Implicit learning of fifth- and sixth-order sequential probabilities

Gilbert Remillard
Morehead State University, Morehead, Kentucky

Serial reaction time (SRT) task studies have established that people can implicitly learn sequential contingencies as complex as fourth-order probabilities. The present study examined people's ability to learn fifth-order (Experiment 1) and sixth-order (Experiment 2) probabilities. Remarkably, people learned fifth- and sixth-order probabilities. This suggests that the implicit sequence learning mechanism can operate over a range of at least seven sequence elements.

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 or newly introduced 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, intentional 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., Curran & Keele, 1993; Lewicki, Hill, & Bizot, 1988; McDowall, Lustig, & Parkin, 1995; Reed & Johnson, 1994; Remillard, 2008a; Remillard & Clark, 2001; Stadler, 1989, 1993, 1995).

Implicit sequence learning involves the extraction of sequential contingencies. Surprisingly, there has been very little systematic investigation into the kinds of sequential contingencies that people can learn implicitly. The overwhelming majority of studies have employed sequences in which the next element in the sequence is predictable from elements one or two trials back, with only a handful of studies having used sequences in which the next element is predictable from elements three trials back. One reason for

this state of affairs may be the results of a highly influential study by Cleeremans and McClelland (1991) suggesting that people could not learn to use sequence elements four trials back to anticipate the next element in the sequence.

There are two reasons to believe that people are capable of implicitly learning much more complex sequential contingencies. First, implicit sequence learning, unlike explicit sequence learning, is not bound by the capacity limits of working memory. Implicit sequence learning proceeds independently of individual differences in working memory capacity, whereas explicit sequence learning does not (Frensch & Miner, 1994, Experiment 1; Song, Marks, Howard, & Howard, 2009; Unsworth & Engle, 2005). Also, implicit sequence learning is unaffected by the amount of working memory resources devoted to learning the sequence. This is suggested by studies showing that explicit learning of a sequence is accompanied by parallel implicit learning of the sequence, and the extent of implicit learning is equivalent to that when there is no explicit learning (Song et al., 2009; Song, Howard, & Howard, 2007; Willingham & Goedert-Eschmann, 1999; Willingham, Salidis, & Gabrieli, 2002). Finally, sequence knowledge acquired through explicit learning can be used in a controlled and flexible manner (suggesting the involvement of working memory), whereas sequence knowledge acquired through implicit learning cannot (Jiménez, Vaquero, & Lupiáñez, 2006).

Second, it has been proposed that implicit sequence learning involves the operation of associative processes on sequence elements stored in a short-term memory system (Frensch, 1998, pp. 87–95; Frensch, Buchner, & Lin, 1994; Frensch & Miner, 1994). Information in this memory system appears to decay quite slowly because sequence learning still occurs when the time interval between the response to the current sequence element and the appearance of the next sequence element (i.e., the response–stimulus interval) is as long as 2 sec (see, e.g.,

---

G. Remillard, g.remillard@moreheadstate.edu

Willingham, Greenberg, & Thomas, 1997). A slow decay rate suggests that the capacity of the memory system is rather large, making it possible to learn complex sequential contingencies.

Consistent with the idea that people are capable of implicitly learning complex sequential contingencies, Remillard (2008a), who identified a confound in the sequences used by Cleeremans and McClelland (1991) that could have masked learning, has shown that people can implicitly learn second-, third-, and fourth-order adjacent and nonadjacent probabilities. The probabilities are described in Table 1. As a class, nth-order probabilities are conditional probabilities that involve events as far back as trial $t-n$, where trial t is the current trial. The nth-order probabilities can be subdivided into adjacent and nonadjacent probabilities. Adjacent probabilities involve events from consecutive trials (e.g., lag 2-1 and lag 3-2-1 probabilities). Nonadjacent probabilities involve events that skip over at least one trial (e.g., lag 2-x and lag 4-x-2-x probabilities).

The present study examined whether people can implicitly learn sequential contingencies more complex than those studied by Remillard (2008a). No SRT task studies have examined learning of fifth- or higher order probabilities. Using a different sequence learning paradigm, Millward and Reber (1972) showed that people could learn to use sequence elements as far back as seven trials to explicitly predict the next element in the sequence. Unfortunately, explicit prediction of sequence elements may have encouraged explicit sequence learning. Also, participants were required to learn simpler contingencies before learning more complex contingencies. Learning the simpler contingencies could have bootstrapped learning the more complex contingencies (e.g., see Lany & Gomez, 2008; Lany, Gomez, & Gerken, 2007). Participants in the present study, unlike those in the Millward and Reber study, did not explicitly predict sequence elements and were not exposed to simpler contingencies prior to being exposed to more complex contingencies.

EXPERIMENT 1

Experiment 1 used a six-choice SRT task. The sequences of target locations were structured so that fifth-order probabilities varied (.25 vs. .75) and first- through fourth-order probabilities were constant (.50). Experiment 1 did not attempt to distinguish between learning of the 16 different types (1 adjacent and 15 nonadjacent) of fifth-order probability because these were confounded (e.g., P[3|2-1-3-5-1] = P[3|2-x-x-x-x] = P[3|2-x-3-5-x] = .25 and P[4|2-1-3-5-1] = P[4|2-x-x-x-x] = P[4|2-x-3-5-x] = .75). Thus, shorter RTs on high-probability successors (e.g., on 4 given 2-1-3-5-1, where P[4|2-1-3-5-1] = .75) than on low-probability successors (e.g., on 3 given 2-1-3-5-1, where P[3|2-1-3-5-1] = .25) would indicate learning of at least 1 of the 16 types of fifth-order probability. More specifically, it would indicate that people can learn to use the location of the target on trial $t-5$ to predict its location on trial t.

Method

Participants. Six introductory psychology students, 18 or 19 years of age, participated for course credit and payment of $100.

SRT task. The SRT task was run on a desktop computer with a standard monitor and keyboard. The six target locations were horizontally arranged and each location was marked with a line measuring 0.4×0.1 cm. Adjacent lines were separated by an interval of 2.0 cm. The target was a line measuring 0.1×0.3 cm that appeared 0.2 cm above a line marking a target location. All lines were black and the background was gray (RGB values 210, 210, 210). The viewing distance was approximately 60 cm. The red-stickered X, C, V, M, $<$, and $>$ 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 appeared at one of the six marked locations, 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 300 msec later. RT was measured as the time between target appearance and the first response.

There were 16 sessions. Each session was composed of 18 blocks of trials with 124 trials per block. Session 1 began with a practice block of 100 trials. On a given day, there were 0 or 1 sessions. There were never more than 3 consecutive zero-session days. The 16 sessions were completed in 18 to 24 days.

A performance history was provided at the end of each block of trials in a session. 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-sec 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. Locations 1 and 6 could each be succeeded by Locations 3 and 4, Locations 3 and 4 could each

Table 1
Types of Probabilities

Probability Class	Probability Subclass	Specific Probability	Symbolically
First order	Adjacent	Lag 1	$P(E \mid A_1)$
Second order	Adjacent	Lag 2-1	$P(E \mid A_2\text{-}A_1)$
	Nonadjacent	Lag 2-x	$P(E \mid A_2\text{-}x)$
Third order	Adjacent	Lag 3-2-1	$P(E \mid A_3\text{-}A_2\text{-}A_1)$
	Nonadjacent	Lag 3-2-x	$P(E \mid A_3\text{-}A_2\text{-}x)$
		Lag 3-x-1	$P(E \mid A_3\text{-}x\text{-}A_1)$
		Lag 3-x-x	$P(E \mid A_3\text{-}x\text{-}x)$
Fourth order	Adjacent	Lag 4-3-2-1	$P(E \mid A_4\text{-}A_3\text{-}A_2\text{-}A_1)$
	Nonadjacent	Lag 4-3-2-x	$P(E \mid A_4\text{-}A_3\text{-}A_2\text{-}x)$
		Lag 4-3-x-1	$P(E \mid A_4\text{-}A_3\text{-}x\text{-}A_1)$
		Lag 4-x-2-1	$P(E \mid A_4\text{-}x\text{-}A_2\text{-}A_1)$
		Lag 4-3-x-x	$P(E \mid A_4\text{-}A_3\text{-}x\text{-}x)$
		Lag 4-x-2-x	$P(E \mid A_4\text{-}x\text{-}A_2\text{-}x)$
		Lag 4-x-x-1	$P(E \mid A_4\text{-}x\text{-}x\text{-}A_1)$
		Lag 4-x-x-x	$P(E \mid A_4\text{-}x\text{-}x\text{-}x)$

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 $P(E \mid A_3\text{-}A_2\text{-}A_1)$ 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 $P(E \mid A_4\text{-}x\text{-}A_2\text{-}x)$ 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."

Table 2
Probabilities Inherent in the Sequences of
Target Locations in Experiment 1

Tier	Row	Context	Successor 1	2	3	4	5	6
1	1	2-1-3-2-1	–	–	L	H	–	–
	2	5-1-3-2-1	–	–	H	L	–	–
	3	2-6-3-2-1	–	–	L	H	–	–
	4	5-6-3-2-1	–	–	H	L	–	–
	5	2-1-4-2-1	–	–	L	H	–	–
	6	5-1-4-2-1	–	–	H	L	–	–
	7	2-6-4-2-1	–	–	L	H	–	–
	8	5-6-4-2-1	–	–	H	L	–	–
	.							
	25	2-1-3-5-6	–	–	L	H	–	–
	26	5-1-3-5-6	–	–	H	L	–	–
	27	2-6-3-5-6	–	–	L	H	–	–
	28	5-6-3-5-6	–	–	H	L	–	–
	29	2-1-4-5-6	–	–	L	H	–	–
	30	5-1-4-5-6	–	–	H	L	–	–
	31	2-6-4-5-6	–	–	L	H	–	–
	32	5-6-4-5-6	–	–	H	L	–	–
2	33	3-2-1-3-2	M	–	–	–	–	M
	34	4-2-1-3-2	M	–	–	–	–	M
	35	3-5-1-3-2	M	–	–	–	–	M
	36	4-5-1-3-2	M	–	–	–	–	M
	37	3-2-6-3-2	M	–	–	–	–	M
	38	4-2-6-3-2	M	–	–	–	–	M
	39	3-5-6-3-2	M	–	–	–	–	M
	40	4-5-6-3-2	M	–	–	–	–	M
	.							
	57	3-2-1-4-5	M	–	–	–	–	M
	58	4-2-1-4-5	M	–	–	–	–	M
	59	3-5-1-4-5	M	–	–	–	–	M
	60	4-5-1-4-5	M	–	–	–	–	M
	61	3-2-6-4-5	M	–	–	–	–	M
	62	4-2-6-4-5	M	–	–	–	–	M
	63	3-5-6-4-5	M	–	–	–	–	M
	64	4-5-6-4-5	M	–	–	–	–	M
3	65	1-3-2-1-3	–	H	–	–	L	–
	66	6-3-2-1-3	–	L	–	–	H	–
	67	1-4-2-1-3	–	H	–	–	L	–
	68	6-4-2-1-3	–	L	–	–	H	–
	69	1-3-5-1-3	–	H	–	–	L	–
	70	6-3-5-1-3	–	L	–	–	H	–
	71	1-4-5-1-3	–	H	–	–	L	–
	72	6-4-5-1-3	–	L	–	–	H	–
	.							
	89	1-3-2-6-4	–	H	–	–	L	–
	90	6-3-2-6-4	–	L	–	–	H	–
	91	1-4-2-6-4	–	H	–	–	L	–
	92	6-4-2-6-4	–	L	–	–	H	–
	93	1-3-5-6-4	–	H	–	–	L	–
	94	6-3-5-6-4	–	L	–	–	H	–
	95	1-4-5-6-4	–	H	–	–	L	–
	96	6-4-5-6-4	–	L	–	–	H	–

Note—Across every 16 occurrences of a context with L/H successors, the L and H successors occurred 4 times and 12 times, respectively. Across every 16 occurrences of a context with M successors, each M successor occurred 8 times. A dash indicates that the transition did not occur. L, low-probability successor; H, high-probability successor; M, medium-probability successor.

be succeeded by Locations 2 and 5, and Locations 2 and 5 could each be succeeded by Locations 1 and 6. Consequently, there were 96 (6×2^4) possible contexts of length 5, 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 some of the 96 contexts and the probabilities with which successors followed contexts. The 96 contexts were divided into three tiers, with each tier composed of 32 contexts.

The contexts in Tier 1 (Rows 1–32) began with Locations 2 or 5 and were succeeded by Locations 3 or 4. Every 16 occurrences of a context were followed 4 times by one successor (low-probability successor, L) and 12 times by the other successor (high-probability successor, H). For example, Row 3 indicates that every 16 occurrences of context 2-6-3-2-1 were followed 4 times by Successor 3 and 12 times by Successor 4, so that $P(3|2\text{-}6\text{-}3\text{-}2\text{-}1) = .25$ and $P(4|2\text{-}6\text{-}3\text{-}2\text{-}1) = .75$. Fifth-order probabilities in Tier 1 were .25 or .75. The odd-numbered rows reveal that whenever Location 2 was the target location on trial $t-5$, Locations 3 and 4 were the target locations on trial t with probabilities .25 and .75, respectively. Thus, all 16 types of fifth-order probability with Location 2 on trial $t-5$ and Location 3 (Location 4) on trial t had a value of .25 (.75). For example, $P(3|2\text{-}6\text{-}3\text{-}2\text{-}1) = P(3|2\text{-}x\text{-}x\text{-}x\text{-}x) = P(3|2\text{-}6\text{-}x\text{-}x\text{-}x) = P(3|2\text{-}6\text{-}x\text{-}2\text{-}1) = .25$ and $P(4|2\text{-}6\text{-}3\text{-}2\text{-}1) = P(4|2\text{-}x\text{-}x\text{-}x\text{-}x) = P(4|2\text{-}6\text{-}x\text{-}x\text{-}x) = P(4|2\text{-}6\text{-}x\text{-}2\text{-}1) = .75$. The even-numbered rows reveal that whenever Location 5 was the target location on trial $t-5$, Locations 3 and 4 were the target locations on trial t with probabilities .75 and .25, respectively. Thus, all 16 types of fifth-order probability with Location 5 on trial $t-5$ and Location 3 (Location 4) on trial t had a value of .75 (.25). For example, $P(3|5\text{-}6\text{-}3\text{-}2\text{-}1) = P(3|5\text{-}x\text{-}x\text{-}x\text{-}x) = P(3|5\text{-}6\text{-}x\text{-}x\text{-}x) = P(3|5\text{-}6\text{-}x\text{-}2\text{-}1) = .75$ and $P(4|5\text{-}6\text{-}3\text{-}2\text{-}1) = P(4|5\text{-}x\text{-}x\text{-}x\text{-}x) = P(4|5\text{-}6\text{-}x\text{-}x\text{-}x) = P(4|5\text{-}6\text{-}x\text{-}2\text{-}1) = .25$. Importantly, all first- through fourth-order probabilities in Tier 1 were .50 (e.g., $P[3|1] = P[4|1] = .50$ and $P[3|6\text{-}x\text{-}x\text{-}x] = P[4|6\text{-}x\text{-}x\text{-}x] = .50$), and Locations 3 and 4 were each a target location equally often (i.e., $P[3] = P[4] = .17$). Therefore, shorter RTs on H than on L successors would indicate that participants learned to use the location of the target on trial $t-5$, alone or in conjunction with subsequent target locations, to predict the location of the target on trial t (i.e., participants learned at least 1 of the 16 types of fifth-order probability). Tier 3 (Rows 65–96) was similar to Tier 1 except that contexts began with Locations 1 or 6 and were succeeded by Locations 2 or 5.

The contexts in Tier 2 (Rows 33–64) began with Locations 3 or 4 and were succeeded by Locations 1 or 6. Every 16 occurrences of a context were followed 8 times by one successor (medium-probability successor, M) and 8 times by the other successor (medium-probability successor, M). For example, Row 35 indicates that every 16 occurrences of context 3-5-1-3-2 were followed 8 times by Successor 1 and 8 times by Successor 6 so that $P(1|3\text{-}5\text{-}1\text{-}3\text{-}2) = .50$ and $P(6|3\text{-}5\text{-}1\text{-}3\text{-}2) = .50$. All first- through fifth-order probabilities in Tier 2 were .50.

A number of factors other than transition probabilities can influence RTs on H and L successors. One factor is the pattern of within-hand (W) and between-hand (B) transitions that make up a context and successor. For example, the context 2-6-4-2-1 and its Successor 3 form a BWBWW pattern because 2-6, 6-4, 4-2, 2-1, and 1-3 are B, W, B, W, and W transitions, respectively. RT to the last element of a pattern can vary across the different patterns (Remillard, 2008a). The contexts and L (H) successors in Rows 1–32 of Table 2 were matched to the contexts and H (L) successors in Rows 65–96, respectively, with respect to the pattern of W/B transitions and the response hand for the successor. For example, the context and L successor in Row 4 and the context and H successor in Row 68 both form a WBWWB pattern, and both successors require a right-hand response.

Sequential priming effects are another factor that can influence RTs on H and L successors. Table 3 lists eight types of six-element runs on the basis of the first, second, and third elements being equal (E) or unequal (U) to the fourth, fifth, and sixth elements, respectively. RT to the last element of a run can vary across the different types of runs (Remillard & Clark, 2001). The contexts and L (H) successors in Rows 1–32 of Table 2 were matched to the contexts and H (L) successors in Rows 65–96, respectively, with respect to type of run. For example, the context and L successor in Row 4 and the context and H successor in Row 68 are both UUU runs.

Table 3
Run Types

Run	Example
EEE	1-3-2-1-3-2
UEE	6-3-2-1-3-2
EUE	1-4-2-1-3-2
UUE	6-4-2-1-3-2
EEU	1-3-5-1-3-2
UEU	6-3-5-1-3-2
EUU	1-4-5-1-3-2
UUU	6-4-5-1-3-2

Note—Six-element runs were categorized as a function of the first, second, and third elements being equal (E) or unequal (U) to the fourth, fifth, and sixth elements, respectively.

Software developed by Remillard (2008b) was used to generate the sequences of target locations. For each participant, a 34,277-element sequence was randomly generated, with the constraint that across every 16 occurrences of a context with L/H (M) successors, the L successor occurred 4 times and the H successor occurred 12 times (each M successor occurred 8 times). Elements 1–124, 120–243, 239–362, and so forth to 34,154–34,277 each constituted a block of 124 trials, for a total of 288 blocks (16 sessions × 18 blocks per session). A computer program went over each 34,277-element sequence and determined the exact values of all first- through fifth-order probabilities to ensure that the values were as expected. The practice block of 100 trials at the beginning of Session 1 was randomly generated, with the constraint that across every 2 occurrences of a context, each of its successors occurred once.

There were six versions of Table 2. Version 1 was Table 2. Version 2 was formed from Version 1 by exchanging L and H successors. Version 3 was created by having Tiers 1, 2, and 3 of Table 2 describe M, L/H, and L/H successors, respectively. Version 4 was formed from Version 3 by exchanging L and H successors. Version 5 was created by having Tiers 1, 2, and 3 of Table 2 describe L/H, L/H, and M successors, respectively. Version 6 was formed from Version 5 by exchanging L and H successors.

Prediction task. Awareness of the fifth-order probabilities was assessed using a prediction task. There were 64 prediction trials corresponding to the 64 contexts with L/H successors (e.g., Tiers 1 and 3 in Table 2). A prediction trial began with a press of the space bar in response to a prompt on the screen. This was followed by the disappearance of the prompt and the appearance of the six horizontal lines marking the six target locations. After another 1,800 msec, participants observed the target move across five locations (i.e., a context) followed by the appearance of a question mark below each of the context's two possible successors. Participants indicated at which of the two marked locations the target was most likely to have appeared next during training, given the preceding five target locations, by pressing the corresponding response key. The sequence of target movements could be repeated any number of times by pressing the K key if participants felt that they needed to see the sequence again prior to making a prediction response. In a sequence of target movements, target duration was 275 msec and the interstimulus interval was 300 msec. Following a prediction response, the screen was cleared and the prompt to press the space bar to begin the next trial appeared.

The 64 prediction trials were presented in a random order for each participant. Participants performed two practice prediction trials prior to starting the 64 prediction trials. Scores greater than 50% correct (random guessing performance) on the 64 trials would suggest an awareness of the fifth-order probabilities. Each participant was informed prior to beginning the prediction task that there were a total of 6 participants in the study and that the individual with the highest score on the prediction task would receive a bonus payment of $30.[1]

Procedure. One participant was randomly assigned to each of the six versions of Table 2. At the beginning of Session 1, the SRT task was described to participants, and they 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 16, participants performed the prediction task.

Data analysis. The 16 sessions were divided into eight epochs that each spanned two sessions. For each of the six versions of Table 2, there were 64 contexts with L/H successors. For each participant, the median RT of responses in an epoch (excluding incorrect responses and the first five trials of each block) was determined for each of the 64 L and H successors. The 64 data points for L successors were averaged, as were the 64 data points for H successors. The averaged scores were submitted to an ANOVA with successor (L, H) and epoch (1–8) as within-subjects factors. Tests for the effect of successor and the successor × epoch lin interaction (where *epoch lin* is the linear component of the epoch factor) were one-tailed. I was interested in detecting shorter RTs on H than on L successors and an increase in the RT difference across epochs. Error rates were analyzed in a manner identical to that for RTs. Alpha was .05. The effect size measure, $r_{contrast}$, is reported for all F tests with numerator $df = 1$ (Rosenthal, Rosnow, & Rubin, 2000).

Results and Discussion

Reaction time. The results appear in Figure 1. For the RT data, there was an effect of successor [$F(1,5) = 7.68$, $MS_e = 49.91$, $p = .020$; $r_{contrast} = .78$]. The successor × epoch lin interaction was not significant [$F(1,5) = 0.00$, $MS_e = 4.62$, $p = .490$; $r_{contrast} = 0$]. Thus, RT was shorter on H than on L successors and the difference did not increase significantly across epochs. This suggests that there was learning of at least 1 of the 16 types of fifth-order probability and that most of the learning occurred early in training. Indeed, there was an effect of successor in the first epoch [$F(1,5) = 5.22$, $MS_e = 11.26$, $p = .036$; $r_{contrast} = .71$].

A reviewer expressed concern that in the absence of a successor × epoch lin interaction, the effect of successor might be the result of an artifact rather than learning sequential contingencies. This is unlikely for three reasons. First, the sequences of target locations in the present experiment were rigorously controlled. Given the rigorous methodological controls, it is difficult to imagine what the artifact might be.

Second, I separated the M successors into two sets, M1 and M2, just like the successors in the other tiers were separated into two sets, L and H. For example, in Table 2, the M successors in the left column in Tier 2 were labeled *M1* if they were in an odd-numbered row and *M2* if they were in an even-numbered row. The labeling was reversed for M successors in the right column. The RT difference between M1 and M2 successors, averaged across the eight epochs, was not significant [$M = 0.5$ msec, $p = .152$ (one-tailed)]. In contrast, the RT difference between L and H successors was significant in each of the two tiers with L/H successors.[2] It is difficult to see how an artifact could selectively affect the two tiers with L/H successors, but not the tier with M successors.

Third, other studies using the same six-choice SRT task in which each target location is followed by two possible successors have also obtained effects of successor in the absence of successor × session interactions (Remillard, 2008a; Remillard & Clark, 2001). Those studies exam-
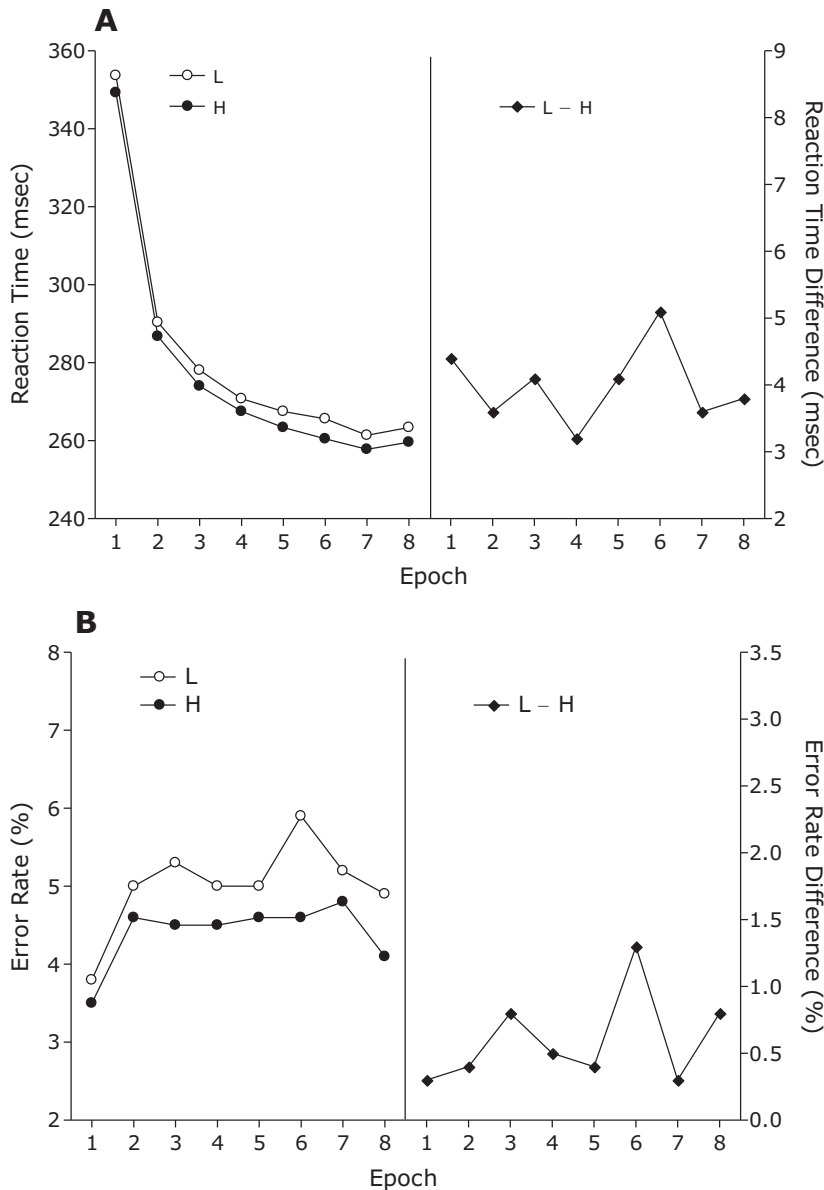
Figure 1. Reaction time (top panel) and error rate (bottom panel) as a function of successor (low probability, L; high probability, H) and epoch (1–8) in Experiment 1. The right side of each panel plots the difference between L and H successors. Each epoch spanned two sessions (i.e., 4,464 trials).

ined learning of first- through fourth-order probabilities. Across five experiments involving 11 probability conditions, the successor × session interaction failed to reach significance in 7 of the 11 conditions. Despite the absence of successor × session interactions, the learning effects (i.e., RT differences between L and H successors) were very orderly. For example, the learning effect decreases as one progresses from first-order to second-order to third-order to fourth-order adjacent probabilities of .40 versus .60. Also, the learning effect is greater when second-order adjacent and nonadjacent probabilities of .40 versus .60

are confounded than when second-order adjacent probabilities vary and nonadjacent probabilities are held constant. This is also the case with third- and fourth-order probabilities. Finally, there is no learning effect when fourth-order adjacent probabilities are .40 versus .60, but a learning effect emerges when the probabilities are widened to .33 versus .67. It is difficult to see how an artifact could produce such orderly results.

Because the current paradigm tends to produce rapid learning of sequential contingencies, the appropriate measure of learning is not an increasing difference in RT be-

tween L and H successors, but rather the difference in RT between L and H successors. Rowland and Shanks (2006) make a similar point when they note that "learning tends to be quite rapid in the probabilistic SRT task . . . so the critical statistical evidence for learning is a main effect of target probability rather than a probability $\times$ block interaction" (p. 291).

Error rate. For the error rate data, there was an effect of successor [$F(1,5) = 11.33$, $MS_e = 0.76$, $p = .010$; $r_{contrast} = .83$]. The successor $\times$ epoch lin interaction was not significant [$F(1,5) = 0.57$, $MS_e = 0.75$, $p = .242$; $r_{contrast} = .32$]. Thus, error rate was smaller on H than on L successors and the difference did not increase significantly across epochs. These results parallel the RT results.

Awareness. The percentage of the 64 trials on the prediction task that received a correct response was determined for each participant. The mean percentage correct (52.9%) did not differ significantly from what would be expected by random guessing (50%) [$F(1,5) = 1.60$, $MS_e = 30.68$, $p = .261$; $r_{contrast} = .49$]. Thus, there was no evidence for awareness of the fifth-order probabilities.

One could argue that the mean percentage correct would have been significantly greater than 50% with a larger sample of participants. To ensure that RT differences between L and H successors were not due to contexts for which participants may have had explicit knowledge of the fifth-order probabilities, two steps were taken. First, sets of matched contexts were identified that, on average, yielded correct responses on the prediction task. For example, the contexts in Rows 6, 38, and 70 of Table 2 was one such set of matched contexts. Each of these contexts and their successors are matched with respect to the pattern of W/B transitions and to type of run (see Table 3). For a participant, two of the three contexts in the set had L/H successors. The identity of the two contexts depended on the version of Table 2 that the participant received. For this set of contexts, participants selected, on average, H successors over L successors on the prediction task. Second, all analyses described above were rerun with the exclusion of the 16 sets of matched contexts identified in the first step. Removing sets of matched contexts ensured that the remaining contexts were also matched. For example, the contexts in Rows 5, 37, and 69 and their successors were retained and are matched with respect to the pattern of W/B transitions and to type of run. Also, for the contexts that were retained, the assignment of successors as L or H was fully counterbalanced across the six versions of Table 2.

Not surprisingly, the mean percentage correct on the prediction task dropped from 52.9% in the original analysis to 40.1% in the new analysis. The latter value was less than 50% [$F(1,5) = 13.18$, $MS_e = 44.60$, $p = .015$; $r_{contrast} = .85$]. Removing contexts that tended to elicit correct responses on the prediction task had little impact on the RT difference between L and H successors. In the new analysis, there was an effect of successor [$F(1,5) = 8.23$, $MS_e = 72.88$, $p = .018$; $r_{contrast} = .79$], and the successor $\times$ epoch lin interaction was not significant [$F(1,5) = 0.27$, $MS_e = 16.26$, $p = .314$; $r_{contrast} = .23$]. The RT difference between L and H successors, averaged across the eight epochs, was 4.0 msec in the original analysis and 5.0 msec in the new analysis. Thus, it is unlikely that RT differences between L and H successors were due to contexts for which participants may have had explicit knowledge of the fifth-order probabilities.

I also calculated the RT difference between L and H successors, averaged across the eight epochs, for each of the 32 sets of matched contexts and selected the 8 sets that yielded the largest RT differences (i.e., the top quartile). For example, the contexts in Rows 5, 37, and 69 of Table 2 was one such set and the RT difference between L and H successors for this set was 12.1 msec (averaged across participants). The RT differences between L and H successors over the 8 sets ranged from 7.7 msec to 15.2 msec, with a mean of 11.4 msec. Limiting the analyses to the 8 sets of matched contexts, there was an effect of successor [$F(1,5) = 36.66$, $MS_e = 85.28$, $p = .001$; $r_{contrast} = .94$], and the successor $\times$ epoch lin interaction was not significant [$F(1,5) = 0.52$, $MS_e = 116.46$, $p = .252$; $r_{contrast} = .31$]. Also, the mean percentage correct on the prediction task (52.1%) did not differ significantly from 50% [$F(1,5) = 0.19$, $MS_e = 135.42$, $p = .679$; $r_{contrast} = .19$]. Thus, there was no evidence for awareness of the fifth-order probabilities for those contexts for which RT differences between L and H successors were largest.

EXPERIMENT 2

Experiment 2 used a six-choice SRT task. The sequences of target locations were structured such that sixth-order probabilities varied (.15 vs. .85) and first-through fifth-order probabilities were constant (.50). Experiment 2 did not attempt to distinguish between learning of the 32 different types (1 adjacent and 31 nonadjacent) of sixth-order probability because these were confounded (e.g., P[3 | 3-2-6-3-2-1] = P[3 | 3-x-x-x-x-x] = P[3 | 3-2-x-x-x-1] = .15 and P[4 | 3-2-6-3-2-1] = P[4 | 3-x-x-x-x-x] = P[4 | 3-2-x-x-x-1] = .85). Thus, shorter RTs on high-probability successors (e.g., on 4, given 3-2-6-3-2-1, where P[4 | 3-2-6-3-2-1] = .85) than on low-probability successors (e.g., on 3, given 3-2-6-3-2-1, where P[3 | 3-2-6-3-2-1] = .15) would indicate learning of at least 1 of the 32 types of sixth-order probability. More specifically, it would indicate that people can learn to use the location of the target on trial $t-6$ to predict its location on trial t.

If people can learn sixth-order probabilities, then, on the basis of the results of Experiment 1 and studies by Remillard (2008a) and Remillard and Clark (2001), I expected much of the learning to occur early in training. Nonetheless, training was extended from 16 sessions in Experiment 1 to 30 sessions in Experiment 2, and successor probabilities were widened from .25/.75 in Experiment 1 to .15/.85 in Experiment 2. If participants showed no learning of the sixth-order probabilities, I wanted to be confident that this was due to a fundamental limitation of the sequence learning mechanism and not to the use of overly narrow successor probabilities or to a lack of training.

Awareness of the sixth-order probabilities was not assessed in Experiment 2. This would have required extending the prediction task to 128 trials, making it relatively lengthy.

Method

Participants. Six introductory psychology students, 18 to 21 years of age, participated for course credit and payment of $180.

SRT task. The SRT task was identical to that in Experiment 1 except that there were 30 sessions. Each session was composed of 18 blocks of trials with 131 trials per block. Session 1 began with a practice block of 110 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 30 sessions were completed in 38 to 46 days.

Structure of the sequences of target locations. Each target location had two possible successors, as in Experiment 1. Consequently, there were 192 (6×2^5) possible contexts of length 6, each followed by two possible successors. Letting the numbers 1 to 6 represent the six target locations from left to right, respectively, Table 4 presents some of the 192 contexts and the probabilities with which successors followed contexts. The 192 contexts were divided into three tiers, with each tier composed of 64 contexts.

The contexts in Tier 1 (Rows 1–64) began with Locations 3 or 4 and were succeeded by Locations 3 or 4. Every 20 occurrences of a context were followed 3 times by one successor (low-probability successor, L) and 17 times by the other successor (high-probability successor, H). For example, Row 3 indicates that every 20 occurrences of context 3-5-1-3-2-1 were followed 3 times by Successor 3 and 17 times by Successor 4, so that $P(3|3-5-1-3-2-1) = .15$ and $P(4|3-5-1-3-2-1) = .85$. Sixth-order probabilities in Tier 1 were .15 or .85. The odd-numbered rows reveal that whenever Location 3 was the target location on trial $t-6$, Locations 3 and 4 were the target locations on trial t with probabilities .15 and .85, respectively. Thus all 32 types of sixth-order probability with Location 3 on trial $t-6$ and Location 3 (Location 4) on trial t had a value of .15 (.85). For example, $P(3|3-5-1-3-2-1) = P(3|3-x-x-x-x-x) = P(3|3-5-x-x-x-x) = P(3|3-5-x-x-2-1) = .15$ and $P(4|3-5-1-3-2-1) = P(4|3-x-x-x-x-x) = P(4|3-5-x-x-x-x) = P(4|3-5-x-x-2-1) = .85$. The even-numbered rows reveal that whenever Location 4 was the target location on trial $t-6$, Locations 3 and 4 were the target locations on trial t with probabilities .85 and .15, respectively. Thus, all 32 types of sixth-order probability with Location 4 on trial $t-6$ and Location 3 (Location 4) on trial t had a value of .85 (.15). For example, $P(3|4-5-1-3-2-1) = P(3|4-x-x-x-x-x) = P(3|4-5-x-x-x-x) = P(3|4-5-x-x-2-1) = .85$ and $P(4|4-5-1-3-2-1) = P(4|4-x-x-x-x-x) = P(4|4-5-x-x-x-x) = P(4|4-5-x-x-2-1) = .15$. Importantly, all first- through fifth-order probabilities in Tier 1 were .50 (e.g., $P[3|1] = P[4|1] = .50$ and $P[3|5-x-x-x] = P[4|5-x-x-x] = .50$), and Locations 3 and 4 were each a target location equally often (i.e., $P[3] = P[4] = .17$). Thus, shorter RTs on H than on L successors would indicate that participants learned to use the location of the target on trial $t-6$, alone or in conjunction with subsequent target locations, to predict the location of the target on trial t (i.e., participants learned at least 1 of the 32 types of sixth-order probability). Tier 3 (Rows 129–192) was similar to Tier 1 except that contexts began with Locations 2 or 5 and were succeeded by Locations 2 or 5.

The contexts in Tier 2 (Rows 65–128) began with Locations 1 or 6 and were succeeded by Locations 1 or 6. Every 20 occurrences of a context were followed 10 times by one successor (medium-probability successor, M) and 10 times by the other successor (medium-probability successor, M). For example, Row 67 indicates that every 20 occurrences of context 1-4-2-1-3-2 was followed 10 times by Successor 1 and 10 times by Successor 6, so that $P(1|1-4-2-1-3-2) = .50$ and $P(6|1-4-2-1-3-2) = .50$. All first-through sixth-order probabilities in Tier 2 were .50.

The contexts and L (H) successors in Rows 1–64 of Table 4 were matched to the contexts and H (L) successors in Rows 129–192,

Table 4
Probabilities Inherent in the Sequences of Target Locations in Experiment 2

Tier	Row	Context	Successor
			1	2	3	4	5	6
1	1	3-2-1-3-2-1	–	–	L	H	–	–
	2	4-2-1-3-2-1	–	–	H	L	–	–
	3	3-5-1-3-2-1	–	–	L	H	–	–
	4	4-5-1-3-2-1	–	–	H	L	–	–
	5	3-2-6-3-2-1	–	–	L	H	–	–
	6	4-2-6-3-2-1	–	–	H	L	–	–
	7	3-5-6-3-2-1	–	–	L	H	–	–
	8	4-5-6-3-2-1	–	–	H	L	–	–
	.							
	57	3-2-1-4-5-6	–	–	L	H	–	–
	58	4-2-1-4-5-6	–	–	H	L	–	–
	59	3-5-1-4-5-6	–	–	L	H	–	–
	60	4-5-1-4-5-6	–	–	H	L	–	–
	61	3-2-6-4-5-6	–	–	L	H	–	–
	62	4-2-6-4-5-6	–	–	H	L	–	–
	63	3-5-6-4-5-6	–	–	L	H	–	–
	64	4-5-6-4-5-6	–	–	H	L	–	–
2	65	1-3-2-1-3-2	M	–	–	–	–	M
	66	6-3-2-1-3-2	M	–	–	–	–	M
	67	1-4-2-1-3-2	M	–	–	–	–	M
	68	6-4-2-1-3-2	M	–	–	–	–	M
	69	1-3-5-1-3-2	M	–	–	–	–	M
	70	6-3-5-1-3-2	M	–	–	–	–	M
	71	1-4-5-1-3-2	M	–	–	–	–	M
	72	6-4-5-1-3-2	M	–	–	–	–	M
	.							
	121	1-3-2-6-4-5	M	–	–	–	–	M
	122	6-3-2-6-4-5	M	–	–	–	–	M
	123	1-4-2-6-4-5	M	–	–	–	–	M
	124	6-4-2-6-4-5	M	–	–	–	–	M
	125	1-3-5-6-4-5	M	–	–	–	–	M
	126	6-3-5-6-4-5	M	–	–	–	–	M
	127	1-4-5-6-4-5	M	–	–	–	–	M
	128	6-4-5-6-4-5	M	–	–	–	–	M
3	129	2-1-3-2-1-3	–	H	–	–	L	–
	130	5-1-3-2-1-3	–	L	–	–	H	–
	131	2-6-3-2-1-3	–	H	–	–	L	–
	132	5-6-3-2-1-3	–	L	–	–	H	–
	133	2-1-4-2-1-3	–	H	–	–	L	–
	134	5-1-4-2-1-3	–	L	–	–	H	–
	135	2-6-4-2-1-3	–	H	–	–	L	–
	136	5-6-4-2-1-3	–	L	–	–	H	–
	.							
	185	2-1-3-5-6-4	–	H	–	–	L	–
	186	5-1-3-5-6-4	–	L	–	–	H	–
	187	2-6-3-5-6-4	–	H	–	–	L	–
	188	5-6-3-5-6-4	–	L	–	–	H	–
	189	2-1-4-5-6-4	–	H	–	–	L	–
	190	5-1-4-5-6-4	–	L	–	–	H	–
	191	2-6-4-5-6-4	–	H	–	–	L	–
	192	5-6-4-5-6-4	–	L	–	–	H	–

Note—Across every 20 occurrences of a context with L/H successors, the L and H successors occurred 3 times and 17 times, respectively. Across every 20 occurrences of a context with M successors, each M successor occurred 10 times. A dash indicates that the transition did not occur. L, low-probability successor; H, high-probability successor; M, medium-probability successor.

respectively, with respect to the pattern of W/B transitions and the response hand for the successor. For example, the context and L successor in Row 3 and the context and H successor in Row 131 both form a BBWWWW pattern, and both successors require a left-hand

response. Also, contexts and L (H) successors in Rows 1–64 of Table 4 were matched to contexts and H (L) successors in Rows 129–192, respectively, with respect to type of run (see Table 3). For example, the context (excluding the first element) and L successor in Row 3 and the context (excluding the first element) and H successor in Row 131 are both UEE runs.

For each participant, a 67,506-element sequence was randomly generated with the constraint that across every 20 occurrences of a context with L/H (M) successors, the L successor occurred 3 times and the H successor occurred 17 times (each M successor occurred 10 times). Elements 1–131, 126–256, 251–381, and so forth to 67,376–67,506 each constituted a block of 131 trials, for a total of 540 blocks (30 sessions × 18 blocks per session). A computer program went over each 67,506-element sequence and determined the exact values of all first- through sixth-order probabilities to ensure the values were as expected. The practice block of 110 trials at the beginning of Session 1 was randomly generated with the constraint that across every 2 occurrences of a context, each of its successors occurred once. There were six versions of Table 4. These were generated in a manner similar to that described in Experiment 1.

Procedure. One participant was randomly assigned to each of the six versions of Table 4. At the beginning of Session 1, the SRT task was described to participants and they 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.

Data analysis. The 30 sessions were divided into 10 epochs that each spanned three sessions. For each of the six versions of Table 4, there were 128 contexts with L/H successors. Ideally, one would like to obtain median RTs for each of the 128 L and H successors. Unfortunately, the number of observations for each of the L successors was quite small. Consequently, the 128 L and H successors were paired on the basis of their similarity with respect to the pattern of W/B transitions and to type of run. For example, the L (H) successors in Rows 1 and 64 of Tier 1 of Table 4 were paired because both, along with their contexts, form a WWWWWW (WWWWWB) pattern and, excluding the first element of the contexts, both complete EEE (EEU) runs. Similarly, L/H successors in Rows 2 and 63, 3 and 62, and so forth to 32 and 33 were paired. A similar strategy was employed for the other tiers of Table 4.

For each participant, the median RT of responses in an epoch (excluding incorrect responses and the first six trials of each block) was determined for each of the 64 paired L and H successors. The 64 data points for L successors were averaged, as were the 64 data points for H successors. The averaged scores were submitted to an ANOVA with successor (L, H) and epoch (1–10) as within-subjects factors. Tests for the effect of successor and the successor × epoch lin interaction were one-tailed. Error rates were analyzed in a manner identical to that for RTs.

Results and Discussion

The results appear in Figure 2. For the RT data, there was an effect of successor [$F(1,5) = 39.82$, $MS_e = 30.88$, $p < .001$; $r_{contrast} = .94$]. The successor × epoch lin interaction approached significance [$F(1,5) = 3.78$, $MS_e = 2.13$, $p = .055$; $r_{contrast} = .66$]. Thus, RT was shorter on H than on L successors and the difference tended to increase across epochs. This suggests that there was learning of at least 1 of the 32 types of sixth-order probability. There was an effect of successor in the first epoch [$F(1,5) = 25.76$, $MS_e = 6.50$, $p = .002$; $r_{contrast} = .92$], indicating the presence of learning relatively early in training.

For the error rate data, there was an effect of successor [$F(1,5) = 32.66$, $MS_e = 4.24$, $p = .001$; $r_{contrast} = .93$]. The successor × epoch lin interaction was significant [$F(1,5) = 5.75$, $MS_e = 2.11$, $p = .031$; $r_{contrast} = .73$]. Thus, error rate was smaller on H than on L successors and the difference increased across epochs. These results parallel the RT results.

The RT and error rate difference between L and H successors tended to increase across epochs in Experiment 2 but not in Experiment 1. The reason for the discrepancy is not clear. Participants in Experiment 2 experienced 30 sessions of training and successor probabilities of .15 versus .85. Perhaps a successor × epoch lin interaction would have emerged in Experiment 1 had participants in that experiment experienced more training sessions or wider successor probabilities. Unfortunately, almost nothing is known about the conditions under which implicit sequence learning reaches asymptote quickly or more slowly.

The RT and error rate difference between L and H successors were also numerically greater in Experiment 2 than in Experiment 1. This may seem surprising, given that the sequences of target locations involved higher order probabilities in Experiment 2 than in Experiment 1. However, the result was undoubtedly due to the use of wider successor probabilities in Experiment 2 (.15/.85) than in Experiment 1 (.25/.75). The RT and error rate difference between L and H successors would likely have been smaller in Experiment 2 than in Experiment 1 had both experiments employed the same successor probabilities (e.g., see Remillard, 2008a; Remillard & Clark, 2001).

GENERAL DISCUSSION

Participants in the present study learned to use the location of the target on trial $t-5$ (Experiment 1) or $t-6$ (Experiment 2) to predict its location on trial t, and much of the learning occurred in the first epoch. Learning in Experiment 1 was implicit. Awareness of the sixth-order probabilities was not assessed in Experiment 2. However, if awareness is assumed to have been nonexistent in the first epoch, then awareness was likely not responsible for the RT differences between L and H successors because most of the difference emerged in the first epoch.

The results of Experiment 2 suggest that the implicit sequence learning mechanism operates over a range of at least seven sequence elements (the six-element context and successor). This exceeds the capacity of spatial working memory, which is estimated to be five or six spatial locations in college students (Nelson, Dickson, & Banos, 2000; Orsini et al., 1986; Vandierendonck, Kemps, Fastame, & Szmalec, 2004). When presented with a sequence of seven spatial locations, college students recall, on average, 50% to 70% of the sequence elements in the correct position (Fischer, 2001). Thus, the results of Experiment 2 support the hypothesis presented in the beginning of this article that implicit sequence learning is not bound by the capacity limits of working memory.

It has been proposed that implicit sequence learning involves the operation of associative processes on sequence elements stored in a short-term memory system (Frensch, 1998, pp. 87–95; Frensch et al., 1994; Frensch & Miner, 1994). Computational models of sequence learning called
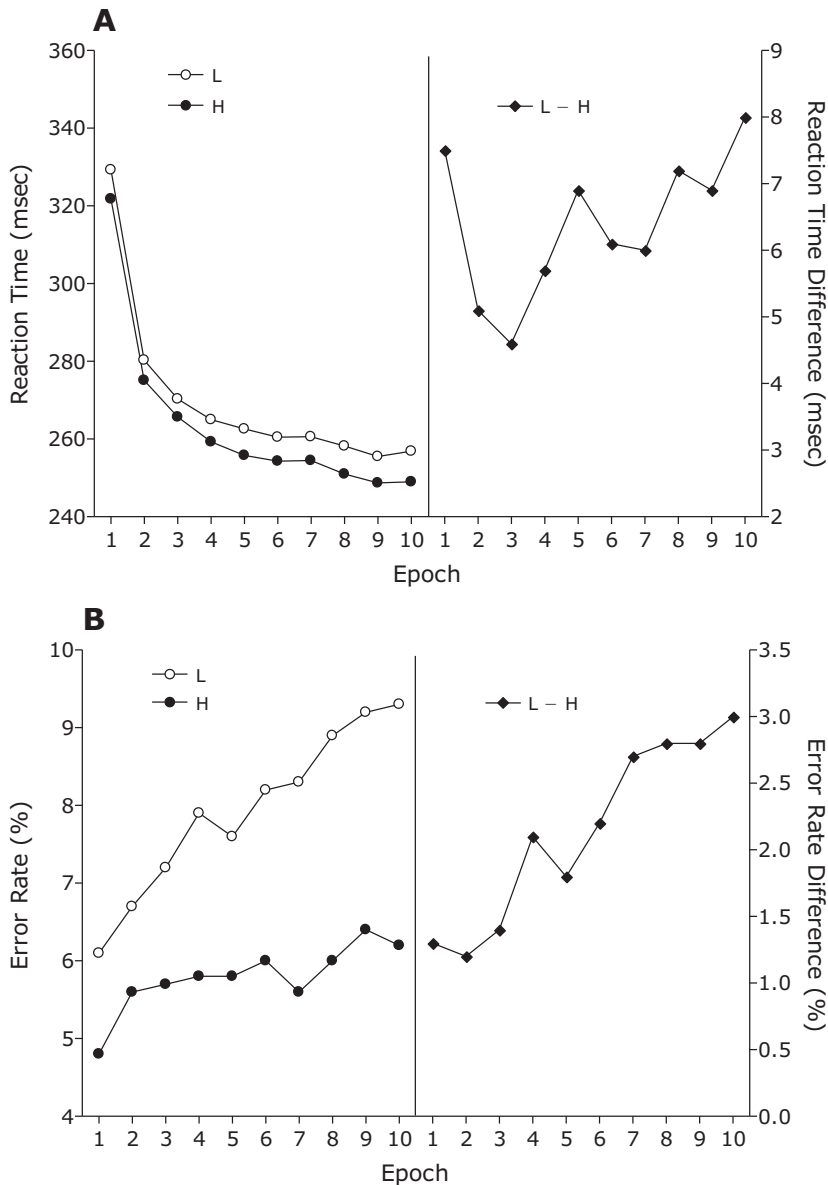
Figure 2. Reaction time (top panel) and error rate (bottom panel) as a function of successor (low probability, L; high probability, H) and epoch (1–10) in Experiment 2. The right side of each panel plots the difference between L and H successors. Each epoch spanned three sessions (i.e., 7,074 trials).

buffer networks implement such a short-term memory system (Cleeremans, 1993; Kuhn & Dienes, 2008). In Experiment 2, the time interval between the offset of the first element in a six-element context and the onset of the successor was approximately 3.4 sec (given a response–stimulus interval of 300 msec and an RT of 320 msec). Thus, the results of Experiment 2 suggest that the short-term memory system can hold at least seven sequence elements, with each element being held for at least 3.4 sec.

There is evidence that mechanisms of implicit sequence learning play a role in language acquisition (Conway,

Bauernschmidt, Huang, & Pisoni, 2010; Saffran, 2003). However, some aspects of language acquisition require learning remote sequential contingencies (Gomez & Maye, 2005), and it has not been clear whether mechanisms of implicit sequence learning are capable of acquiring such remote contingencies. The results of Experiments 1 and 2 suggest that the mechanism responsible for implicitly learning contingencies in visuospatial/motor sequences is capable. Whether the mechanism responsible for implicitly learning contingencies in auditory/verbal sequences is equally capable is an empirical question.

Earlier work indicates that the implicit sequence learning mechanism operates over adjacent as well as nonadjacent elements (Remillard, 2008a) and operates independently of individual differences in intelligence (Gebauer & Mackintosh, 2007). This, coupled with evidence that the mechanism is not bound by the capacity limits of working memory and can operate over a range of at least seven sequence elements, suggests that the implicit sequence learning mechanism is remarkably powerful. However, the range of sequence elements over which the mechanism can operate is clearly limited. The fact that the learning effect (i.e., the RT difference between L and H successors) decreases as one progresses from first-order to second-order to third-order to fourth-order adjacent probabilities (Remillard, 2008a; Remillard & Clark, 2001) suggests that implicit sequence learning becomes more difficult as the range of sequence elements that must be considered increases. More research will be required to establish the ultimate range of sequence elements over which the implicit sequence learning mechanism can operate and to identify factors that influence the operational range.

AUTHOR NOTE

REFERENCES

CLEEREMANS, A. (1993). *Mechanisms of implicit 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.

CONWAY, C. M., BAUERNSCHMIDT, A., HUANG, S. S., & PISONI, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, **114**, 356-371.

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

FISCHER, M. H. (2001). Probing spatial working memory with the Corsi blocks task. *Brain & Cognition*, **45**, 143-154.

FRENSCH, P. A. (1998). One concept, multiple meanings. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 47-104). Thousand Oaks, CA: Sage.

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

FRENSCH, P. A., & MINER, C. S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition*, **22**, 95-110.

FU, Q., FU, X., & DIENES, Z. (2008). Implicit sequence learning and conscious awareness. *Consciousness & Cognition*, **17**, 185-202.

GEBAUER, G. F., & MACKINTOSH, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 34-54.

GOMEZ, R., & MAYE, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, **7**, 183-206.

JIMÉNEZ, L., VAQUERO, J. M. M., & LUPIÁÑEZ, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 475-490.

KNEE, R., THOMASON, S., ASHE, J., & WILLINGHAM, D. T. (2007). The representation of explicit motor sequence knowledge. *Memory & Cognition*, **35**, 326-333.

KUHN, G., & DIENES, Z. (2008). Learning non-local dependencies. *Cognition*, **106**, 184-206.

LANY, J., & GOMEZ, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science*, **19**, 1247-1252.

LANY, J., GOMEZ, R. L., & GERKEN, L. A. (2007). The role of prior experience in language acquisition. *Cognitive Science*, **31**, 481-507.

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.

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.

MILLWARD, R. B., & REBER, A. S. (1972). Probability learning: Contingent-event schedules with lags. *American Journal of Psychology*, **85**, 81-98.

NELSON, R. E., DICKSON, A. L., & BANOS, J. H. (2000). An automated administration of Corsi's block-tapping test. *Perceptual & Motor Skills*, **91**, 578-580.

ORSINI, A., CHIACCHIO, L., CINQUE, M., COCCHIARO, C., SCHIAPPA, O., & GROSSI, D. (1986). Effects of age, education and sex on two tests of immediate memory: A study of normal subjects from 20 to 99 years of age. *Perceptual & Motor Skills*, **63**, 727-732.

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

REMILLARD, G. (2008a). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *Quarterly Journal of Experimental Psychology*, **61**, 400-424.

REMILLARD, G. (2008b). A program for generating randomized simple and context-sensitive sequences. *Behavior Research Methods*, **40**, 484-492.

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

ROSENTHAL, R., ROSNOW, R. L., & RUBIN, D. B. (2000). *Contrasts and effects sizes in behavioral research: A correlational approach*. Cambridge: Cambridge University Press.

ROWLAND, L. A., & SHANKS, D. R. (2006). Sequence learning and selection difficulty. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 287-299.

SAFFRAN, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, **12**, 110-114.

SONG, S., HOWARD, J. H., JR., & HOWARD, D. V. (2007). Implicit probabilistic sequence learning is independent of explicit awareness. *Learning & Memory*, **14**, 167-176.

SONG, S., MARKS, B., HOWARD, J. H., JR., & HOWARD, D. V. (2009). Evidence for parallel explicit and implicit sequence learning systems in older adults. *Behavioural Brain Research*, **196**, 328-332.

STADLER, M. A. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, & 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, & Cognition*, **21**, 674-685.

UNSWORTH, N., & ENGLE, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, **33**, 213-220.

VANDIERENDONCK, A., KEMPS, E., FASTAME, M. C., & SZMALEC, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, **95**, 57-79.

WILLINGHAM, D. B., & GOEDERT-ESCHMANN, K. (1999). The relation between implicit and explicit learning: Evidence for parallel development. *Psychological Science*, **10**, 531-534.

WILLINGHAM, D. B., GREELEY, T., & BARDONE, A. M. (1993). Disso-

ciation in a serial response time task using a recognition measure: Comment on Perruchet and Amorim (1992). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1424-1430.

WILLINGHAM, D. B., GREENBERG, A. R., & THOMAS, R. C. (1997). Response-to-stimulus interval does not affect implicit motor sequence learning, but does affect performance. *Memory & Cognition*, **25**, 534-542.

WILLINGHAM, D. B., SALIDIS, J., & GABRIELI, J. D. E. (2002). Direct comparison of neural systems mediating conscious and unconscious skill learning. *Journal of Neurophysiology*, **88**, 1451-1460.

NOTES

1. One might argue that the prediction task is not optimally sensitive to explicit knowledge of the fifth-order probabilities because participants do not respond to the location of the target as they do in the SRT task. The empirical evidence suggests otherwise. Willingham, Greeley, and Bardone (1993) showed that a recognition test requiring participants to observe a sequence of target locations is as sensitive to explicit sequence knowledge as is a recognition test requiring participants to respond to the location of target, as in the SRT task. This makes sense in light of evidence suggesting that explicit sequence knowledge is coded in terms of visuospatial locations and not in terms of motor responses or response locations (Knee, Thomason, Ashe, & Willingham, 2007). Also, a monetary reward was offered for good performance on the prediction task. A monetary reward can increase a measure's sensitivity to explicit sequence knowledge (Fu, Fu, & Dienes, 2008).

2. The two tiers with L/H successors were labeled Tier A and Tier B. L/H successors in Tier A immediately followed M successors in the sequence of target locations (e.g., Tier 1 in Table 2), and L/H successors in Tier B immediately preceded M successors (e.g., Tier 3 in Table 2). The RT difference between L and H successors, averaged across the eight epochs, was significant in Tier A ($M = 3.9$ msec, $p = .020$, one-tailed) and Tier B ($M = 4.1$ msec, $p = .035$, one-tailed).