runs, RTs did not differ across the two sets. This suggests that there was no learning of fourth-order information. However, there is another explanation for the absence of an RT difference. The last elements in the former set of runs occurred more recently in the sequence (i.e., four trials back) than the last elements in the latter set of runs (i.e., more than four trials back). RT to a target location tends to decrease as the number of trials since that target location's last occurrence increases (Boyer, Destrebecqz, & Cleeremans, 2005; Curran, Smith, DiFranco, & Daggy, 2001). It also could be that when the preceding four target locations are all different (e.g., AJCM), people expect a different target location on the next trial (e.g., AJCM−X rather than AJCM−A). In either case, the result would be speeded last-element RTs in the latter set of runs or slowed last-element RTs in the former set and a masking of evidence for learning of fourth-order information. Counterbalancing the assignment of subsequences (e.g., AJCM−A, AJCM−X) to high- and low-probability transitions might have revealed learning of fourth-order information.

Cleeremans (1993, pp. 144−167) also examined learning of fourth-order information. When A occurred on trial $t-4$, C was more likely than B to occur on trial *t*. Despite the fact that the run AMNN was equally likely to be followed by C and B, RT was faster to C than to B following AMNN. This suggests that subjects had learned lag 4−*x*−*x*−*x* probabilities. However, learning of second-order or third-order information could explain the results. In the sequence of target locations, C was more likely than B to follow the runs AM and AMN. C was also more likely than B to occur on trial *t* when A occurred on trials $t-2$ or $t-3$. Thus faster responding to C than to B following AMNN may have been the result of greater priming of C than B after encountering AM or AMN, which carried over to AMNN. A response that is primed but not executed may remain primed over the next couple of trials (Cleeremans & McClelland, 1991, p. 245). Thus the results from the two preceding studies are ambiguous as to whether or not people can learn fourth-order information.

In summary, it is unclear whether people can implicitly learn nonadjacent dependencies or fourth-order information in a continuous sequence of events. Clarifying these issues is important for three reasons. First, it would substantially increase our understanding of the capabilities of the implicit sequence learning mechanism. Second, elucidating the sequential dependencies that can be learned places constraints on models of sequence learning. For example, if people are capable of learning fourth-order information, then a model of sequence learning must be equally capable. Finally, a number of authors have proposed that sequence learning involves the formation of associations between adjacent events (e.g., Keele et al., 2003; Koch & Hoffmann, 2000; Perruchet & Vinter, 1998; St. John & Shanks, 1997). Obtaining evidence for learning of nonadjacent dependencies would suggest that sequence learning is not limited to the formation of associations between adjacent events and would be inconsistent with some current models of sequence learning (Perruchet & Pacton, 2006). Consequently, the goal of the present study was to examine whether people can implicitly learn second-, third-, and fourth-order adjacent and nonadjacent dependencies.

EXPERIMENT 1

Experiment 1 used a six-choice SRT task to examine implicit learning of lag $2-x$ probabilities of .60 versus .40. The approach was straightforward. For example, consider the two runs $3-2-1$ and $3-2-6$ where $P(1|3-2) = .60$ (high-probability transition), $P(6|3-2) = .40$ (low-probability transition), and $P(1|3-x) = P(6|3-x) = .50$. Lag $2-1$ probabilities vary across the two runs whereas lag $2-x$ probabilities remain constant. Now consider the two runs $5-1-4$ and $5-1-3$ where $P(4|5-1) = P(4|5-x) = .60$ and $P(3|5-1) = P(3|5-x) = .40$. Lag $2-1$ and lag $2-x$ probabilities are confounded, and both vary across the two runs. If people are incapable of learning lag $2-x$ probabilities, then the RT difference between high- and low-probability transitions should be similar in the two sets of runs. Therefore, a greater RT difference between high- and low-probability transitions for the second set of runs than for the first set of runs would implicate learning of the lag $2-x$ probabilities in the second set of runs.

Method

Participants

The participants were 24 University of Winnipeg (Winnipeg, Manitoba, Canada) undergraduates ranging in age from 18 to 31 years.

SRT task

The SRT task was run on a personal computer with standard monitor and keyboard. Millisecond timing was implemented using Bovens and Brysbaert's (1990) routine. The lines marking the six horizontally arranged target locations were 0.3 cm in length and were separated by intervals of 2.0 cm. Viewing distance was approximately 60 cm. The red-stickered *S*, *D*, *F*, *J*, *K*, and *L* response keys, on which were placed the left ring, left middle, left index, right index, right middle, and right ring fingers, corresponded to the first through sixth target locations from the left, respectively.

On each trial, the target, a lowercase *o*, appeared 0.5 cm above one of the six lines, and participants pressed the corresponding response key. If the correct key was pressed, the target immediately disappeared. Otherwise, the target remained in its location until the correct key was pressed. After the target disappeared, it reappeared 250 ms later. RT was measured as the time between target appearance and the first response.

There was one session on each of 4 consecutive days. Each session was composed of 18 blocks of trials with 103 trials per block. Session 1 began with a practice block of 99 trials.

A performance history was provided at the end of each block. The numbers 1 to 18, corresponding to the number of blocks in a session, appeared vertically along the side of the screen. Beside the number for a completed block, one of two types of information was displayed. If 6% or more of the responses in the block were incorrect, the message *too many errors* and the error rate were displayed. Otherwise, a horizontal line, its length representing the average RT of correct responses, and the average RT were displayed. After a 10-s break, participants initiated the next block of trials at their discretion by pressing a key in response to a prompt on the screen.

Structure of the sequences of target locations

Each target location had two possible successors. For example, the appearance of the target in Location 1 could be followed by its appearance in Locations 3 or 4, and the appearance of the target in Location 2 could be followed by its appearance in Locations 1 or 6. Consequently, there were 24 (6 × 2 × 2) possible sequences of length 3 (contexts) each followed by two possible successors. Letting the numbers 1 to 6 represent the six target locations from left to right, respectively, Table 2 presents the contexts and the probabilities and frequencies with which successors followed contexts across every six blocks of trials. For example, row 1 indicates that context 3−2−1 was followed 18 times by Successor 3 and 12 times by Successor 4—that is, P(3|3−2−1) = .60 (high-probability transition, H) and P(4|3−2−1) = .40 (low-probability transition, L). Row 17 indicates that context 2−1−3 was followed 15 times by Successor 2 and 15 times by Successor 5—that is, P(2|2−1−3) = .50 (medium-probability transition, M) and P(5|2−1−3) = .50 (medium-probability transition, M).

In the first tier (rows 1−8), lag 2−1 and lag 2−*x* probabilities were confounded. For example, rows 1 and 2 indicate that 2−1 was followed 30 times by Successor 3 and 20 times by Successor 4 so that P(3|2−1) = .60 and P(4|2−1) = .40. Rows 1, 2, 5, and 6 indicate that the occurrence of location 2 on trial $t - 2$ was followed 60 times by Location 3 on trial t and 40 times by Location 4 on trial t so that P(3|2−*x*) = .60 and P(4|2−*x*) = .40. Thus P(3|2−1) = P(3|2−*x*) = .60 and P(4|2−1) = P(4|2−*x*) = .40. The first tier was labelled AN2 because second-order adjacent and nonadjacent dependencies were confounded.

In the second tier (rows 9−16), lag 2−1 probabilities varied whereas lag 2−*x* probabilities were constant. For example, rows 9 and 10 indicate that 3−2 was followed 30 times by Successor 1 and 20 times by Successor 6 so that P(1|3−2) = .60

Table 2. *Probabilities and frequencies inherent in the sequences of target locations across every six blocks of trials in Experiment 1*

			Successor					
		Context	*1*	*2*	*3*	*4*	*5*	*6*
Lag 2−1 + Lag 2−*x* (AN2)	1.	3−2−1	—	—	H (18)	L (12)	—	—
	2.	4−2−1	—	—	H (12)	L (8)	—	—
	3.	3−5−1	—	—	L (8)	H (12)	—	—
	4.	4−5−1	—	—	L (12)	H (18)	—	—
	5.	3−2−6	—	—	H (12)	L (8)	—	—
	6.	4−2−6	—	—	H (18)	L (12)	—	—
	7.	3−5−6	—	—	L (12)	H (18)	—	—
	8.	4−5−6	—	—	L (8)	H (12)	—	—
Lag 2−1 (A2)	9.	1−3−2	H (15)	—	—	—	—	L (10)
	10.	6−3−2	H (15)	—	—	—	—	L (10)
	11.	1−4−2	L (10)	—	—	—	—	H (15)
	12.	6−4−2	L (10)	—	—	—	—	H (15)
	13.	1−3−5	L (10)	—	—	—	—	H (15)
	14.	6−3−5	L (10)	—	—	—	—	H (15)
	15.	1−4−5	H (15)	—	—	—	—	L (10)
	16.	6−4−5	H (15)	—	—	—	—	L (10)
Neutral	17.	2−1−3	—	M (15)	—	—	M (15)	—
	18.	5−1−3	—	M (10)	—	—	M (10)	—
	19.	2−6−3	—	M (15)	—	—	M (15)	—
	20.	5−6−3	—	M (10)	—	—	M (10)	—
	21.	2−1−4	—	M (10)	—	—	M (10)	—
	22.	5−1−4	—	M (15)	—	—	M (15)	—
	23.	2−6−4	—	M (10)	—	—	M (10)	—
	24.	5−6−4	—	M (15)	—	—	M (15)	—

Note: AN2 = second-order adjacent and nonadjacent; A2 = second-order adjacent only; H = high-probability transition; L = low-probability transition; M = medium-probability transition.

and P(6|3−2) = .40. Rows 9, 10, 13, and 14 indicate that the occurrence of Location 3 on trial $t - 2$ was followed 50 times by Location 1 on trial t and 50 times by Location 6 on trial t so that P(1|3−x) = P(6|3−x) = .50. The second tier was labelled A2 because only second-order adjacent dependencies varied. In the third tier (rows 17−24), lag 2−1 and lag 2−x probabilities were constant.

The sequential structure was controlled so that certain types of information were not systematically confounded with lag 2−1 and lag 2−x probabilities. Each location was a target location equally often, lag 1 and lag 3−x−x probabilities were .50, and lag 3−x−1 probabilities were .48, .50, or .52. Moreover, lag 3−2−1 probabilities were redundant with lag 2−1 probabilities in tiers AN2 and A2, thus adding no information over and above that provided by the lag 2−1 probabilities—for example, P(3|4−5−1) = P(3|5−1) = .40. Lag 3−2−x probabilities were redundant with lag 2−x probabilities in tier AN2, thus adding no information over and above that provided by the lag 2−x probabilities—for example, P(3|4−5−x) = P(3|5−x) = .40. Finally, lag 3−2−x probabilities were .50 in tier A2.

Given the structure of the sequences of target locations, shorter RTs on A2−H than A2−L transitions would be evidence for learning the lag 2−1 probabilities. This result is expected based on prior research (e.g., Jimenez & Mendez, 1999; Remillard & Clark, 2001). More importantly, a larger RT difference between AN2−H and AN2−L transitions than between A2−H and A2−L transitions would be evidence for learning the lag 2−x probabilities.

For each participant and successive set of six blocks of trials, the sequence of target locations was generated by submitting the frequencies in Table 2 to a sequence-generation algorithm that randomly generated a 603-element sequence with the specified frequencies (Remillard & Clark, 1999). Elements 1−103, 101−203, 201−303, 301−403, 401−503, and 501−603 each constituted a block of 103 trials. For the practice block of 99 trials at the beginning of Session 1, the frequencies in Table 2 were replaced with the number 2. Thus the sequence of target locations in the practice block was unstructured in that all probabilities were .50.

There were 12 versions of Table 2. Version 1 was Table 2. Version 2 was formed from Version 1 by exchanging H and L transitions. Version 3 was created by having the top, middle, and bottom tiers of Table 2 describe A2, neutral, and AN2 transitions, respectively. Version 4 was formed from Version 3 by exchanging H and L transitions. Version 5 was created by having the top, middle, and bottom tiers of Table 2 describe neutral, AN2, and A2 transitions, respectively. Version 6 was formed from Version 5 by exchanging H and L transitions.

In Versions 1−6, neutral transitions were always followed by A2 transitions, which were always followed by AN2 transitions, which in turn were always followed by neutral transitions, and so on. For example, in the sequence 5−1−3−2−6−3−5, 5−1−3−2 is a neutral transition (refer to Table 2), 1−3−2−6 is an A2 transition, 3−2−6−3 is an AN2 transition, and 2−6−3−5 is a neutral transition. This ordering was reversed in Versions 7−12. Version 7 was created by having the top, middle, and bottom tiers of Table 2 describe AN2, neutral, and A2 transitions, respectively. Version 8 was formed from Version 7 by exchanging H and L transitions. Version 9 was created by having the top, middle, and bottom tiers of Table 2 describe neutral, A2, and AN2 transitions, respectively. Version 10 was formed from Version 9 by exchanging H and L transitions. Version 11 was created by having the top, middle, and bottom tiers of Table 2 describe A2, AN2, and neutral transitions, respectively. Version 12 was formed from Version 11 by exchanging H and L transitions. The frequencies for each version appear in Appendix A.

Awareness survey

The survey was a six-item paper-and-pencil test. The four items assessing awareness of the lag 2−1 probabilities were taken from tier A2, and the two items assessing awareness of the lag 2−x probabilities were taken from tier AN2. There were two options per item. For participants who received Version 1 of the sequential structure (i.e.,

Table 2), the four items for lag 2−1 probabilities were 3 → 2 → 1 6, 4 → 2 → 1 6, 3 → 5 → 1 6, and 4 → 5 → 1 6 (the numbers 1 and 6 in each item were arranged vertically in the survey and not horizontally as shown here). For each item, numbers represented target locations, and participants had to choose the high-probability transition. For example, the first item required an indication of whether the target letter *o* was more likely to appear in Location 1 or Location 6 after having appeared in Location 3 followed by Location 2. For participants' reference, the six target locations on the computer screen were marked with the short lines that were present during the SRT task. Scores greater than 50% correct (random guessing performance) on the four items would suggest awareness of the lag 2−1 probabilities.

For participants who received Version 1 of the sequential structure (i.e., Table 2), the two items for lag 2−*x* probabilities were 2 → x → 3 4 and 5 → x → 3 4 (the numbers 3 and 4 in each item were arranged vertically in the survey and not horizontally as shown here). Participants had to choose the high-probability transition. For example, the first item required an indication of whether the target letter *o* was more likely to appear in Location 3 or Location 4 after having appeared in Location 2 two trials earlier. Scores greater than 50% correct on the two items would suggest awareness of the lag 2−*x* probabilities.

Procedure

A total of 2 participants were randomly assigned to each of the 12 versions of the sequential structure. At the beginning of Session 1, the SRT task was described, and participants were instructed to try to improve their RT with practice while keeping their error rate below 6%. The structure underlying the sequence of target locations was not mentioned. Immediately following the last block of Session 4, the awareness survey was administered.

Data analysis

A number of factors other than transition probabilities can influence RTs. Table 3 lists four types of five-element runs on the basis of the first and second elements being equal (E) or unequal (U) to the fourth and fifth elements, respectively. RT to the last element may be shorter for EE runs, in which repetition of a bigram is correctly primed (e.g., 3−2−1−3 primes 2, and 2 occurs), than for UE runs (Remillard & Clark, 2001). Likewise, RT to the last element may be longer for EU runs, in which repetition of a bigram is incorrectly primed (e.g., 3−5−1−3 primes 5, but 2 occurs), than for UU runs. To ensure that L and H transitions were not differentially affected by the different run types, RTs were calculated as a function of type of run.

Table 3. *Types of run*

Run	*Example*
EE (fast)	3−2−1−3−2
UE (slow)	4−2−1−3−2
EU (slow)	3−5−1−3−2
UU (fast)	4−5−1−3−2

Note: Five-element runs were categorized as a function of the first and second elements being equal (E) or unequal (U) to the fourth and fifth elements, respectively. RT to the last element may be faster for EE than UE runs, and slower for EU than UU runs.

For each participant, the median RT of correct responses (excluding the first five trials of each block) was determined for each of the eight H and L transitions in a tier, and also as a function of the type of run that a transition completed. Four of the eight H transitions in a tier completed EE and UE runs, and the remaining four H transitions completed EU and UU runs. For example, 1−4−2−6 is an A2−H transition (with respect to Table 2) that completes the EU run 2−1−4−2−6 and the UU run 5−1−4−2−6. The sequence 6−3−5−6 is an A2−H transition that completes the EE run 5−6−3−5−6 and the UE run 2−6−3−5−6. The situation was similar for L transitions. Thus there were 16 data points for H transitions (8 H transitions with each completing 2 types of run) and 16 data points for L transitions (8 L transitions with each completing 2 types of run) in each of the tiers A2 and AN2.

RTs are also affected by specific hand transitions. For example, the four-element sequence

4−2−6−4 is an AN2−L transition (with respect to Table 2) involving two between-hand transitions (4−2 and 2−6) and one within-hand transition (6−4). I have found that different permutations of within-hand and between-hand transitions tend to produce different RTs to the last element of the sequence. Across all four-element sequences, there were 8 (2 × 2 × 2) possible permutations of within-hand and between-hand transitions. The eight permutations each occurred once across the eight H transitions and across the eight L transitions in tier A2. However, in tier AN2 the permutations WWW, BWW, BBB, and WBB (where W is within-hand and B is between-hand) each occurred twice across the eight H transitions, and the permutations WWB, BWB, BBW, and WBW each occurred twice across the eight L transitions for participants receiving Versions 1, 3, 5, 7, 9, and 11 of the sequential structure. The assignments were reversed for participants receiving the even-numbered versions of the sequential structure. This state of affairs can lead to RT differences between H and L transitions that differ for odd-numbered and even-numbered versions of the sequential structure. This can increase the error variability in an analysis of variance (ANOVA) when examining the effect of probability (H versus L) thereby reducing power. To remove the error variability, version (odd, even) was introduced as a between-subjects factor in the ANOVAs.

For A2 transitions, the 16 data points for H transitions were averaged as were the 16 data points for L transitions. The averaged scores were submitted to an ANOVA with probability (L, H) and session (1−4) as within-subject factors, and version (odd, even) as a between-subjects factor. A similar analysis was conducted for AN2 transitions. When comparing A2 and AN2 transitions, type (A2, AN2) was introduced as a within-subject factor. Analyses involving the session factor focused on the linear component of session (session L) because decreases in RT across sessions were of interest. Alpha was .05.

The effect size measure, r_{contrast}, is reported for all F-tests with numerator $df = 1$ (Rosenthal, Rosnow, & Rubin, 2000). I also report 95% confidence intervals for some of the tests. A confidence interval for a contrast was constructed using the *MSE* from the corresponding F-test and is reported in the form X ± Y.

Results and discussion

The results appear in Figure 1. There was a Probability × Session L interaction for AN2 transitions, $F(1, 22) = 6.17$, $p = .021$, $r_{\text{contrast}} = .47$, and A2 transitions, $F(1, 22) = 6.99$, $p = .015$, $r_{\text{contrast}} = .49$. Thus the RT difference between H and L transitions increased across sessions for both types of transitions. Limiting the analysis to Sessions 3 and 4 (the last half of training), there was a Probability × Type interaction, $F(1, 22) = 8.94$, $p = .007$, $r_{\text{contrast}} = .54$, 8.6 ms ± 5.9 ms. This reflects a larger L − H difference for AN2 than A2 transitions and indicates that there was learning of the lag 2−x probabilities. Finally, RT was shorter on A2−H than A2−L transitions across Sessions 3 and 4, $F(1, 22) = 93.66$, $p < .001$, $r_{\text{contrast}} = .90$, 12.2 ms ± 2.6 ms. This indicates that there was learning of the lag 2−1 probabilities.[2]

Error rates were also examined. The L − H difference increased from Session 1 to Session 4 for AN2 transitions (0.6% to 2.9%), $F(1, 22) = 5.79$, $p = .025$, $r_{\text{contrast}} = .46$, and A2 transitions (− 0.4% to 1.8%), $F(1, 22) = 15.72$, $p = .001$, $r_{\text{contrast}} = .65$. Although the L − H difference was numerically greater for AN2 than A2

[2] Limiting the analysis to Sessions 3 and 4, the Probability × Version interaction was significant for AN2 transitions, $F(1, 22) = 16.32$, $p = .001$, $r_{\text{contrast}} = .65$, but not A2 transitions, $F(1, 22) < 1$, $r_{\text{contrast}} = .09$. For AN2 transitions, the RT difference between H and L transitions was greater for odd-numbered (33 ms) than even-numbered (8 ms) versions of the sequential structure. This is because for AN2 transitions (but not A2 transitions) the specific permutations of within-hand and between-hand transitions comprising H and L transitions were different for odd-numbered and even-numbered versions (see Data Analysis subsection in Method section). Clearly, specific permutations of within-hand and between-hand transitions can have a strong influence on RTs.

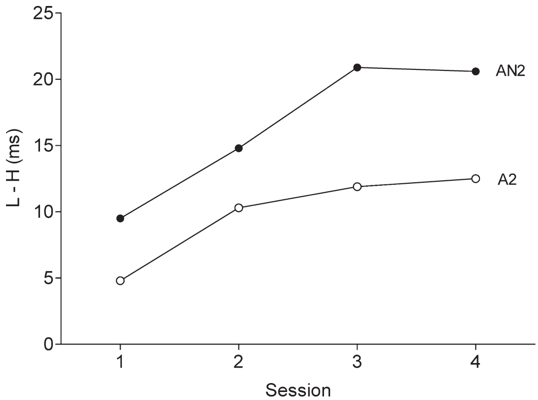

Figure 1. *The difference in reaction time between H and L transitions as a function of type of transition (AN2, A2) and session (1–4) in Experiment 1. AN2 = second-order adjacent and nonadjacent; A2 = second-order adjacent only; L = low probability; H = high probability.*

transitions in Sessions 3 and 4, the advantage was not significant, $F(1, 22) < 1$, $r_{\text{contrast}} = .17$.

Awareness

The percentage of the four (two) survey items pertaining to lag 2−1 (lag 2−x) probabilities receiving correct responses (i.e., for which H transitions were chosen) was determined for each participant. The mean percentages were 50.0 and 43.8 for lag 2−1 and lag 2−x probabilities, respectively. The percentages did not differ significantly from what would be expected by random guessing (50%); $F(1, 23) < 1$, $r_{\text{contrast}} = 0$, 50.0% $\pm$ 9.8% for lag 2−1 probabilities, and $F(1, 23) < 1$, $r_{\text{contrast}} = .15$, 43.8% $\pm$ 18.0% for lag 2−x probabilities. Thus there was no evidence for awareness of the lag 2−1 and lag 2−x probabilities.

EXPERIMENT 2

The previous experiment produced solid evidence that people can implicitly learn lag 2−x probabilities. Experiment 2 examined implicit learning of third-order adjacent and nonadjacent dependencies of .60 versus .40. The approach was similar to that in Experiment 1. For example, consider the two runs 1−3−5−1 and 1−3−5−6 where P(1|1−3−5) = .60, P(6|1−3−5) = .40, and P(1|1−3−x) = P(6|1−3−x) = P(1|1−x−5) = P(6|1−x−5) = P(1|1−x−x) = P(6|1−x−x) = .50. Lag 3−2−1 probabilities (i.e., adjacent dependencies) vary across the two runs whereas lag 3−2−x, lag 3−x−1, and lag 3−x−x probabilities (i.e., nonadjacent dependencies) remain constant. Now consider the two runs 5−1−3−2 and 5−1−3−5 where P(2|5−1−3) = P(2|5−1−x) = P(2|5−x−3) = P(2|5−x−x) = .60, and P(5|5−1−3) = P(5|5−1−x) = P(5|5−x−3) = P(5|5−x−x) = .40. Adjacent and nonadjacent dependencies are confounded, and both vary across the two runs. If people are incapable of learning nonadjacent dependencies, then the RT difference between high- and low-probability transitions should be similar in the two sets of runs. Therefore, a greater RT difference between high- and low-probability transitions for the second set of runs than the first set of runs would implicate learning of at least one of the three nonadjacent dependencies. Establishing which of the three nonadjacent dependencies are learned is a task for future research. Experiment 2 also examined learning of lag 4−3−2−1 probabilities (i.e., fourth-order adjacent dependencies).

Method

Participants

The participants were 12 Morehead State University (Morehead, Kentucky) undergraduates ranging in age from 18 to 24 years.

SRT task

The SRT task was identical to that in Experiment 1 except that the response keys were *X*, *C*, *V*, *M*, <, and >, and the response–stimulus interval was 300 ms. There were 10 sessions. Each session was composed of 16 blocks of trials with 123 trials per block. Session 1 began with a practice block of 100 trials. On a given day, there were 0, 1, or 2 sessions (with at least 60 min between sessions). There were never more than 3 consecutive zero-session days. The 10 sessions were completed in 10 to 19 days. As in Experiment 1, a performance history was provided at the end of each block of trials.

Structure of the sequences of target locations

Table 4 presents the probabilities that were inherent in the sequences of target locations. For example, row 1 indicates that every 15 occurrences of context 1−3−2−1 was followed 9 times by Successor 4 and 6 times by Successor 3—that is, P(4|1−3−2−1) = .60 (high-probability transition, H), and P(3|1−3−2−1) = .40 (low-probability transition, L). Row 25 indicates that every 15 occurrences of context 2−1−3−5 was followed 9 times by Successor 1 and 6 times by Successor 6—that is, P(1|2−1−3−5) = .60, and P(6|2−1−3−5) = .40.

In the first tier (rows 1–16), lag 4−3−2−1 probabilities were .60 or .40—for example, P(4|1−3−2−1) = .60 and P(3|1−3−2−1) = .40. First-order, second-order, and third-order dependencies were .48, .50, or .52 and were not confounded with lag 4−3−2−1 probabilities.[3] Similarly, fourth-order nonadjacent dependencies were .48, .50, or .52 and were not confounded with lag 4−3−2−1 probabilities. For example, rows 1, 3, 5, 7, 9, 11, 13, and 15 indicate that when Location 1 occurred on trial $t-4$, Locations 3 and 4 on trial t were each an H transition 4 times and an L transition 4 times. Consequently, P(3|1−x−x−x) should be approximately equal to P(4|1−x−x−x). Thus shorter RTs on H than L transitions would be evidence for learning the lag 4−3−2−1 probabilities. The first tier was labelled A4 because only fourth-order adjacent dependencies varied.

In the second tier (rows 17–32), lag 3−2−1 probabilities were .60 or .40—for example, from rows 21 and 22, P(1|1−4−2) = .60 and P(6|1−4−2) = .40. Lag 4−3−2−1 probabilities were redundant with lag 3−2−1 probabilities, thus adding no information over and above that provided by the lag 3−2−1 probabilities—for example, P(1|5−1−4−2) = P(1|1−4−2) = .60. All other probabilities were .48, .50, or .52 and were not confounded with lag 3−2−1 probabilities. Thus shorter RTs on H than L transitions would be evidence for learning the lag 3−2−1 probabilities. The second tier was labelled A3 because only third-order adjacent dependencies varied.

In the third tier (rows 33–48), lag 3−2−1 probabilities (i.e., adjacent dependencies) were .60 or .40, and lag 3−2−x, lag 3−x−1, and lag 3−x−x probabilities (i.e., nonadjacent dependencies) were confounded with lag 3−2−1 probabilities—for example, P(2|5−1−3) = P(2|5−x−3) = P(2|5−1−x) = P(2|5−x−x) = .60. For example, rows 35, 36, 43, and 44 indicate that when Locations 5 and 1 occurred on trials $t-3$ and $t-2$, respectively, Location 2 on trial t was an H transition every time so that P(2|5−1−x) = .60. Lag 4−3−2−1, lag 4−3−x−x, lag 4−3−x−1, and lag 4−3−2−x probabilities were redundant with lag 3−2−1, lag 3−x−x, lag 3−x−1, and lag 3−2−x probabilities, respectively, thus adding no information over and above that provided by the third-order dependencies—for example, P(2|4−5−x−3) = P(2|5−x−3) = .60. All other probabilities were .48, .50, or .52 and were not confounded with the third-order dependencies. The third tier was labelled AN3 because third-order adjacent and nonadjacent dependencies were confounded. A larger RT difference between AN3−H and AN3−L transitions than between A3−H and A3−L transitions would be evidence for learning at least 1 of the 3 third-order nonadjacent dependencies.

For each participant, a 19,044-element sequence of target locations was randomly generated with the constraint that across every 15 occurrences of a context (e.g., 1−3−2−1), one successor (e.g., 4) occurred 9 times, and the other successor (e.g., 3) occurred 6 times. Elements 1−123, 120−242, 239−361, . . . , and 18,922−19,044 each constituted a block of 123 trials for a total of 160 blocks (10 sessions × 16 blocks per session). The practice block of 100 trials at the beginning of Session 1 was randomly generated with the constraint that each context in Table 4 was followed by each of its two possible successors once. Thus the sequence of target locations in the practice block was unstructured in that all probabilities were .50.

[3] After the sequences of target locations had been generated, a computer program went over the sequences and determined the exact values of all first- through fourth-order dependencies.

Table 4. *Probabilities inherent in the sequences of target locations in Experiment 2*

			Successor					
		Context	*1*	*2*	*3*	*4*	*5*	*6*
Lag 4–3–2–1 (A4)	1.	1–3–2–1	—	—	L	H	—	—
	2.	6–3–2–1	—	—	H	L	—	—
	3.	1–4–2–1	—	—	H	L	—	—
	4.	6–4–2–1	—	—	L	H	—	—
	5.	1–3–5–1	—	—	H	L	—	—
	6.	6–3–5–1	—	—	L	H	—	—
	7.	1–4–5–1	—	—	L	H	—	—
	8.	6–4–5–1	—	—	H	L	—	—
	9.	1–3–2–6	—	—	H	L	—	—
	10.	6–3–2–6	—	—	L	H	—	—
	11.	1–4–2–6	—	—	L	H	—	—
	12.	6–4–2–6	—	—	H	L	—	—
	13.	1–3–5–6	—	—	L	H	—	—
	14.	6–3–5–6	—	—	H	L	—	—
	15.	1–4–5–6	—	—	H	L	—	—
	16.	6–4–5–6	—	—	L	H	—	—
Lag 3–2–1 (A3)	17.	2–1–3–2	L	—	—	—	—	H
	18.	5–1–3–2	L	—	—	—	—	H
	19.	2–6–3–2	H	—	—	—	—	L
	20.	5–6–3–2	H	—	—	—	—	L
	21.	2–1–4–2	H	—	—	—	—	L
	22.	5–1–4–2	H	—	—	—	—	L
	23.	2–6–4–2	L	—	—	—	—	H
	24.	5–6–4–2	L	—	—	—	—	H
	25.	2–1–3–5	H	—	—	—	—	L
	26.	5–1–3–5	H	—	—	—	—	L
	27.	2–6–3–5	L	—	—	—	—	H
	28.	5–6–3–5	L	—	—	—	—	H
	29.	2–1–4–5	L	—	—	—	—	H
	30.	5–1–4–5	L	—	—	—	—	H
	31.	2–6–4–5	H	—	—	—	—	L
	32.	5–6–4–5	H	—	—	—	—	L
Lag 3–2–1 + Lag 3–2–x + Lag 3–x–1 + Lag 3–x–x (AN3)	33.	3–2–1–3	—	L	—	—	H	—
	34.	4–2–1–3	—	L	—	—	H	—
	35.	3–5–1–3	—	H	—	—	L	—
	36.	4–5–1–3	—	H	—	—	L	—
	37.	3–2–6–3	—	L	—	—	H	—
	38.	4–2–6–3	—	L	—	—	H	—
	39.	3–5–6–3	—	H	—	—	L	—
	40.	4–5–6–3	—	H	—	—	L	—
	41.	3–2–1–4	—	L	—	—	H	—
	42.	4–2–1–4	—	L	—	—	H	—
	43.	3–5–1–4	—	H	—	—	L	—
	44.	4–5–1–4	—	H	—	—	L	—
	45.	3–2–6–4	—	L	—	—	H	—
	46.	4–2–6–4	—	L	—	—	H	—
	47.	3–5–6–4	—	H	—	—	L	—
	48.	4–5–6–4	—	H	—	—	L	—

Note: Every 15 occurrences of a context (e.g., 4–5–1–3 in row 36) was followed by one successor (e.g., 2) 9 times and by the other successor (e.g., 5) 6 times. A4 = fourth-order adjacent only; A3 = third-order adjacent only; AN3 = third-order adjacent and nonadjacent; H = high-probability transition; L = low-probability transition.

There were 12 versions of Table 4. These were generated in a manner similar to that described in Experiment 1.

Awareness survey

The survey was a 16-item paper-and-pencil test. For A3 transitions, 8 items assessed awareness of the lag 3−2−1 probabilities. There were two options per item. For participants who received Version 1 of the sequential structure (i.e., Table 4), 2 of the 8 items were 1 → 3 → 2 → 1 6 and 6 → 3 → 2 → 1 6 (the last two numbers, 1 and 6, in each item were arranged vertically in the survey and not horizontally as shown here). For each item, numbers represented target locations, and participants had to choose the high-probability transition. For example, the first item required an indication of whether the target letter *o* was more likely to appear in Location 1 or Location 6 after having appeared in Location 1 followed by Location 3 followed by Location 2. For participants' reference, the six target locations on the computer screen were marked with the short lines that were present during the SRT task. Scores greater than 50% correct (random guessing performance) on the 8 items would suggest awareness of the lag 3−2−1 probabilities.

For AN3 transitions, 8 items assessed awareness of the nonadjacent and adjacent dependencies. For participants who received Version 1 of the sequential structure (i.e., Table 4), 2 of the 8 items were 2 → 1 → 3 → 2 5 and 5 → 1 → 3 → 2 5 (the last two numbers, 2 and 5, in each item were arranged vertically in the survey and not horizontally as shown here). Scores greater than 50% correct on the 8 items would suggest awareness of the nonadjacent or adjacent dependencies associated with AN3 transitions. However, it would not be possible to establish the exact nature of the dependencies of which participants were aware. The 16 survey items were mixed such that items for A3 transitions alternated with items for AN3 transitions.

Procedure

One participant was randomly assigned to each of the 12 versions of the sequential structure. At the beginning of session 1, the SRT task was described, and participants were instructed to try to improve their RT with practice while keeping their error rate below 6%. The structure underlying the sequence of target locations was not mentioned. Immediately following the last block of session 10, the awareness survey was administered.

Data analysis

The eight permutations of within-hand and between-hand transitions each occurred twice across the 16 H transitions and across the 16 L transitions in tier A4. For example, 1−3−2−1−4 is an H transition (row 1 in Table 4) with 3−2, 2−1, and 1−4 being within-hand, within-hand, and between-hand transitions, respectively (i.e., a WWB permutation). However, in tier A3 the permutations WWB, BWW, BBW, and WBB each occurred four times across the 16 H transitions, and the permutations WWW, BWB, BBB, and WBW each occurred four times across the 16 L transitions for participants receiving Versions 1, 3, 5, 7, 9, and 11 of the sequential structure. The assignments were reversed for participants receiving the even-numbered versions of the sequential structure. Finally, in tier AN3 the permutations WWB, BWW, BBB, and WBW each occurred four times across the 16 H transitions, and the permutations WWW, BWB, BBW, and WBB each occurred four times across the 16 L transitions for participants receiving Versions 1, 3, 5, 7, 9, and 11 of the sequential structure. The assignments were reversed for participants receiving the even-numbered versions of the sequential structure. Consequently, version (odd, even) was introduced as a between-subjects factor in the ANOVAs for reasons outlined in Experiment 1.

The four types of five-element runs in Table 3 were each completed by 4 H transitions and 4 L transitions in tiers A4 and A3. In tier AN3, EU and UU runs were each completed by 8 H transitions, and EE and UE runs were each completed by 8 L transitions for participants receiving the odd-numbered versions of the sequential structure. The assignments were reversed for participants receiving the even-numbered versions of the sequential structure. Thus RTs on L and H transitions were not differentially affected by the different run types within a version for tiers A4 and A3 and across versions for tier AN3.

For each participant, the median RT of correct responses (excluding the first five trials of each block) was determined for each of the 16 H and L transitions in a tier. The 16 data points for H transitions in a tier were averaged, as were the 16 data points for L transitions. The averaged scores were submitted to an ANOVA with probability (L, H) and session (1–10) as within-subject factors and version (odd, even) as a between-subjects factor. When comparing A3 and AN3 transitions, type (A3, AN3) was introduced as a within-subject factor. Analyses involving the session factor focused on the linear component of session (session L).

A3 and A4 transitions were also compared. In tier A3, lag 4–3–2–1 and lag 3–2–1 probabilities were confounded. In tier A4, lag 3–2–1 probabilities were constant. Thus it was expected that the RT difference between A3–H and A3–L transitions would be greater than that between A4–H and A4–L transitions. When comparing A3 and A4 transitions, type (A3, A4) was introduced as a within-subject factor. Alpha was .05. Effect sizes and confidence intervals are reported as in Experiment 1.

Results and discussion

The results appear in Figure 2. The Probability × Session L interaction was not significant for AN3 transitions, $F(1, 10) = 1.36$, $p = .270$, $r_{\text{contrast}} = .35$, nor for A3 transitions, $F(1, 10) < 1$, $r_{\text{contrast}} = .17$, and A4 transitions, $F(1, 10) = 1.30$, $p = .281$, $r_{\text{contrast}} = .34$. Thus the RT difference between H and L transitions did not increase

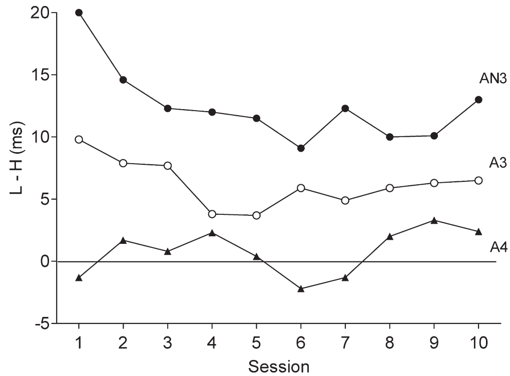

Figure 2. *The difference in reaction time between H and L transitions as a function of type of transition (AN3, A3, A4) and session (1–10) in Experiment 2. AN3 = third-order adjacent and nonadjacent; A3 = third-order adjacent only; A4 = fourth-order adjacent only; L = low probability; H = high probability.*

significantly across sessions for any of the three types of transition. The analyses that follow are limited to Sessions 6–10 (the last half of training).

There was a Probability × Type (AN3, A3) interaction, $F(1, 10) = 5.42$, $p = .042$, $r_{\text{contrast}} = .59$, 5.0 ms ± 4.8 ms. This reflects a larger L − H difference for AN3 than A3 transitions and indicates that there was learning of at least 1 of the 3 third-order nonadjacent dependencies. RT was shorter on A3–H than A3–L transitions, $F(1, 10) = 14.65$, $p = .003$, $r_{\text{contrast}} = .77$, 5.9 ms ± 3.4 ms. This indicates that there was learning of the lag 3–2–1 probabilities.

There was a Probability × Type (A3, A4) interaction, $F(1, 10) = 11.93$, $p = .006$, $r_{\text{contrast}} = .74$, 5.0 ms ± 3.3 ms. This reflects a larger L − H difference for A3 than A4 transitions. The RT difference between A4–H and A4–L transitions was not significant, $F(1, 10) = 3.18$, $p = .105$, $r_{\text{contrast}} = .49$, 0.8 ms ± 1.0 ms. Thus there was no evidence for learning of the lag 4–3–2–1 probabilities.[4]

[4] Limiting the analysis to Sessions 6–10, the Probability × Version interaction was significant for A3 transitions, $F(1, 10) = 7.47$, $p = .021$, $r_{\text{contrast}} = .65$, marginally significant for AN3 transitions, $F(1, 10) = 4.57$, $p = .058$, $r_{\text{contrast}} = .56$, and not significant for A4 transitions, $F(1, 10) < 1$, $r_{\text{contrast}} = .21$. For A3 and AN3 transitions, the RT difference between H and L transitions was greater for even-numbered than odd-numbered versions (10 ms vs. 2 ms for A3 transitions, 16 ms vs. 6 ms for AN3 transitions). This is because for A3 and AN3 transitions (but not A4 transitions) the specific permutations of within-hand and between-hand transitions comprising H and L transitions were different for odd-numbered and even-numbered versions. Again, specific permutations of within-hand and between-hand transitions can exert a strong influence on RTs.

Error rates were also examined. The L − H difference increased significantly from Session 1 to Session 10 for AN3 transitions (1.1% to 2.8%), $F(1, 10) = 5.21$, $p = .046$, $r_{\text{contrast}} = .59$, but not for A3 transitions (1.4% to 1.7%), $F(1, 10) = 1.76$, $p = .214$, $r_{\text{contrast}} = .39$, and A4 transitions (− 0.1% to −1.0%), $F(1, 10) < 1$, $r_{\text{contrast}} = .13$. The L − H difference was numerically greater for AN3 than A3 transitions across Sessions 6–10, and the advantage approached significance, $F(1, 10) = 3.37$, $p = .096$, $r_{\text{contrast}} = .50$.

Awareness

For A3 and AN3 transitions, the percentage of the eight survey items receiving correct responses (i.e., for which H transitions were chosen) was determined for each participant. The mean percentages were 46.9 and 58.3 for A3 and AN3 transitions, respectively. The percentage did not differ significantly from what would be expected by random guessing (50%) for A3 transitions, $F(1, 11) < 1$, $r_{\text{contrast}} = .21$, 46.9% ± 9.7%, and AN3 transitions, $F(1, 11) = 1.91$, $p = .194$, $r_{\text{contrast}} = .38$, 58.3% ± 13.3%. Thus there was no evidence for awareness of the third-order adjacent and nonadjacent dependencies.

EXPERIMENT 3

The previous experiment showed that people can implicitly learn third-order adjacent and nonadjacent dependencies. Experiment 3 examined learning of fourth-order adjacent and nonadjacent dependencies of .67 versus .33. The probabilities were widened because the results from Experiment 2 suggested that lag 4−3−2−1 probabilities of .60 versus .40 might be difficult to learn. As in the previous experiments, the approach consisted of comparing the L − H difference in RT for AN transitions to the L − H difference in RT for A transitions.

For a given version of the sequential structure in Experiments 1 and 2, the permutations of within-hand and between-hand transitions comprising H transitions were different for A and AN transitions, as were the permutations comprising L transitions. The assignment of permutations to H and L transitions was counterbalanced across the 12 versions of the sequential structure. It would be desirable to have the permutations comprising H and L transitions be the same for A and AN transitions within each version. This could potentially reduce error variability and increase power when comparing L − H differences across A and AN transitions. Experiment 3 equated A and AN transitions with respect to the permutations comprising H and L transitions.

Method

Participants

The participants were 12 Morehead State University (Morehead, Kentucky) undergraduates ranging in age from 18 to 21 years.

SRT task

The SRT task was identical to that in Experiment 2. There were 10 sessions. Each session was composed of 16 blocks of trials with 123 trials per block. Session 1 began with a practice block of 100 trials. On a given day, there were 0, 1, or 2 sessions (with at least 60 min between sessions). There were never more than 3 consecutive zero-session days. The 10 sessions were completed in 11 to 18 days.

Structure of the sequences of target locations

Table 5 presents the probabilities that were inherent in the sequences of target locations. For example, row 1 indicates that every 15 occurrences of context 1−3−2−1 was followed 10 times by Successor 4 and 5 times by Successor 3—that is, $P(4|1-3-2-1) = .67$ (high-probability transition, H), and $P(3|1-3-2-1) = .33$ (low-probability transition, L). The second and third tiers contain uppercase and lowercase letters. Both H and h reflect high-probability transitions, and both L and l reflect low-probability transitions. The meaning of the different cases are explained below.

In the first tier (rows 1–16), lag 4−3−2−1 probabilities were .67 or .33—for example, $P(4|1-3-2-1) = .67$, and $P(3|1-3-2-1) = .33$. First-order, second-order, and third-order dependencies were .50, as were fourth-order nonadjacent dependencies. Thus shorter RTs on H than L transitions would be evidence for learning the lag 4−3−2−1 probabilities. The first tier was labelled A4 because only fourth-order adjacent dependencies varied.

In the second tier (rows 17−32), lag 4−3−2−1 probabilities (i.e., adjacent dependencies) were .67 or .33. Fourth-order nonadjacent dependencies were confounded with lag 4−3−2−1 probabilities. For example, row 17 indicates that P(1|2−1−3−2) = .33. The odd-numbered rows indicate that when Location 2 occurred on trial $t - 4$, Location 1 on trial t was an L (or l) transition every time so that P(1|2−x−x−x) = .33. Thus P(1|2−1−3−2) = P(2−x−x−x) = .33. First-order, second-order, and third-order dependencies were .50. The second tier was labelled AN4 because fourth-order adjacent and nonadjacent dependencies were confounded. The third tier (rows 33−48) was similar to the second tier.

The second and third tiers contain the uppercase letters H and L and the lowercase letters h and l. RTs were analysed only for H and L transitions, and not for h and l transitions. This is because H and L transitions in tiers AN4 were matched to H and L transitions in tier A4 with respect to the pattern of within-hand and between-hand transitions. For example, the AN4−H transition 3−5−1−3−2 in row 35 is matched to the A4−H transition 1−4−2−1−3 in row 3, and both involve the same pattern of within-hand and between-hand transitions—for example, 3−5 and 1−4 (between-hand, left hand to right land), 5−1 and 4−2 (between-hand, right hand to left hand), 1−3 and 2−1 (within-hand, left hand), 3−2 and 1−3 (within-hand, left hand). Tiers AN4 rows 17, 18, 35, 36, 37, 38, 23, 24, 41, 42, 27, 28, 29, 30, 47, 48 were matched to tier A4 rows 1−16, respectively. A larger RT difference between AN4−H and AN4−L transitions than between A4−H and A4−L transitions would be evidence for learning at least 1 of the 7 fourth-order nonadjacent dependencies. Establishing which of the nonadjacent dependencies are learned is a task for future research.

The sequences of target locations were generated as in Experiment 2 except that across every 15 occurrences of a context (e.g., 1−3−2−1), one successor (e.g., 4) occurred 10 times, and the other successor (e.g., 3) occurred 5 times. There were 12 versions of Table 5. These were generated in a manner similar to that described in Experiment 1. However, one tier described A4 transitions, and two tiers described AN4 transitions in Versions 1−6, whereas one tier described AN4 transitions, and two tiers described A4 transitions in Versions 7−12. The probabilities for each version appear in Appendix B.

Awareness measure

To keep the measure relatively short, only awareness of lag 4−3−2−1 probabilities associated with A4 transitions was assessed. Awareness of the fourth-order nonadjacent dependencies associated with AN4 transitions was not assessed. The measure consisted of 16 prediction trials. Each trial began with the appearance of the six horizontal lines marking the six target locations. One second later, participants observed the target move across four locations (e.g., 1−3−2−1 from row 1 of Table 5) followed by a carat symbol (∧) that appeared below each of the two possible successors (e.g., 3 and 4). Participants indicated at which of the two marked locations the target was most likely to have appeared next during training by pressing the corresponding key. The sequence of target movements could be repeated by pressing the space bar. In a sequence of target movements, target duration and the interstimulus interval were each 300 ms. Following a prediction response, six *X*s overwrote the six horizontal lines for a duration of 2 s. The next trial then began.

For participants receiving Versions 1−6 of the sequential structure, the 16 prediction trials were the 16 rows of tier A4. For participants receiving Versions 7−12, the 16 prediction trials were supposed to be the 8 rows with uppercase letters H and L in each of the two A4 tiers. However, due to human error, the 16 prediction trials were the 16 rows from one of the two A4 tiers. This is not a serious error because awareness of the lag 4−3−2−1 probabilities associated with A4 transitions is still being assessed. The 16 prediction trials were presented in a random order for each participant. Scores greater than 50% correct

Table 5. *Probabilities inherent in the sequences of target locations in Experiment 3*

			Successor					
		Context	*1*	*2*	*3*	*4*	*5*	*6*
Lag 4–3–2–1 (A4)	1.	1–3–2–1	—	—	L	H	—	—
	2.	6–3–2–1	—	—	H	L	—	—
	3.	1–4–2–1	—	—	H	L	—	—
	4.	6–4–2–1	—	—	L	H	—	—
	5.	1–3–5–1	—	—	H	L	—	—
	6.	6–3–5–1	—	—	L	H	—	—
	7.	1–4–5–1	—	—	L	H	—	—
	8.	6–4–5–1	—	—	H	L	—	—
	9.	1–3–2–6	—	—	H	L	—	—
	10.	6–3–2–6	—	—	L	H	—	—
	11.	1–4–2–6	—	—	L	H	—	—
	12.	6–4–2–6	—	—	H	L	—	—
	13.	1–3–5–6	—	—	L	H	—	—
	14.	6–3–5–6	—	—	H	L	—	—
	15.	1–4–5–6	—	—	H	L	—	—
	16.	6–4–5–6	—	—	L	H	—	—
Lag 4–3–2–1 + Nonadjacent (AN4)	17.	2–1–3–2	L	—	—	—	—	H
	18.	5–1–3–2	H	—	—	—	—	L
	19.	2–6–3–2	l	—	—	—	—	h
	20.	5–6–3–2	h	—	—	—	—	l
	21.	2–1–4–2	l	—	—	—	—	h
	22.	5–1–4–2	h	—	—	—	—	l
	23.	2–6–4–2	L	—	—	—	—	H
	24.	5–6–4–2	H	—	—	—	—	L
	25.	2–1–3–5	l	—	—	—	—	h
	26.	5–1–3–5	h	—	—	—	—	l
	27.	2–6–3–5	L	—	—	—	—	H
	28.	5–6–3–5	H	—	—	—	—	L
	29.	2–1–4–5	L	—	—	—	—	H
	30.	5–1–4–5	H	—	—	—	—	L
	31.	2–6–4–5	l	—	—	—	—	h
	32.	5–6–4–5	h	—	—	—	—	l
Lag 4–3–2–1 + Nonadjacent (AN4)	33.	3–2–1–3	—	h	—	—	l	—
	34.	4–2–1–3	—	l	—	—	h	—
	35.	3–5–1–3	—	H	—	—	L	—
	36.	4–5–1–3	—	L	—	—	H	—
	37.	3–2–6–3	—	H	—	—	L	—
	38.	4–2–6–3	—	L	—	—	H	—
	39.	3–5–6–3	—	h	—	—	l	—
	40.	4–5–6–3	—	l	—	—	h	—
	41.	3–2–1–4	—	H	—	—	L	—
	42.	4–2–1–4	—	L	—	—	H	—
	43.	3–5–1–4	—	h	—	—	l	—
	44.	4–5–1–4	—	l	—	—	h	—
	45.	3–2–6–4	—	h	—	—	l	—
	46.	4–2–6–4	—	l	—	—	h	—
	47.	3–5–6–4	—	H	—	—	L	—
	48.	4–5–6–4	—	L	—	—	H	—

Note: Every 15 occurrences of a context (e.g., 4–5–1–3 in row 36) was followed by one successor (e.g., 2) 5 times and by the other successor (e.g., 5) 10 times. AN4 = fourth-order adjacent and nonadjacent; A4 = fourth-order adjacent only; H, h = high-probability transition; L, l = low-probability transition. Reaction times and error rates on h and l transitions were not analysed.

(random guessing performance) on the 16 trials would suggest an awareness of the lag 4–3–2–1 probabilities.

Procedure
The procedure followed that of Experiment 2.

Data analysis
The eight permutations of within-hand and between-hand transitions each occurred twice across the 16 H transitions and across the 16 L transitions for tiers A4 and AN4. This was the case for each version of the sequential structure. Consequently, RT differences between H and L transitions should not be very different for odd-numbered and even-numbered versions. Thus version (odd, even) was not a between-subjects factor in the ANOVAs. This has the advantage of increasing the denominator degrees of freedom and therefore power.

The four types of five-element runs in Table 3 were each completed by 4 H transitions and 4 L transitions in tiers A4 and AN4. Thus RTs on L and H transitions were not differentially affected by the different run types within a version.

For each participant, the median RT of correct responses (excluding the first 5 trials of each block) was determined for each of the 16 H and L transitions in a tier. The 16 data points for H transitions in a tier were averaged, as were the 16 data points for L transitions. The averaged scores were submitted to an ANOVA with probability (L, H) and session (1—10) as within-subject factors. When comparing A4 and AN4 transitions, type (A4, AN4) was introduced as a within-subject factor. Analyses involving the session factor focused on the linear component of session (session L). Alpha was .05. Effect sizes and confidence intervals are reported as in Experiment 1.

Results and discussion

The results appear in Figure 3. The Probability × Session L interaction was not significant for AN4 transitions, $F(1, 11) = 1.98$, $p = .187$, $r_{\text{contrast}} = .39$, nor for A4 transitions, $F(1, 11) < 1$, $r_{\text{contrast}} = .09$. Thus the RT difference between H and L

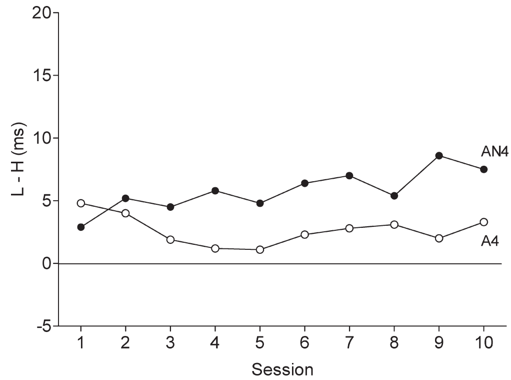

Figure 3. *The difference in reaction time between H and L transitions as a function of type of transition (AN4, A4) and session (1–10) in Experiment 3. AN4 = fourth-order adjacent and nonadjacent; A4 = fourth-order adjacent only; L = low probability; H = high probability.*

transitions did not increase significantly across sessions for the two types of transition. The analyses that follow are limited to Sessions 6–10 (the last half of training).

There was a Probability × Type interaction, $F(1, 11) = 11.88$, $p = .005$, $r_{\text{contrast}} = .72$, 4.2 ms $\pm$ 2.7 ms. This reflects a larger L − H difference for AN4 than A4 transitions and indicates that there was learning of at least 1 of the 7 fourth-order nonadjacent dependencies. RT was shorter on A4–H than A4–L transitions, $F(1, 11) = 12.34$, $p = .005$, $r_{\text{contrast}} = .73$, 2.7 ms $\pm$ 1.7 ms. This indicates that there was learning of the lag 4–3–2–1 probabilities.

Error rates were also examined. The L − H difference increased significantly from Session 1 to Session 10 for AN4 transitions (−0.1% to 3.0%), $F(1, 11) = 24.71$, $p < .001$, $r_{\text{contrast}} = .83$, but not for A4 transitions (0% to 0.4%), $F(1, 11) = 1.27$, $p = .283$, $r_{\text{contrast}} = .32$. The L − H difference was greater for AN4 than A4 transitions across Sessions 6–10, $F(1, 11) = 11.47$, $p = .006$, $r_{\text{contrast}} = .71$.

Awareness
For A4 transitions, the percentage of the 16 prediction trials receiving correct responses (i.e., for which H transitions were chosen) was determined

for each participant. The mean percentage was 51.0, which did not differ significantly from what would be expected by random guessing (50%), $F(1, 11) < 1$, $r_{\text{contrast}} = .09$, 51.0 $\pm$ 7.5%. Thus there was no evidence for awareness of the lag 4−3−2−1 probabilities.

GENERAL DISCUSSION

Prior research has established that people can learn first- and second-order adjacent dependencies, as well as third-order information. In the case of third-order information, however, it was not clear whether adjacent or nonadjacent dependencies were learned because the two types of information were confounded. The present study adds considerably to this body of knowledge by showing that people can learn second-, third-, and fourth-order nonadjacent dependencies, as well as third- and fourth-order adjacent dependencies. This suggests that the sequence-learning mechanism can form associations between nonadjacent events and use information four trials back to differentially predict events on the current trial. The mechanism would appear to be quite powerful.

If learning in the present study was explicit (i.e., the result of conscious, intentional processes) then participants would presumably have some awareness of the dependencies embedded in the sequences of target locations. There was no evidence from the awareness measures that participants had such awareness, suggesting that learning was implicit. A potential criticism is that the awareness measures, particularly the paper-and-pencil surveys in Experiments 1 and 2, were not very sensitive because they did not reinstate many of the cues that were present during the SRT task. The existing evidence, however, suggests that the awareness measures used in the present study are sensitive. First, a survey like those in Experiments 1 and 2 can detect awareness of lag 1 probabilities of .60 versus .40 (83% correct, $r_{\text{contrast}} = .82$) and lag 2−1 probabilities of .60 versus .40 (71% correct, $r_{\text{contrast}} = .62$) after extended training on the SRT task (Remillard & Clark, 2001, Exp. 4). Second, Willingham, Greeley, and Bardone (1993) showed that a recognition test where sequences of target locations are presented as sequences of digits is as sensitive as a recognition test that requires participants to respond to the sequences as in the SRT task. Third, a recognition test with sequences presented as digits is capable of detecting conscious sequence knowledge after very limited exposure to the sequence during the SRT task (Perruchet, Bigand, & Benoit-Gonin, 1997, Exp. 3). This suggests that such a recognition test is sensitive to relatively low levels of awareness. Finally, a recognition test with sequences presented as digits can discriminate participants who are instructed to try to explicitly learn a sequence while performing the SRT task from participants who are not given such instructions (Curran, 1997).

The present study indicates that people can learn lag $2-x$ probabilities of .60 versus .40 in a sequence of events when there are no experimenter-induced pauses in the sequence (i.e., the sequence is continuous), and elements in the sequence are from the same domain (i.e., spatial locations). This is in contrast to the results from word segmentation studies, which suggest that people can learn lag $2-x$ probabilities of 1.0 between elements in a word only when there are short pauses between words in the sequence (Gomez, 2002; Perruchet et al., 2004), or odd and even elements in a word are from different domains (Bonatti et al., 2005; Creel et al., 2004; Newport & Aslin, 2004; Onnis et al., 2005).

There are a number of differences between the present study and word segmentation studies. Consequently, the reasons for the discrepancies in lag $2-x$ learning between the two types of study are not obvious. One possibility is that learning in the present study was primarily response based (participants responded to the location of the target with a corresponding keypress), whereas learning in word segmentation studies can only be perceptual (participants simply listen to a novel language). Remillard (2003, Exp. 4) failed to obtain evidence for perceptual-based learning of lag 2−1 probabilities of .67 versus .33 in a continuous sequence of spatial locations (also see Deroost & Soetens, 2006). In contrast, people can learn nonadjacent dependencies in sequences of musical tones that have timing gaps at strategic locations (Dienes & Longuet-Higgins, 2004; Kuhn & Dienes, 2005). Thus perceptual-based learning of second- or higher-order probability information might be very difficult in the absence of salient cues such as temporal pauses.

The larger RT differences between AN−H and AN−L transitions than between A−H and A−L transitions in the present study were taken as evidence that participants learned nonadjacent dependencies. For example, in Table 2, a larger RT difference between 2−1−3 and 2−1−4 (rows 1 and 2 in tier AN2) than between 3−2−1 and 3−2−6 (rows 9 and 10 in tier A2) suggests learning of the nonadjacent dependencies $P(3|2-x) = .60 > P(4|2-x) = .40$. However, there is an alternative explanation that does not involve learning of nonadjacent dependencies. An encounter with 2−1 (rows 1 and 2) may activate the chunk 2−6 (rows 5 and 6) because both have element 2 in the same position. The chunks 2−1 and 2−6 both prime 3 (an H transition), producing a relatively large RT difference between 2−1−3 and 2−1−4. In contrast, an encounter with 3−2 (rows 9 and 10) may activate the chunk 3−5 (rows 13 and 14) with 3−2 priming 1 (an H transition) and 3−5 priming 6 (an H transition). The result is a smaller RT difference between 3−2−1 and 3−2−6 than between 2−1−3 and 2−1−4. A similar analysis can be applied to learning of third- and fourth-order nonadjacent dependencies. Thus learning may involve only adjacent dependencies, and the L − H advantage for AN transitions relative to A transitions may reflect activation of similar chunks that prime the same successor in the case of AN transitions and different successors in the case of A transitions.

The chunk overlap hypothesis predicts that L − H differences in RT and error rate for AN transitions should asymptote at the same time as that for A transitions. Once learning of adjacent dependencies has reached asymptote, as indicated by differences between A−H and A−L transitions, there should be no reason for differences between AN−H and AN−L transitions to keep growing. The results from the present study are generally inconsistent with this prediction. For example, in Experiment 3, RT and error rate differences between A4−H and A4−L transitions did not change significantly across the 10 sessions ($F < 1$ and $p = .283$, respectively). In contrast, the error rate difference between AN4−H and AN4−L transitions increased across sessions ($p < .001$), and the RT difference also tended to increase across sessions (see Figure 3) although not significantly ($p = .187$). Thus L − H differences in RT and error rate appeared to asymptote sooner for A4 than AN4 transitions, contrary to the chunk overlap hypothesis. Similarly, in Experiment 2, the error rate difference between A3−H and A3−L transitions did not change significantly across the 10 sessions ($p = .214$), whereas the difference between AN3−H and AN3−L transitions increased across sessions ($p = .046$). Finally, in Experiment 1, the RT and error rate differences between A2−H and A2−L transitions did not change significantly across Sessions 2 to 4 ($p = .250$ and $F < 1$, respectively), whereas the differences between AN2−H and AN2−L transitions increased across Sessions 2 to 4 ($p = .004$ and $p = .019$, respectively). Thus L − H differences in RT and error rate seemed to asymptote sooner for A than AN transitions. These results are inconsistent with the chunk overlap hypothesis, but consistent with the idea that participants learn adjacent and nonadjacent dependencies and that adjacent dependencies are learned more quickly than nonadjacent dependencies. The preceding analysis is clearly post hoc, and further studies will be needed to distinguish the nonadjacency hypothesis from the chunk overlap hypothesis.

The results of the present study are summarized in Figure 4 and are generally consistent with the results of similar studies. For example, Remillard and Clark (2001, Exp. 2) examined learning of second-order adjacent dependencies of .60 versus .40 using a four-choice SRT task and obtained an L − H difference of 14 ms. This is similar to the 12-ms difference for A2 transitions in Experiment 1. J. H. Howard and Howard (1997; also see D. V. Howard et al., 2004) examined

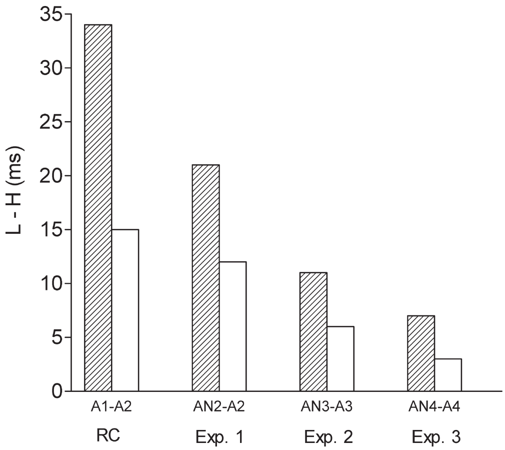

Figure 4. *The difference in reaction time between H and L transitions as a function of type of transition and experiment. Reaction times were averaged across Sessions 3 and 4 for Experiment 1 and Sessions 6–10 for Experiments 2 and 3, and for RC (i.e., Remillard & Clark, 2001, Exp. 4). The results from Remillard and Clark are included for the sake of completeness. They examined learning of first- and second-order adjacent dependencies of .60 versus .40 (A1 and A2 transitions, respectively) using a six-choice SRT task like the one in the present study.*

learning of second-order information using a four-choice SRT task where contexts of length 2 were followed by one successor with probability .625 and three other successors each with probability .125. Adjacent and nonadjacent dependencies were confounded and the L − H difference was 32 ms. However, when participants were required to maintain high accuracy, as in the present study, the L − H difference dropped to 24 ms. This is similar to the 21-ms difference for AN2 transitions in Experiment 1. Finally, D. V. Howard et al. (2004) examined learning of third-order information using a four-choice SRT task where contexts of length 3 were followed by one successor with probability .500 and three other successors each with probability .167. Adjacent and nonadjacent dependencies were confounded, and the L − H difference was a significant 3 ms. This is smaller than the 11-ms difference for AN3 transitions in Experiment 2, but understandable given that high transition probabilities were .50 rather than .60 as in Experiment 2.

The present study has provided evidence that people can learn second-, third-, and fourth-order nonadjacent dependencies as well as third- and fourth-order adjacent dependencies, and that the learning can be implicit. Models of implicit sequence learning will have to account for these new findings.

Original manuscript received 6 August 2006
Accepted revision received 10 January 2007
First published online 16 May 2007

REFERENCES

Bonatti, L. L., Pena, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations. *Psychological Science*, *16*, 451–459.

Bovens, N., & Brysbaert, M. (1990). IBM PC/XT/AT and PS/2 Turbo Pascal timing with extended resolution. *Behavior Research Methods, Instruments, & Computers*, *22*, 332–334.

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

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

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

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *5*, 1119–1130.

Curran, T. (1997). Effects of aging on implicit sequence learning: Accounting for sequence structure and explicit knowledge. *Psychological Research*, *60*, 24–41.

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

Curran, T., Smith, M. D., DiFranco, J. M., & Daggy, A. T. (2001). Structural influences on implicit and explicit sequence learning. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 40, pp. 147–182). San Diego, CA: Academic Press.

Deroost, N., & Soetens, E. (2006). Perceptual or motor learning in SRT tasks with complex sequence structures. *Psychological Research*, *70*, 88–102.

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

Frensch, P. A. (1998). One concept multiple meanings: On how to define the concept of implicit learning. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 47–104). Thousand Oaks, CA: Sage Publications.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

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

Howard, D. V., Howard, J. H., Jr., Japikse, K., Di Yanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: Effects of level of structure, adult age, and extended practice. *Psychology and Aging*, *19*, 79–92.

Howard, J. H., Jr., & Howard, D. V. (1997). Age differences in implicit learning in higher order dependencies in serial patterns. *Psychology and Aging*, *12*, 634–656.

Jimenez, L., & Mendez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 236–259.

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

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review*, *110*, 316–339.

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

Kuhn, G., & Dienes, Z. (2005). Implicit learning of nonlocal musical rules: Implicitly learning more than chunks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1417–1432.

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

Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 350–364.

McDowall, J., Lustig, A., & Parkin, G. (1995). Indirect learning of event sequences: The effects of divided attention and stimulus continuity. *Canadian Journal of Experimental Psychology*, *49*, 415–435.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, *53*, 225–237.

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

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233–238.

Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology: General*, *133*, 573–583.

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*, 246–263.

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.

Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 581–597.

Remillard, G., & Clark, J. M. (1999). Generating fixed-length sequences satisfying any given nth-order transition probability matrix. *Behavior Research Methods, Instruments, and Computers*, *31*, 235–243.

Remillard, G., & Clark, J. M. (2001). Implicit learning of first-, second-, and third-order transition probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 483–498.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge, UK: Cambridge University Press.

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

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.

Shin, J. C., & Ivry, R. B. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 445–457.

Stadler, M. A. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1061–1069.

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

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

St. John, M. F., & Shanks, D. R. (1997). Implicit learning from an information processing standpoint. In D. C. Berry (Ed.), *How implicit is implicit learning* (pp. 125–161). New York: University Press.

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

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

APPENDIX A

Frequencies for the 12 versions of Table 2

		Version											
Context	*S*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
3–2–1	3	18	8	15	10	15	10	15	10	15	10	18	8
	4	12	12	10	15	15	10	10	15	15	10	12	12
4–2–1	3	12	12	15	10	10	15	15	10	10	15	12	12
	4	8	18	10	15	10	15	10	15	10	15	8	18
3–5–1	3	8	18	10	15	15	10	10	15	10	15	12	12
	4	12	12	15	10	15	10	15	10	10	15	18	8
4–5–1	3	12	12	10	15	10	15	10	15	15	10	8	18
	4	18	8	15	10	10	15	15	10	15	10	12	12
3–2–6	3	12	12	10	15	10	15	15	10	10	15	8	18
	4	8	18	15	10	10	15	10	15	10	15	12	12
4–2–6	3	18	8	10	15	15	10	15	10	15	10	12	12
	4	12	12	15	10	15	10	10	15	15	10	18	8
3–5–6	3	12	12	15	10	10	15	10	15	15	10	12	12
	4	18	8	10	15	10	15	15	10	15	10	8	18
4–5–6	3	8	18	15	10	15	10	10	15	10	15	18	8
	4	12	12	10	15	15	10	15	10	10	15	12	12
1–3–2	1	15	10	15	10	18	8	15	10	18	8	15	10
	6	10	15	15	10	12	12	15	10	12	12	10	15
6–3–2	1	15	10	10	15	12	12	10	15	12	12	15	10
	6	10	15	10	15	8	18	10	15	8	18	10	15
1–4–2	1	10	15	15	10	8	18	10	15	12	12	10	15
	6	15	10	15	10	12	12	10	15	18	8	15	10

(*Continued overleaf*)

Appendix A. (Continued)

		Version											
Context	*S*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
6–4–2	1	10	15	10	15	12	12	15	10	8	18	10	15
	6	15	10	10	15	18	8	15	10	12	12	15	10
1–3–5	1	10	15	10	15	12	12	10	15	8	18	15	10
	6	15	10	10	15	8	18	10	15	12	12	10	15
6–3–5	1	10	15	15	10	18	8	15	10	12	12	15	10
	6	15	10	15	10	12	12	15	10	18	8	10	15
1–4–5	1	15	10	10	15	12	12	15	10	12	12	10	15
	6	10	15	10	15	18	8	15	10	8	18	15	10
6–4–5	1	15	10	15	10	8	18	10	15	18	8	10	15
	6	10	15	15	10	12	12	10	15	12	12	15	10
2–1–3	2	15	10	18	8	15	10	18	8	15	10	15	10
	5	15	10	12	12	10	15	12	12	10	15	15	10
5–1–3	2	10	15	12	12	15	10	12	12	15	10	10	15
	5	10	15	8	18	10	15	8	18	10	15	10	15
2–6–3	2	15	10	8	18	10	15	12	12	10	15	10	15
	5	15	10	12	12	15	10	18	8	15	10	10	15
5–6–3	2	10	15	12	12	10	15	8	18	10	15	15	10
	5	10	15	18	8	15	10	12	12	15	10	15	10
2–1–4	2	10	15	12	12	10	15	8	18	15	10	10	15
	5	10	15	8	18	15	10	12	12	10	15	10	15
5–1–4	2	15	10	18	8	10	15	12	12	15	10	15	10
	5	15	10	12	12	15	10	18	8	10	15	15	10
2–6–4	2	10	15	12	12	15	10	12	12	10	15	15	10
	5	10	15	18	8	10	15	8	18	15	10	15	10
5–6–4	2	15	10	8	18	15	10	18	8	10	15	10	15
	5	15	10	12	12	10	15	12	12	15	10	10	15

Note: The frequencies for Version 1 are those listed in Table 2. S = successor.

APPENDIX B

Probabilities for the 12 versions of Table 5

	Context	*Successor*																							
		3												4											
1.	1–3–2–1	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h	H	L	H	L	l	h
2.	6–3–2–1	H	L	H	L	l	h	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l
3.	1–4–2–1	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L
4.	6–4–2–1	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H
5.	1–3–5–1	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L
6.	6–3–5–1	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H
7.	1–4–5–1	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h	H	L	H	L	l	h
8.	6–4–5–1	H	L	H	L	l	h	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l
9.	1–3–2–6	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L
10.	6–3–2–6	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H
11.	1–4–2–6	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h	H	L	H	L	l	h
12.	6–4–2–6	H	L	H	L	l	h	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l
13.	1–3–5–6	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h	H	L	H	L	l	h
14.	6–3–5–6	H	L	H	L	l	h	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l
15.	1–4–5–6	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L
16.	6–4–5–6	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H
		1												6											
17.	2–1–3–2	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L
18.	5–1–3–2	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H
19.	2–6–3–2	l	h	H	L	H	L	h	l	L	H	L	H	h	l	L	H	L	H	l	h	H	L	H	L
20.	5–6–3–2	h	l	L	H	L	H	l	h	H	L	H	L	l	h	H	L	H	L	h	l	L	H	L	H
21.	2–1–4–2	l	h	H	L	H	L	h	l	L	H	L	H	h	l	L	H	L	H	l	h	H	L	H	L
22.	5–1–4–2	h	l	L	H	L	H	l	h	H	L	H	L	l	h	H	L	H	L	h	l	L	H	L	H
23.	2–6–4–2	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L
24.	5–6–4–2	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H
25.	2–1–3–5	l	h	H	L	H	L	h	l	L	H	L	H	h	l	L	H	L	H	l	h	H	L	H	L
26.	5–1–3–5	h	l	L	H	L	H	l	h	H	L	H	L	l	h	H	L	H	L	h	l	L	H	L	H
27.	2–6–3–5	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L
28.	5–6–3–5	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H
29.	2–1–4–5	L	H	h	l	L	H	L	H	h	l	L	H	H	L	l	h	H	L	H	L	l	h	H	L
30.	5–1–4–5	H	L	l	h	H	L	H	L	l	h	H	L	L	H	h	l	L	H	L	H	h	l	L	H
31.	2–6–4–5	l	h	H	L	H	L	h	l	L	H	L	H	h	l	L	H	L	H	l	h	H	L	H	L
32.	5–6–4–5	h	l	L	H	L	H	l	h	H	L	H	L	l	h	H	L	H	L	h	l	L	H	L	H
		2												5											
33.	3–2–1–3	h	l	L	H	L	H	h	l	L	H	L	H	l	h	H	L	H	L	l	h	H	L	H	L
34.	4–2–1–3	l	h	H	L	H	L	l	h	H	L	H	L	h	l	L	H	L	H	h	l	L	H	L	H
35.	3–5–1–3	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h
36.	4–5–1–3	L	H	L	H	h	l	H	L	H	L	l	h	H	L	H	L	l	h	L	H	L	H	h	l
37.	3–2–6–3	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h
38.	4–2–6–3	L	H	L	H	h	l	H	L	H	L	l	h	H	L	H	L	l	h	L	H	L	H	h	l
39.	3–5–6–3	h	l	L	H	L	H	h	l	L	H	L	H	l	h	H	L	H	L	l	h	H	L	H	L
40.	4–5–6–3	l	h	H	L	H	L	l	h	H	L	H	L	h	l	L	H	L	H	h	l	L	H	L	H
41.	3–2–1–4	H	L	H	L	l	h	L	H	L	H	h	l	L	H	L	H	h	l	H	L	H	L	l	h

(*Continued overleaf*)

Appendix B. (Continued)

	Context	*Successor*	
		2	5
42.	4–2–1–4	L H L H h l H L H L l h	H L H L l h L H L H h l
43.	3–5–1–4	h l L H L H h l L H L H	l h H L H L l h H L H L
44.	4–5–1–4	l h H L H L l h H L H L	h l L H L H h l L H L H
45.	3–2–6–4	h l L H L H h l L H L H	l h H L H L l h H L H L
46.	4–2–6–4	l h H L H L l h H L H L	h l L H L H h l L H L H
47.	3–5–6–4	H L H L l h L H L H h l	L H L H h l H L H L l h
48.	4–5–6–4	L H L H h l H L H L l h	H L H L l h L H L H h l

Note: The first through 12th letters in each 12-letter string represent the probabilities for Versions 1–12, respectively. For example, the first letters in the 12-letter strings are the probabilities for Version 1 (i.e., Table 5). For Versions 1–6, the two tiers containing both uppercase and lowercase letters describe AN4 transitions, and the tier containing only uppercase letters describes A4 transitions. For Versions 7–12, the two tiers containing both uppercase and lowercase letters describe A4 transitions, and the tier containing only uppercase letters describes AN4 transitions.