Psychological Research (2011) 75:307–323
DOI 10.1007/s00426-010-0309-0

ORIGINAL ARTICLE

Pure perceptual-based learning of second-, third-, and fourth-order sequential probabilities

Gilbert Remillard

Received: 17 May 2010 / Accepted: 30 August 2010 / Published online: 15 September 2010

Abstract There is evidence that sequence learning in the traditional serial reaction time task (SRTT), where target location is the response dimension, and sequence learning in the perceptual SRTT, where target location is not the response dimension, are handled by different mechanisms. The ability of the latter mechanism to learn sequential contingencies that can be learned by the former mechanism was examined. Prior research has established that people can learn second-, third-, and fourth-order probabilities in the traditional SRTT. The present study reveals that people can learn such probabilities in the perceptual SRTT. This suggests that the two mechanisms may have similar architectures. A possible neural basis of the two mechanisms is discussed.

Introduction

Implicit sequence learning is sequence learning that is not the result of conscious, intentional processes and has been studied using the serial reaction time task (SRTT). On each trial in the traditional SRTT, 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 (e.g., Nissen & Bullemer, 1987). 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 (e.g., Remillard, 2008a). Sequence learning occurs when, given previous target locations, more probable succeeding locations elicit shorter RTs than do less probable succeeding locations.

Implicit sequence learning is usually established 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).

The sequence of target locations is correlated with the sequence of responses in the traditional SRTT because the key corresponding to the location of the target must be pressed. In the perceptual SRTT, target location is not the response dimension and the sequence of target locations is uncorrelated with the sequence of responses. For example, people might respond to the identity of the target on each trial, not its location, and the sequence of target identities (and hence responses) is uncorrelated with the sequence of target locations (e.g., Mayr, 1996; Remillard, 2003). Under these circumstances, people can implicitly learn the sequence of target locations.

The presence of a distractor that captures visuospatial attention impairs sequence learning in the perceptual SRTT (Remillard, 2009), but not in the traditional SRTT (Deroost, Coomans, & Soetens, 2009; Rowland & Shanks, 2006a, b; also see Jimenez & Vazquez, 2008). This suggests that sequence learning in the perceptual SRTT

G. Remillard (✉)
Department of Psychology, Morehead State University,
Morehead, KY 40351, USA
e-mail: g.remillard@moreheadstate.edu

depends on the visuospatial attention orienting system,[1] whereas sequence learning in the traditional SRTT does not. Thus, there appears to be at least two sequence learning mechanisms. One mechanism learns sequences of visuospatial locations and depends on the visuospatial attention system. The other mechanism presumably learns sequences of responses and depends on a system other than the visuospatial attention system. Indeed, some studies suggest that learning in the traditional SRTT is primarily of a sequence of response (e.g., key) locations (Bischoff-Grethe, Goedert, Willingham, & Grafton, 2004; Willingham, 1999; Willingham, Wells, Farrell, & Stemwedel, 2000) and that learning depends on response selection processes (Deroost & Soetens, 2006a; Schumacher & Schwarb, 2009; Schwarb & Schumacher, 2009).[2] Also consistent with the idea of at least two mechanisms is evidence that learning a sequence of visuospatial locations and learning a sequence of responses can proceed independently of one another (Mayr, 1996) and that individuals with Parkinson's disease are impaired at learning a sequence of responses, but not a sequence of visuospatial locations (Helmuth, Mayr, & Daum, 2000).

The architectural similarity of the two sequence learning mechanisms is unknown. One way to begin to explore the issue is to examine the kinds of sequential contingencies that each mechanism can learn. If both mechanisms can learn the same kinds of contingencies, this would suggest that the two mechanisms have similar architectures. Conversely, if one mechanism can learn more complex contingencies than the other mechanism, this would suggest that the two mechanisms have dissimilar architectures.

Using the traditional SRTT, Remillard (2008a) and Remillard and Clark (2001) have shown that people can learn first-, second-, third-, and fourth-order probabilities embedded in probabilistic sequences. *N*th-order probabilities are a class of conditional probabilities that involve events as far back as trial $t - n$, where trial t is the current trial (see Table 1). *N*th-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). Nonadjacent probabilities involve events that skip over at least one trial (e.g., lag 2-*x* and lag 4-*x*-2-*x* probabilities). More recently, Remillard (2010) has shown that people can learn fifth- and sixth-order probabilities. Thus, the mechanism for learning a sequence of responses in the traditional SRTT is remarkably powerful. The goal of the present study was to determine if the mechanism for learning a sequence of visuospatial locations in the perceptual SRTT is capable of learning second- or higher-order probabilities.

People can learn first-order probabilities in the perceptual SRTT (e.g., Remillard, 2003, 2009). However, in none of the three perceptual SRTT paradigms has it been clearly established that people can learn second-order probabilities. The *observation* paradigm has people observe a sequence of target locations without making any kind of response to the location of the target. Studies employing the observation paradigm have either failed to obtain evidence for learning of a second-order sequence of target locations (Kelly, Burton, Riedel, & Lynch, 2003; Song, Howard, & Howard, 2008, Experiment 1), or found that such learning is accompanied by substantial awareness of the sequence, raising the possibility that learning was explicit rather than implicit (Bird, Osman, Saggerson, & Heyes, 2005; Kelly et al., 2003). An apparent exception is Song et al. (2008, Experiment 2) who had participants observe a complex second-order sequence of target locations and try to explicitly learn the sequence. Although participants were unable to explicitly learn the sequence, a transfer test in which participants responded to the location of the target with a corresponding keypress suggested that participants had learned the sequence through observation. Unfortunately, participants may have responded to the location of the target during observation with subvocal responses, such as "1", "2", "3", and "4" (where numbers represent target locations) in an attempt to explicitly learn the sequence. Thus, the task may not have been purely observational. The observation paradigm is problematic because it is difficult to know what people are doing during observation. People may be (1) ignoring the target, (2) trying to explicitly learn the sequence of target locations, or (3) responding in a manner other than visual orienting (e.g., subvocally) to the location of the target. The next two paradigms are less likely to be affected by these issues.

[1] There is a growing body of evidence that the visuospatial attention system and the oculomotor system are tightly linked. For example, a covert shift of attention to a location is accompanied by the preparation of an eye movement to that location (Belopolsky & Theeuwes, 2009; Van der Stigchel & Theeuwes, 2007). Also, covert shifts of attention activate brain regions that are nearly identical to those activated by saccadic eye movements (de Haan, Morgan, & Rorden, 2008; Ikkai & Curtis, 2008). And, a covert shift of attention cannot be directed to a location that cannot be accessed physically by an eye movement (Craighero, Nascimben, & Fadiga, 2004; Smith, Ball, Ellison, & Schenk, 2010). Thus sequence learning in the perceptual SRTT might involve the oculomotor system. However, until more is known about the relationship between the visuospatial attention system and the oculomotor system, it will be difficult to establish the latter system's role in sequence learning in the perceptual SRTT.

[2] There may also be some learning of the sequence of visuospatial locations in the traditional SRTT (Clegg, 2005; Keele, Jennings, Jones, Caulton, & Cohen, 1995; Stadler, 1989). Such learning, however, appears to have minimal influence on performance measures of sequence learning because, as noted earlier, the presence of a distractor that captures visuospatial attention impairs learning of a sequence of visuospatial locations but has no effect on performance measures of sequence learning in the traditional SRTT.

Table 1 Types of probabilities

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

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 $P(E|A_3\text{-}A_2\text{-}A_1)$ 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 $P(E|A_4\text{-}x\text{-}A_2\text{-}x)$ 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

The *target identity* paradigm has people respond to the target's identity, not its location, and the sequence of target identities (and hence responses) is uncorrelated with the sequence of target locations. One study employing the target identity paradigm failed to obtain evidence for learning of a second-order sequence of target locations (Deroost & Soetens, 2006b, Experiment 3), whereas a second study did obtain such evidence (Mayr, 1996). However, participants in the latter study were exposed to a simple repeating sequence of target locations (ABCACB), where the locations were the vertices of a triangle. The sequence could have been learned as alternating clockwise and counterclockwise patterns of movements (ABCA and ACBA).

Finally, the *target-marked locations* (TML) paradigm is identical to the target identity paradigm except that the locations are marked with the targets rather than with short lines or outline boxes. For example, Remillard (2003; also see Deroost & Soetens, 2006c) marked six locations with the bigrams *xo* and *ox*. On each trial, an underline appeared below a bigram marking one of the locations and participants pressed the key corresponding to the identity of the underlined target bigram (left key for *xo* and right key for *ox*). Immediately following a correct response, the underline disappeared and the bigrams marking the locations were pseudorandomly reordered. The next trial began 400 ms later (the response–stimulus interval) with the appearance of the underline. Remillard showed that the TML paradigm is better able to detect learning of the sequence of target locations than is the target identity paradigm. This is because during the response–stimulus interval, participants orient to the anticipated target location and process the bigram marking that location. The result is an RT benefit if the underline appears at the anticipated location or cost if the underline appears at an unanticipated location marked with a different bigram. Remillard examined learning of second-order adjacent probabilities using the TML paradigm. The learning effect was very small and inconsistent across the four sessions of training. Thus, it was not clear whether the effect was an artifact or the result of real, but weak learning.

Experiment 1

People can learn second-order adjacent probabilities in the traditional SRTT (e.g., Remillard, 2008a; Remillard & Clark, 2001). It is not known whether people can learn such probabilities in the perceptual SRTT. Experiment 1 used the TML paradigm and examined learning of second-order adjacent probabilities. There were two targets. One target was a pair of vertical lines with a broken left line and a solid right line. This target required a left key response. The other target had a solid left line and a broken right line, and required a right key response. There were six horizontally arranged target locations. The sequence of target locations contained second-order adjacent probabilities. Specifically, given the target locations on trials $t-2$ and $t-1$, there was one low (0.33) and one high (0.67) probability successor on trial t. The sequence of targets and hence responses was unstructured and independent of the sequence of target locations. Thus, shorter RTs on high- than on low-probability successors would indicate learning of the second-order adjacent probabilities in the sequence of target locations.

On a trial, each of the six target locations was marked with one of the two pairs of vertical lines chosen randomly with the constraint that locations 1 versus 6, 2 versus 5, and 3 versus 4 were marked with different pairs (see Fig. 1). This ensured that there were three of each pair and that low- and high-probability successors were marked with different pairs. The latter follows from the fact that locations 1 versus 6, 2 versus 5, and 3 versus 4 were complements; that is, given the target locations on trials $t-2$ and $t-1$, if one location (e.g., location 3) was the low-probability successor on trial *t*, then its complement (e.g., location 4) was the high-probability successor. After a 400-ms delay, an underline appeared below the pair marking the next location in the sequence of target locations. Immediately following a correct response to the identity of the underlined target, the next trial began. Because the pair of vertical lines marking a location was chosen pseudorandomly on each trial, each pair was equally likely to mark each location.

Remillard (2003, Experiment 4) used sequences of target locations similar in structure to those in Experiment 1 and found that the RT difference between low- and high-probability successors was very small and inconsistent across the four sessions of training. Thus, it was not clear whether the RT difference was an artifact or the result of real, but weak learning of the second-order adjacent probabilities. Experiment 1 increased training to nine sessions to get a better indication of whether second-order adjacent probabilities can be learned.

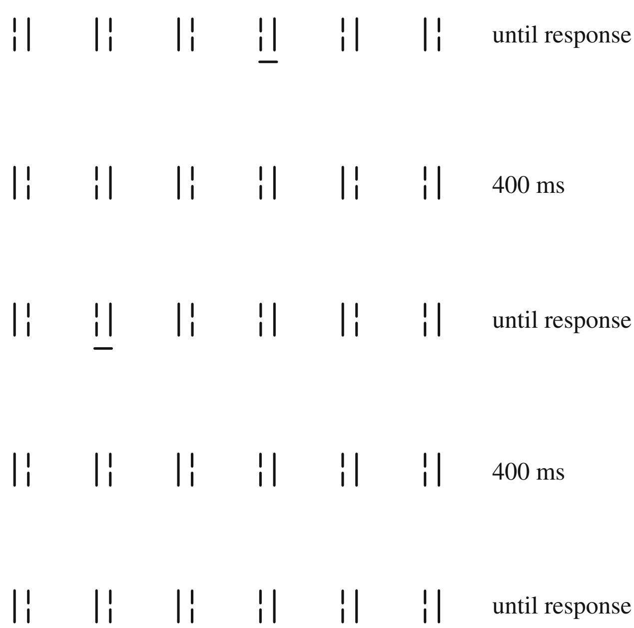

Fig. 1 An example sequence of events corresponding to the sequence of target locations 4-2-1 and the sequence of targets left broken line–left broken line–right broken line. Location 4 could be followed by locations 2 or 5, which are marked by different pairs of vertical lines (*row* 2), and location 2 could be followed by locations 1 or 6, which are marked by different pairs of vertical lines (*row* 4)

The traditional SRTT studies by Remillard (2008a) and Remillard and Clark (2001) employed six horizontally arranged target locations with each target location having two possible successors. The present experiments also employed six horizontally arranged target locations with each target location having the same two possible successors as in those studies. Thus, any difference between the traditional SRTT studies and the present experiments with respect to learning second- or higher-order probabilities could not be attributed to a difference in the sequences of target locations that were used.

Method

Participants

The participants were 12 university undergraduates ranging in age from 18 to 26 years.

Perceptual SRTT

The perceptual SRTT was run on a personal computer with standard monitor and keyboard. The six target locations were horizontally arranged and each location was marked with a pair of vertical lines. Each line in a pair was 0.50 cm in height and 0.04 cm in width. There was a 0.08-cm gap halfway up either the left line or the right line. The other line was solid. The two lines were separated by an interval of 0.16 cm. Adjacent pairs were separated by an interval of 0.80 cm. The distance between the left line of the pair marking the leftmost target location and the right line of the pair marking the rightmost target location was 5.30 cm. The underline was 0.24 cm in width and 0.04 cm in height. The underline appeared 0.20 cm below a pair of vertical lines. All lines were white and the background was black. The viewing distance was approximately 50 cm. The red-stickered *F* and *J* response keys, on which were placed the left and right index fingers, corresponded to the target pair with the broken left and right line, respectively.

Trial *t* began with the appearance of the underline. The location of the underline was determined by the sequence of target locations. Participants pressed the key corresponding to the target pair above the underline. Immediately following a correct response, the underline was erased and the location markers were changed as follows: if the target location and target on trial $t+1$ were location *X* and pair Y, respectively, then pair Y marked location X. Pairs to mark the remaining locations were chosen randomly with the constraint that locations 1 versus 6, 2 versus 5, and 3 versus 4 were marked with different pairs. This ensured that low- versus high-probability successors were marked with different pairs (see the next section). After a 400-ms delay, trial $t + 1$ began with the appearance of the underline below the pair marking location X (see Fig. 1). RT was measured as the time between the appearance of the underline and the first response regardless of the response's correctness.

Participants performed the perceptual SRTT over nine sessions. Each session was composed of 15 blocks of trials with 113 trials per block. Session 1 began with a practice block of 99 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 nine sessions were completed in 8–16 days.

A performance history was provided at the end of each block of trials in a session. The numbers 1–15 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 underline in location 1 could be followed by its appearance in locations 3 or 4, and the appearance of the underline in location 2 could be followed by its appearance in locations 1 or 6. Consequently, there were 24 ($6 \times 2 \times 2$) possible contexts of length 3 each followed by two possible successors. Letting the numbers 1–6 represent the six target locations from left to right, respectively, Table 2 presents the contexts and the probabilities with which successors followed contexts. The 24 contexts were divided into three tiers with each tier composed of eight contexts.

The contexts in tier 1 (rows 1–8) began with locations 3 or 4 and were succeeded by locations 3 or 4. Every 15 occurrences of a context was followed 5 times by one successor (low-probability successor, L) and 10 times by the other successor (high-probability successor, H). For example, row 1 indicates that every 15 occurrences of context 3-2-1 was followed 5 times by successor 3 and 10 times by successor 4 so that $P(3|3\text{-}2\text{-}1) = 0.33$ and $P(4|3\text{-}2\text{-}1) = 0.67$. Second-order adjacent probabilities in tier 1 were 0.33 or 0.67. For example, the contexts in rows 5 and 6 indicate that when locations 2 and 6 occurred on trials $t - 2$ and $t - 1$, respectively, location 3 on trial t was an H successor every time so that $P(3|2\text{-}6) = 0.67$. Third-order adjacent probabilities were redundant with second-order adjacent probabilities (e.g., $P[3|3\text{-}2\text{-}6] = P[3|2\text{-}6] = 0.67$), thus adding no information over and above that provided by the second-order adjacent probabilities. First-order adjacent probabilities and lag 2-x probabilities were 0.50. For example, rows 5–8 indicate that when location 6 occurred on trial $t - 1$, location 3 on trial t was an H successor for two contexts and an L successor for two contexts so that $P(3|6) = 0.50$. Similarly, rows 1, 2, 5, and 6 indicate that when location 2 occurred on trial $t - 2$, location 3 on trial t was an L successor for two contexts and an H successor for two contexts so that $P(3|2\text{-}x) = 0.50$. Lag 3-x-x, 3-x-1, and 3-2-x probabilities

Table 2 Probabilities inherent in the sequences of target locations in Experiment 1

Row	Context	Successor					
		1	2	3	4	5	6
1	3-2-1	–	–	L	H	–	–
2	4-2-1	–	–	L	H	–	–
3	3-5-1	–	–	H	L	–	–
4	4-5-1	–	–	H	L	–	–
5	3-2-6	–	–	H	L	–	–
6	4-2-6	–	–	H	L	–	–
7	3-5-6	–	–	L	H	–	–
8	4-5-6	–	–	L	H	–	–
9	1-3-2	M	–	–	–	–	M
10	6-3-2	M	–	–	–	–	M
11	1-4-2	M	–	–	–	–	M
12	6-4-2	M	–	–	–	–	M
13	1-3-5	M	–	–	–	–	M
14	6-3-5	M	–	–	–	–	M
15	1-4-5	M	–	–	–	–	M
16	6-4-5	M	–	–	–	–	M
17	2-1-3	–	H	–	–	L	–
18	5-1-3	–	H	–	–	L	–
19	2-6-3	–	L	–	–	H	–
20	5-6-3	–	L	–	–	H	–
21	2-1-4	–	L	–	–	H	–
22	5-1-4	–	L	–	–	H	–
23	2-6-4	–	H	–	–	L	–
24	5-6-4	–	H	–	–	L	–

Across every 15 occurrences of a context with L/H successors, the L and H successors occurred 5 times and 10 times, respectively. Across every 14 occurrences of a context with M successors, each M successor occurred 7 times

L low-probability successor, *H* high-probability successor, *M* medium probability successor were 0.44, 0.50, or 0.56 and were not confounded with second-order adjacent probabilities. Finally, locations 3 and 4 were each a target location equally often (i.e., $P[3] = P[4] = 0.17$). Thus, shorter RTs on H than on L successors would be evidence for learning the second-order adjacent probabilities. Tier 3 (rows 17–24) 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 9–16) began with locations 1 or 6 and were succeeded by locations 1 or 6. Every 14 occurrences of a context was followed 7 times by one successor (medium probability successor, M) and 7 times by the other successor (medium probability successor, M). For example, row 11 indicates that every 14 occurrences of context 1-4-2 was followed 7 times by successor 1 and 7 times by successor 6 so that $P(1|1\text{-}4\text{-}2) = 0.50$ and $P(6|1\text{-}4\text{-}2) = 0.50$. All first- through third-order probabilities were 0.50 in tier 2.

Locations 1–3 and 4–6 form the left and right halves of the display, respectively. Contexts and L (H) successors in rows 17–24 of tier 3 were matched to contexts and H(L) successors in rows 1–8 of tier 1, respectively, with respect to the pattern of within-half (W) and between-half (B) transitions. For example, context 4-5-1 and H successor 3 in row 4 and context 5-6-3 and L successor 2 in row 20 both form the transition pattern WBW and both successors occur in the left half of the display.

Software developed by Remillard (2008b) was used to generate the sequences of target locations. For each participant, a 14,853-element sequence of target locations was randomly generated with the constraint that across every 15 (14) occurrences of a context with L/H(M) successors, the L successor occurred 5 times and the H successor occurred 10 times (each M successor occurred 7 times). Elements 1–113, 111–223, 221–333, and so forth to 14,741–14,853 each constituted a block of 113 trials for a total of 135 blocks (9 sessions × 15 blocks per session). 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 third-order adjacent and nonadjacent probabilities. The practice block of 99 trials at the beginning of session 1 was randomly generated with the constraint that each context in Table 2 was followed by each of its two possible successors twice. Thus, the sequence of target locations in the practice block was unstructured in that all probabilities were 0.50.

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.

Table 3 Probabilities inherent in the sequences of targets in Experiment 1

Row	Context	Successor	
		1	2
1	1-1-1	M	M
2	1-1-2	M	M
3	1-2-1	M	M
4	1-2-2	M	M
5	2-1-1	M	M
6	2-1-2	M	M
7	2-2-1	M	M
8	2-2-2	M	M

The numbers 1 and 2 correspond to the target pair with the broken left and right line, respectively. Across every 16 occurrences of a context, each M successor occurred 8 times

M medium probability successor

Structure of the sequences of targets

The sequences of targets, and hence left and right key responses, were unstructured and independent of the sequences of target locations. Letting the numbers 1 and 2 correspond to the target pair with the broken left and right line, respectively, Table 3 presents the eight possible contexts of length 3 and the probabilities with which successors followed contexts. For example, row 2 indicates that every 16 occurrences of context 1-1-2 was followed 8 times by successor 1 and 8 times by successor 2 so that $P(1|1\text{-}1\text{-}2) = P(2|1\text{-}1\text{-}2) = 0.50$ (medium probability successors, M).

For each participant, a 14,853-element sequence of targets was randomly generated with the constraint that across every 16 occurrences of a context in Table 3, one successor occurred 8 times and the other successor occurred 8 times. Elements 1–113, 111–223, 221–333, and so forth to 14,741–14,853 each constituted a block of 113 trials for a total of 135 blocks (9 sessions × 15 blocks per session). The practice block of 99 trials at the beginning of session 1 was randomly generated with the constraint that each context in Table 3 was followed by each of its two possible successors six times.

Awareness survey

The survey to assess awareness of the second-order adjacent probabilities was a 12-item paper-and-pencil test. An item described a context of length 2 and the context's two possible successors (e.g., 2 → 1 → 3 4 and 6 → 3 → 2 5, where the pairs of numbers following the last arrow in an 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 successor. For example, the item 2 → 6 → 3 4 required participants to imagine the underline moving from location 2 to location 6 and then to indicate whether the underline was more likely to have appeared next in location 3 or location 4 during training. For participants' reference while they completed the survey, each of the six target locations on the computer screen was marked with a pair of vertical lines as in the perceptual SRTT. Eight items pertained to L/H successors and four to M successors. Scores greater than 50% correct (random guessing performance) on the eight items pertaining to L/H successors would indicate awareness of the second-order adjacent probabilities.[3]

Procedure

Two participants were randomly assigned to each of the six versions of Table 2. At the beginning of session 1, the perceptual SRTT 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 9, the awareness survey was administered.

Data analysis

Table 4 identifies four types of five-element runs 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 is shorter for EE runs, where repetition of a bigram is correctly primed (e.g., 1-3-2-1 primes 3, and 3 occurs), than for UE runs (Remillard, 2003; Remillard & Clark, 2001). Likewise, RT to the last element is longer for EU runs, where repetition of a bigram is incorrectly primed (e.g., 1-4-2-1 primes 4, but 3 occurs), than for UU runs.

Table 4 Types of runs

Run	Example
EE	1-3-2-1-3
UE	6-3-2-1-3
EU	1-4-2-1-3
UU	6-4-2-1-3

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

To ensure that L and H successors were not differentially affected by the type of run and the transition pattern, the following procedure was employed. First, each of the 16 contexts of length 3 with L/H successors was extended to two contexts of length 4 by considering the two possible locations that could precede a context of length 3. For example, context 3-2-1 in row 1 of Table 2 could be preceded by 1 or 6 resulting in the contexts 1-3-2-1 and 6-3-2-1. The resulting 16 contexts and their L(H) successors in one tier of Table 2 were matched to the resulting 16 contexts and their H(L) successors in the other tier of Table 2 with respect to type of run and transition pattern. For example, context 1-3-2-1 with L successor 3 (extending the context in row 1) and context 3-2-1-3 with H successor 2 (extending the context in row 17) are both EE runs and both form a WWWW transition pattern. Similarly, context 6-4-2-1 with H successor 4 (extending the context in row 2) and context 4-5-1-3 with L successor 5 (extending the context in row 18) are both UE runs and both form a WBWB transition pattern. Second, the 9 sessions were divided into 3 epochs that each spanned three sessions and, for each participant, the median RT of responses in an epoch (excluding incorrect responses and the first four trials of each block) was determined for each of the 32 L and H successors that followed the 32 contexts. The 32 data points for L successors were averaged as were the 32 data points for H successors. The averaged scores were submitted to an analysis of variance (ANOVA) with successor (L, H) and epoch (1–3) as within-subject factors. Analyses involving epoch focused on the linear component of epoch (epoch Lin). Error rates were analyzed in a manner identical to that for RTs. Alpha was 0.05.

Detecting shorter RTs on H than on L successors was of primary interest. Also, there was no reason to expect the opposite pattern of longer RTs on H than on L successors because, to my knowledge, no SRTT study employing probabilistic sequences of target locations has ever

[3] One might argue that the awareness survey is not optimally sensitive to explicit knowledge of the second-order adjacent probabilities because it does not reinstate all of the cues that were present during the perceptual SRTT (e.g., the underline moving from one target location to the next). The empirical evidence, however, suggests the survey is sensitive. A recognition test where sequences of target locations are presented as sequences of digits is (a) as sensitive as a recognition test that requires participants to respond to the sequences of target locations as in the SRTT (Willingham, Greeley, & Bardone, 1993), (b) capable of detecting explicit sequence knowledge after very limited exposure to the sequence during the SRTT, suggesting that such a recognition test can pick up relatively low levels of awareness (Perruchet, Bigand, & Benoit-Gonin, 1997, Experiment 3), and (c) able to discriminate participants who are instructed to try to explicitly learn a sequence while performing the SRTT from participants who are not given such instructions (Curran, 1997). Also, the awareness survey is able to detect explicit knowledge of subtle first- and second-order adjacent dependencies of 0.40 versus 0.60 after extended training on the SRTT (Remillard & Clark, 2001, Experiment 4).

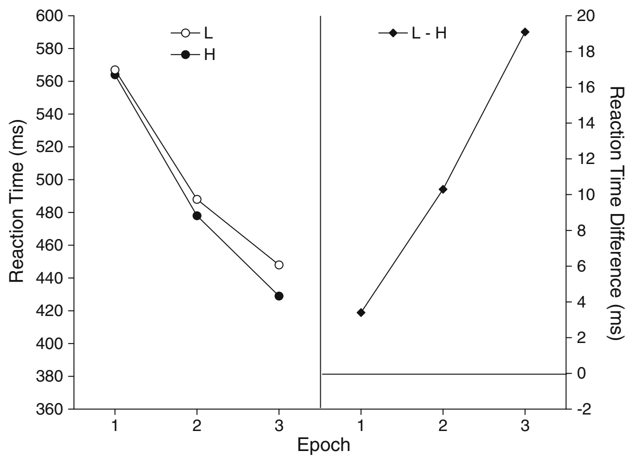

Fig. 2 The *left side* plots reaction time as a function of successor (low probability *L*, high probability *H*) and epoch (1–3) in Experiment 1. The *right side* plots the reaction time difference between L and H successors as a function of epoch. Each epoch spanned three sessions

obtained such a pattern. Therefore, to increase statistical power, the test for the effect of successor was one-tailed.

In addition to the usual statistics, I report 95% confidence intervals (CIs) for the more important analyses, and a 90% CI for the effect of successor. Confidence intervals are reported as (x, y, z) where x and z are the lower and upper bounds, respectively, and y is the central value. The effect of successor is significant if and only if the lower bound of the 90% CI is greater than zero.

Results and discussion

The results appear in Fig. 2. There was an effect of epoch Lin, $F(1, 11) = 132.11$, MSE = 1,461.75, $p < 0.001$, indicating that overall RT declined across epochs. There was an effect of successor, $F(1, 11) = 33.84$, MSE = 63.58, $p < 0.001$, 90% CI = (7.6, 10.9, 14.3 ms), and a Successor × Epoch Lin interaction, $F(1, 11) = 15.91$, MSE = 46.70, $p = 0.002$, 95% CI = (3.5, 7.9, 12.2 ms/epoch).[4] Thus, RT was shorter on H than on L successors and the difference increased across epochs. This indicates that participants learned the second-order adjacent probabilities. There were no significant effects involving error rate.

The percentage of the eight L/H items on the awareness survey receiving a correct response was determined for each participant. The mean percentage was 45.8% and did not differ significantly from what would be expected by random guessing (50%), $F(1, 11) = 1.16$, MSE = 179.92, $p = 0.305$, 95% CI = (37.3, 45.8, 54.4%). Thus, there was no evidence for awareness of the second-order adjacent probabilities.[5]

Experiment 2

Experiment 2 used the TML paradigm and examined learning of third-order adjacent probabilities. In a first attempt involving 12 participants, 10 sessions of training, and successor probabilities of 0.33 and 0.67, RT on H successors did not differ significantly from that on L successors. Averaging across the 5 epochs (with 2 sessions/epoch), the RT difference between L and H successors (i.e., L − H) was −0.5 ms. The negative value indicates that the effect of successor was not significant by a one-tail test. The Successor × Epoch Lin interaction was not significant, $F(1, 11) = 0.21$, $p = 0.656$, indicating that the RT difference between L and H successors did not change significantly with training. Indeed, the RT difference between L and H successors was 0.5 ms over the last 2 epochs of training and was not significant, $F(1, 11) = 0.08$, $p = 0.393$ (one-tail). Thus, there was no evidence that participants learned third-order adjacent probabilities of 0.33 versus 0.67. In a second attempt, reported here as Experiment 2, I used successor probabilities of 0.20 and 0.80, and 15 sessions of training. The goal was to try to establish whether people could learn third-order adjacent probabilities in the perceptual SRTT.

Method

Participants

The participants were 12 university undergraduates ranging in age from 18 to 20 years.

Perceptual SRTT

The perceptual SRTT was identical to that in Experiment 1 except as follows. Participants performed the perceptual SRTT over 15 sessions. Each session was composed of 16 blocks of trials with 113 trials per block. Session 1 began

[4] The central value of 10.9 ms for the 90% CI is the RT difference between L and H successors (i.e., L − H) averaged across the three epochs. The central value of 7.9 ms/epoch for the 95% CI is the slope of the regression line through the L successor RTs minus the slope of the regression line through the H successor RTs.

[5] RTs to the last element of each of the four types of runs in Table 4 were also examined. There were 96 (6×2^4) possible five-element sequences of which 24 were EE runs, 24 were UE runs, 24 were EU runs, and 24 were UU runs. For each participant and epoch, the median RTs to the last element of each of the 96 five-element sequences were computed, and then the 72 RTs for EE runs (24 RTs per epoch × 3 epochs) were averaged, as were the 72 RTs for UE runs, the 72 RTs for EU runs, and the 72 RTs for UU runs. Mean RTs on EE, UE, EU, and UU runs were, respectively, 496, 500, 497, and 490 ms. The RT difference between EE and UE runs was significant, $p = 0.010$, as was the difference between EU and UU runs, $p = 0.001$. The pattern was similar in Experiments 2 and 3.

Table 5 Probabilities inherent in the sequences of target locations in Experiment 2 (outside the parentheses) and Experiment 3 (in parentheses)

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

Across every 15(20) occurrences of a context with L/H successors, the L and H successors occurred 3(3) times and 12(17) times, respectively. Across every 14(20) occurrences of a context with M successors, each M successor occurred 7(10) times

L low-probability successor, *H* high-probability successor, *M* medium probability successor

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 15 sessions were completed in 19–24 days.

Structure of the sequences of target locations

Each target location had the same two possible successors as in Experiment 1. Consequently, there were 48 (6×2^3) possible contexts of length 4 each followed by two possible successors. Letting the numbers 1–6 represent the six target locations from left to right, respectively, Table 5 presents the contexts and the probabilities with which successors followed contexts. The 48 contexts were divided into three tiers with each tier composed of 16 contexts.

The contexts in tier 1 (rows 1–16) began with locations 1 or 6 and were succeeded by locations 3 or 4. Every 15 occurrences of a context was followed 3 times by one successor (low-probability successor, L) and 12 times by the other successor (high-probability successor, H). For example, row 2 indicates that every 15 occurrences of context 6-3-2-1 was followed 3 times by successor 3 and 12 times by successor 4 so that $P(3|6\text{-}3\text{-}2\text{-}1) = 0.20$ and $P(4|6\text{-}3\text{-}2\text{-}1) = 0.80$. Third-order adjacent probabilities in tier 1 were 0.20 or 0.80. For example, the contexts in rows 5 and 6 indicate that when locations 3, 5, and 1 occurred on trials $t - 3$, $t - 2$, and $t - 1$, respectively, location 3 on trial t was an H successor every time so that $P(3|3\text{-}5\text{-}1) = 0.80$. Fourth-order adjacent probabilities were redundant with third-order adjacent probabilities (e.g., $P[3|6\text{-}3\text{-}5\text{-}1] = P[3|3\text{-}5\text{-}1] = 0.80$), thus adding no information over and above that provided by the third-order adjacent probabilities. First- and second-order adjacent probabilities were 0.50, as were lag 2-*x*, lag 3-*x*-*x*, lag 3-2-*x*, and lag 3-*x*-1 probabilities. For example, rows 5–8 indicate that when locations 5 and 1 occurred on trials $t - 2$ and $t - 1$, respectively, location 3 on trial t was an H successor for two contexts and an L successor for two contexts so that

$P(3|5\text{-}1) = 0.50$. Similarly, rows 1, 2, 5, 6, 9, 10, 13, and 14 indicate that when location 3 occurred on trial $t - 3$, location 3 on trial t was an L successor for four contexts and an H successor for four contexts so that $P(3|3\text{-}x\text{-}x) = 0.50$. Fourth-order nonadjacent probabilities were 0.32, 0.50, or 0.68 and were not confounded with third-order adjacent probabilities. Finally, locations 3 and 4 were each a target location equally often (i.e., $P[3] = P[4] = 0.17$). Thus, shorter RTs on H than on L successors would be evidence for learning the third-order adjacent probabilities. Tier 3 (rows 33–48) was similar to tier 1 except that contexts began with locations 3 or 4 and were succeeded by locations 2 or 5.

The contexts in tier 2 (rows 17–32) began with locations 2 or 5 and were succeeded by locations 1 or 6. Every 14 occurrences of a context was followed 7 times by one successor (medium probability successor, M) and 7 times by the other successor (medium probability successor, M). For example, row 18 indicates that every 14 occurrences of context 5-1-3-2 was followed 7 times by successor 1 and 7 times by successor 6 so that $P(1|5\text{-}1\text{-}3\text{-}2) = 0.50$ and $P(6|5\text{-}1\text{-}3\text{-}2) = 0.50$. All first- through fourth-order probabilities were 0.50 in tier 2.

Contexts and L(H) successors in rows 33–48 of tier 3 were matched to contexts and H(L) successors in rows 1–16 of tier 1, respectively, with respect to type of run (see Table 4) and transition pattern. For example, context 6-3-5-1 and H successor 3 in row 6 and context 4-2-6-3 and L successor 2 in row 38 are both UE runs and both form a BBBW transition pattern.

For each participant, a 26,403-element sequence of target locations was randomly generated with the constraint that across every 15(14) occurrences of a context with L/H(M) successors, the L successor occurred 3 times and the H successor occurred 12 times (each M successor occurred 7 times). Elements 1–113, 111–223, 221–333, and so forth to 26,291–26,403 each constituted a block of 113 trials for a total of 240 blocks (15 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 5 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 0.50.

There were six versions of Table 5. These were created in a manner analogous to that described in Experiment 1.

Structure of the sequences of targets

The sequences of targets, and hence left and right key responses, were unstructured and independent of the sequences of target locations. Letting the numbers 1 and 2 correspond to the target pair with the broken left and right line, respectively, Table 6 presents the 16 possible contexts of length 4 and the probabilities with which successors followed contexts. For example, row 4 indicates that every 16 occurrences of context 1-1-2-2 was followed 8 times by successor 1 and 8 times by successor 2 so that $P(1|1\text{-}1\text{-}2\text{-}2) = P(2|1\text{-}1\text{-}2\text{-}2) = 0.50$ (medium probability successors, M).

For each participant, a 26,403-element sequence of targets was randomly generated with the constraint that across every 16 occurrences of a context in Table 6, one successor occurred 8 times and the other successor occurred 8 times. Elements 1–113, 111–223, 221–333, and so forth to 26,291–26,403 each constituted a block of 113 trials for a total of 240 blocks (15 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 6 was followed by each of its two possible successors three times.

Table 6 Probabilities inherent in the sequences of targets in Experiments 2 and 3

Row	Context	Successor	
		1	2
1	1-1-1-1	M	M
2	1-1-1-2	M	M
3	1-1-2-1	M	M
4	1-1-2-2	M	M
5	1-2-1-1	M	M
6	1-2-1-2	M	M
7	1-2-2-1	M	M
8	1-2-2-2	M	M
9	2-1-1-1	M	M
10	2-1-1-2	M	M
11	2-1-2-1	M	M
12	2-1-2-2	M	M
13	2-2-1-1	M	M
14	2-2-1-2	M	M
15	2-2-2-1	M	M
16	2-2-2-2	M	M

The numbers 1 and 2 correspond to the target pair with the broken left and right line, respectively. Across every 16 occurrences of a context, each M successor occurred 8 times

M medium probability successor

Prediction task

Awareness of the third-order adjacent probabilities was assessed using a prediction task. There were 16 prediction trials corresponding to the 16 contexts of length 3 with L/H successors (e.g., contexts 3-2-1 and 4-2-1 in rows 1–4 of Table 5). 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 six pairs of vertical lines marking the six target locations. Then, 1,500 ms later, participants observed an underline move across three locations (i.e., a context) followed by the appearance of two underlines—one at each of the context's two possible successors. Participants indicated which of the two marked locations was the most likely successor during training given the preceding three underline locations. The sequence of underline movements could be repeated any number of times by pressing the *R* key if participants felt that they needed to see the sequence again prior to making a prediction response. In a sequence of underline movements, underline duration was 410 ms and the interstimulus interval was 400 ms. When the underline disappeared to begin the interstimulus interval, the six location markers were randomly reordered with the constraint that locations 1 versus 6, 2 versus 5, and 3 versus 4 were marked with different pairs of vertical lines. Following a prediction response, the screen was cleared and the prompt to press the space bar to begin the next trial appeared.

The 16 prediction trials were presented in a random order for each participant. Participants performed two practice prediction trials prior to starting the 16 prediction trials. Scores greater than 50% correct (random guessing performance) on the 16 trials would suggest an awareness of the third-order adjacent probabilities.

Procedure

Two participants were randomly assigned to each of the six versions of Table 5. At the beginning of session 1, the perceptual SRTT 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 15, participants performed the prediction task.

Data analysis

The 15 sessions were divided into 5 epochs that each spanned three sessions. For each of the six versions of Table 5, there were 32 contexts with L/H successors. For each participant, the median RT of responses in an epoch (excluding incorrect responses and the first four trials of each block) was determined for each of the 32 L and H successors. The 32 data points for L successors were averaged as were the 32 data points for H successors. The averaged scores were submitted to an ANOVA with successor (L, H) and epoch (1–5) as within-subject factors. Statistical analyses proceeded as in Experiment 1.

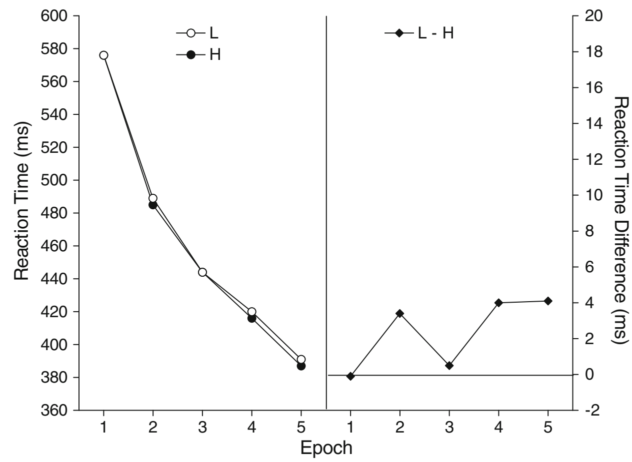

Fig. 3 The *left side* plots reaction time as a function of successor (low probability *L*, high probability *H*) and epoch (1–5) in Experiment 2. The *right side* plots the reaction time difference between L and H successors as a function of epoch. Each epoch spanned three sessions

Results and discussion

The results appear in Fig. 3. There was an effect of epoch Lin, $F(1, 11) = 141.18$, MSE = 3,333.89, $p < 0.001$, indicating that overall RT declined across epochs. There was an effect of successor, $F(1, 11) = 5.09$, MSE = 33.47, $p = 0.023$, 90% CI = (0.5, 2.4, 4.3 ms). The Successor × Epoch Lin interaction was not significant, $F(1, 11) = 1.47$, MSE = 33.05, $p = 0.251$, 95% CI = (−0.7, 0.9, 2.5 ms/epoch). Thus, RT was shorter on H than on L successors and the difference did not change significantly across epochs. The former result indicates that participants learned the third-order adjacent probabilities. There were no significant effects involving error rate.

The percentage of the 16 trials on the prediction task receiving a correct response was determined for each participant. The mean percentage was 51.6% and did not differ significantly from what would be expected by random guessing (50%), $F(1, 11) = 0.14$, MSE = 206.85, $p = 0.714$, 95% CI = (42.4, 51.6, 60.7%). Thus, there was no evidence for awareness of the third-order adjacent probabilities.

Experiment 3

Experiment 3 used the TML paradigm and examined learning of fourth-order probabilities. The previous experiment examined learning of third-order adjacent probabilities and the learning effect (i.e., RT difference between L and H successors) was quite small. This suggests that learning of fourth-order adjacent probabilities could be difficult to detect. Consequently, adjacent and nonadjacent probabilities were confounded in Experiment

3. Remillard (2008a) showed that the learning effect in the traditional SRTT is greater when adjacent and nonadjacent probabilities are confounded than when adjacent probabilities vary and nonadjacent probabilities remain constant. Successor probabilities were also widened from 0.20 versus 0.80 in Experiment 2 to 0.15 versus 0.85 in Experiment 3, and training was extended from 15 sessions in Experiment 2 to 18 sessions in Experiment 3. Thus, the goal of Experiment 3 was to determine whether people can learn to use the location of the target on trial $t-4$ to predict its location on trial t in the perceptual SRTT.

Method

Participants

The participants were 12 university undergraduates ranging in age from 18 to 20 years.

Perceptual SRTT

The perceptual SRTT was identical to that in Experiment 1 except as follows. Participants performed the perceptual SRTT over 18 sessions. Each session was composed of 16 blocks of trials with 114 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 18 sessions were completed in 19–27 days.

Structure of the sequences of target locations

Table 5 presents the contexts and the probabilities (in parentheses) with which successors followed contexts. Every 20 occurrences of a context in tier 1 (rows 1–16) was followed 3 times by one successor (low-probability successor, L) and 17 times by the other successor (high-probability successor, H). For example, row 2 indicates that every 20 occurrences of context 6-3-2-1 was followed 17 times by successor 3 and 3 times by successor 4 so that $P(3|6\text{-}3\text{-}2\text{-}1) = 0.85$ and $P(4|6\text{-}3\text{-}2\text{-}1) = 0.15$. Fourth-order probabilities in tier 1 were 0.15 or 0.85. The odd-numbered rows reveal that whenever location 1 was the target location on trial $t-4$, locations 3 and 4 were the target locations on trial t with probabilities 0.15 and 0.85, respectively. Thus, all eight types of fourth-order probability (see Table 1) with location 1 on trial $t-4$ and location 3 (location 4) on trial t had a value of 0.15 (0.85). For example, $P(3|1\text{-}4\text{-}2\text{-}1) = P(3|1\text{-}x\text{-}x\text{-}x) = P(3|1\text{-}4\text{-}x\text{-}x) = P(3|1\text{-}4\text{-}x\text{-}1) = 0.15$ and $P(4|1\text{-}4\text{-}2\text{-}1) = P(4|1\text{-}x\text{-}x\text{-}x) = P(4|1\text{-}4\text{-}x\text{-}x) = P(4|1\text{-}4\text{-}x\text{-}1) = 0.85$. The even-numbered rows reveal that whenever location 6 was the target location on trial $t-4$, locations 3 and 4 were the target locations on trial t with probabilities 0.85 and 0.15, respectively. Thus, all eight types of fourth-order probability with location 6 on trial $t-4$ and location 3 (location 4) on trial t had a value of 0.85 (0.15). For example, $P(3|6\text{-}4\text{-}2\text{-}1) = P(3|6\text{-}x\text{-}x\text{-}x) = P(3|6\text{-}4\text{-}x\text{-}x) = P(3|6\text{-}4\text{-}x\text{-}1) = 0.85$ and $P(4|6\text{-}4\text{-}2\text{-}1) = P(4|6\text{-}x\text{-}x\text{-}x) = P(4|6\text{-}4\text{-}x\text{-}x) = P(4|6\text{-}4\text{-}x\text{-}1) = 0.15$. Importantly, all first- through third-order probabilities in tier 1 were 0.50. For example, rows 5–8 indicate that when locations 5 and 1 occurred on trials $t-2$ and $t-1$, respectively, location 3 on trial t was an H successor for two contexts and an L successor for two contexts so that $P(3|5\text{-}1) = 0.50$. Similarly, rows 1, 2, 5, 6, 9, 10, 13, and 14 indicate that when location 3 occurred on trial $t-3$, location 3 on trial t was an L successor for four contexts and an H successor for four contexts so that $P(3|3\text{-}x\text{-}x) = 0.50$. Finally, locations 3 and 4 were each a target location equally often (i.e., $P[3] = P[4] = 0.17$). Thus, shorter RTs on H than on L successors would indicate that participants learned to use the location of the target on trial $t-4$, 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 8 types of fourth-order probability). Tier 3 (rows 33–48) was structurally similar to tier 1.

Every 20 occurrences of a context in tier 2 (rows 17–32) was followed 10 times by one successor (medium probability successor, M) and 10 by the other successor (medium probability successor, M). For example, row 18 indicates that every 20 occurrences of context 5-1-3-2 was followed 10 times by successor 1 and 10 times by successor 6 so that $P(1|5\text{-}1\text{-}3\text{-}2) = 0.50$ and $P(6|5\text{-}1\text{-}3\text{-}2) = 0.50$. All first- through fourth-order probabilities were 0.50 in tier 2.

Contexts and L(H) successors in rows 33–48 of tier 3 were matched to contexts and H(L) successors in rows 1–16 of tier 1, respectively, with respect to type of run (see Table 4) and transition pattern. For example, context 1-3-2-6 and H successor 4 in row 9 and context 3-2-1-4 and L successor 5 in row 41 are both UU runs and both form a WWBW transition pattern.

For each participant, a 31,684-element sequence of target locations 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 1–114, 111–224, 221–334, and so forth to 31,571–31,684 each constituted a block of 114 trials for a total of 288 blocks (18 sessions $\times$ 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 5 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 0.50.

There were six versions of Table 5. These were created in a manner analogous to that described in Experiment 1.

Structure of the sequences of targets

The sequences of targets, and hence left and right key responses, were unstructured and independent of the sequences of target locations. The sequences of targets were generated in a manner analogous to that described in Experiment 2.

Prediction task

Awareness of the fourth-order probabilities was assessed using a prediction task similar to that in Experiment 2. There were 32 prediction trials corresponding to the 32 contexts with L/H successors (see Table 5). A prediction trial was identical to that in Experiment 2 except that participants observed the underline move across four locations. The 32 prediction trials were presented in a random order for each participant. Participants performed two practice prediction trials prior to starting the 32 prediction trials. Scores greater than 50% correct (random guessing performance) on the 32 trials would suggest an awareness of the fourth-order probabilities.

Procedure

Two participants were randomly assigned to each of the six versions of Table 5. At the beginning of session 1, the perceptual SRTT 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 18, participants performed the prediction task.

Data analysis

The 18 sessions were divided into 6 epochs that each spanned three sessions. For each of the six versions of Table 5, there were 32 contexts with L/H successors. For each participant, the median RT of responses in an epoch (excluding incorrect responses and the first four trials of each block) was determined for each of the 32 L and H successors. The 32 data points for L successors were averaged as were the 32 data points for H successors. The averaged scores were submitted to an ANOVA with successor (L, H) and epoch (1–6) as within-subject factors. Statistical analyses proceeded as in Experiment 1.

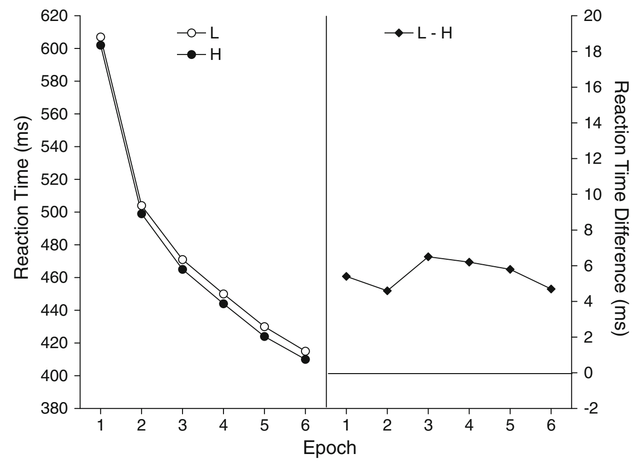

Fig. 4 The *left side* plots reaction time as a function of successor (low probability *L*, high probability *H*) and epoch (1–6) in Experiment 3. The *right side* plots the reaction time difference between L and H successors as a function of epoch. Each epoch spanned three sessions

Results and discussion

The results appear in Fig. 4. There was an effect of epoch Lin, $F(1, 11) = 220.88$, $MSE = 2{,}251.07$, $p < 0.001$, indicating that overall RT declined across epochs. There was an effect of successor, $F(1, 11) = 16.87$, $MSE = 65.71$, $p = 0.001$, 90% CI = (3.1, 5.5, 8.0 ms). The Successor × Epoch Lin interaction was not significant, $F(1, 11) = 0.00$, $MSE = 19.44$, $p = 0.986$, 95% CI = (−0.9, 0.0, 1.0 ms/epoch). Thus, RT was shorter on H than on L successors and the difference did not change significantly across epochs.[6] The former result indicates that participants learned at least 1 of the 8 types of fourth-order probability. There were no significant effects involving error rate.

The percentage of the 32 trials on the prediction task receiving a correct response was determined for each participant. The mean percentage was 48.4% and did not differ significantly from what would be expected by random guessing (50%), $F(1, 11) = 0.42$, $MSE = 70.13$, $p = 0.531$, 95% CI = (43.1, 48.4, 53.8%). Thus, there was no evidence for awareness of the fourth-order probabilities.

I also examined whether performance on the prediction task was correlated with the RT difference between L and H successors. A correlation would not be expected if the RT difference was a product of implicit processes. For this

[6] An effect of successor in the absence of a Successor × Session interaction occurs frequently in the traditional SRTT when there are six target locations and each target location has two possible successors (Remillard, 2008a, 2010; Remillard & Clark, 2001). Thus, this pattern of results is not unusual. Indeed, Rowland and Shanks (2006a) 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 × Block interaction” (p. 291).

analysis, Experiments 2 and 3 were combined because the RT difference between L and H successors in the two experiments was of a similar magnitude and did not change significantly across epochs. Also, the prediction task was similar in the two experiments. The RT difference between L and H successors, averaged across epochs, was calculated for each participant in Experiments 2 and 3 and regressed onto prediction task performance, experiment, and the interaction of prediction task performance and experiment. The regression coefficient for the interaction, $b = -0.10$, was not significantly different from 0, $F(1, 20) = 0.30$, $p = 0.592$. Therefore, the regression analysis was repeated without the interaction term. The regression coefficient for prediction task performance in the second analysis, $b = 0.05$, was not significantly different from 0, $F(1, 21) = 0.49$, $p = 0.492$, 95% CI = (−0.11, 0.05, 0.21 ms/%). Thus, there was no evidence of a correlation between prediction task performance and RT difference between L and H successors when holding experiment constant.

A third regression analysis was performed taking into account the different versions of the sequences of target locations. There were six versions in each experiment for a total of 12 versions. The 12 versions were coded using 11 dummy variables with each variable taking on values of 0 or 1. The RT difference between L and H successors was regressed onto prediction task performance and the 11 dummy variables. There was no need to have experiment as a regressor because experiment was perfectly predictable from version. Again, the regression coefficient for prediction task performance, $b = 0.09$, was not significantly different from 0, $F(1, 11) = 0.76$, $p = 0.403$, 95% CI = (−0.14, 0.09, 0.33 ms/%). Thus, there was no evidence of a correlation between prediction task performance and RT difference between L and H successors when holding version constant.

General discussion

Experiments 1–3 revealed that people can implicitly learn second- and third-order adjacent probabilities in the perceptual SRTT, and also fourth-order probabilities when adjacent and nonadjacent probabilities are confounded. The learning effect (i.e., RT difference between L and H successors) over the latter part of training was 4 ms and a nonsignificant 0.5 ms with third-order adjacent probabilities of 0.20 versus 0.80 and 0.33 versus 0.67, respectively (Experiment 2), and 6 ms with fourth-order probabilities of 0.15 versus 0.85 (Experiment 3). Using the traditional SRTT with sequences of target locations similar to those in the present study, but with narrower successor probabilities, Remillard (2008a) observed a learning effect of 6 ms with third-order adjacent probabilities of 0.40 versus 0.60, and 7 ms with fourth-order probabilities of 0.33 versus 0.67. Thus, the traditional SRTT would likely produce a larger learning effect than the perceptual SRTT given the same sequential contingencies across the two tasks.

The larger learning effect in the traditional SRTT than in the perceptual SRTT could be due to a difference in the expression of learning rather than a difference in learning. In the traditional SRTT, knowledge of the likely target location on the next trial allows one to quickly prepare the corresponding response. Indeed, preparation of a response to an anticipated target location begins 100 ms prior to target onset, and such anticipatory preparation slows preparation of a response to a target that appears at an unanticipated location (Russeler, Hennighausen, & Rosler, 2001). It is also noteworthy that there is a learning effect with a response–stimulus interval of 0 ms (e.g., Rowland & Shanks, 2006a). This suggests that preparation of a response to the likely target location on the next trial can occur while responding to the target location on the current trial. Anticipatory response-preparation is likely slower in the perceptual SRTT because one must orient to the anticipated target location, process the location marker at that location, and then prepare the corresponding response. Thus, greater anticipatory response-preparation in the traditional SRTT than in the perceptual SRTT would lead to a larger learning effect in the former task. Abrahamse, van der Lubbe, and Verwey (2009) note that "sequence effects can not always readily be taken as a clean index for the amount of sequence learning, but rather reflects a combination of the amount of sequence learning and the task-dependent constraints for expressing this knowledge. Therefore, comparing sequence learning across different task-variations should be taken with the necessary caution" (p. 182).

The preceding discussion suggests that a comparison of the learning effect in the traditional SRTT with that in the perceptual SRTT for a given set of sequential contingencies is uninformative with respect to task differences in the ability to learn the contingencies. Therefore, one must depend on the existence of a learning effect rather than the size of a learning effect to make inferences regarding the architectural similarity of the two learning mechanisms. The present experiments show that in the perceptual SRTT, as in the traditional SRTT (Remillard, 2008a; Remillard & Clark, 2001), people can learn second- and third-order adjacent probabilities, and fourth-order probabilities when adjacent and nonadjacent probabilities are confounded. This suggests that the mechanism for learning a sequence of visuospatial locations in the perceptual SRTT may be architecturally similar to the mechanism for learning a sequence of responses in the traditional SRTT.

The idea of two sequence learning mechanisms each with a similar architecture is consistent with neuroanatomy. A number of parallel, corticostriatal loops pass through the basal ganglia (Middleton & Strick, 2001). Each loop begins in a region of the cortex, projects to a region of the striatum (which is the input structure of the basal ganglia and is composed of the putamen and caudate), wends its way through the basal ganglia, and then projects back to the initial cortical region thereby closing the loop. A number of motor loops originate in motor, premotor, and supplementary motor areas of the cortex and project topographically to areas in the putamen. These brain regions, and particularly the putamen, have been implicated in sequence learning in the traditional SRTT (e.g., Badgaiyan, Fischman, & Alpert, 2007; Daselaar, Rombouts, Veltman, Raaijmakers, & Jonkers, 2003; Grafton, Hazeltine, & Ivry, 1995; Peigneux et al., 2000; Rauch et al., 1997; Reiss et al., 2005; Schwarb & Schumacher, 2009). Thus, the motor loops may be part of the mechanism for learning a sequence of responses in the traditional SRTT.

There has been very little research on the neural basis of sequence learning in the perceptual SRTT. The oculomotor loop, which originates in the frontal eye field and projects to the caudate, could, in principle, be part of the mechanism for learning a sequence of visuospatial locations in the perceptual SRTT. The frontal eye field is involved in the overt and covert orienting of visuospatial attention (Awh, Armstrong, & Moore, 2006; Ikkai & Curtis, 2008), and the frontal eye field and caudate are more active when people execute saccades to a visual target that follows a structured sequence of locations relative to a target that follows an unstructured sequence of locations (Gagnon, O'Driscoll, Petrides, & Pike, 2002). There is also evidence that the caudate might be involved in sequence learning in the perceptual SRTT. For example, sequence learning in the perceptual SRTT is unaffected by mild Parkinson's disease, where the caudate is generally unimpaired, but is affected by moderate to severe Parkinson's disease, where the caudate is likely impaired (Price & Shin, 2009). Also, striatal lesions impair sequence learning in the perceptual SRTT (Vakil, Kahan, Huberman, & Osimani, 2000), and the caudate has been implicated in sequence learning in a SRTT that, although not purely perceptual, is mostly visuospatial in nature (Bischoff-Grethe, Martin, Mao, & Berns, 2001). Finally, many of the studies cited in the previous paragraph implicate not only the putamen in sequence learning in the traditional SRTT but also the caudate. Caudate activity likely reflects learning the sequence of visuospatial locations and not the sequence of responses. There is putamen activity, but no caudate activity, related to sequence learning when (a) the visuospatial component is abolished by having participants respond to the identity of centrally presented stimuli with corresponding key presses (Hazeltine, Grafton, & Ivry, 1997), or (b) an attention-capturing visual stimulus is present on each trial in addition to the target (Seidler et al., 2005). The presence of a distractor that captures visuospatial attention is known to impair learning of a sequence of visuospatial locations (Remillard, 2009).

Thus, motor and oculomotor corticostriatal loops may be part of the sequence learning mechanisms in the traditional and perceptual SRTTs, respectively, and the basal ganglia, through which pass the loops, may endow the mechanisms with a common architecture. Ullman (2004, p. 239) has noted that the basal ganglia portion of each corticostriatal loop has a similar synaptic organization and therefore might perform similar computations.

Individuals with damage to Broca's area are impaired at learning a sequence of letters in a modified SRTT, but not at learning a sequence of responses in the traditional SRTT (Goschke, Friederici, Kotz, & van Kampen, 2001). This suggests that the mechanism for learning a sequence of letters is different from that for learning a sequence of responses. Interestingly, the existence of a corticostriatal loop originating in Broca's area has been hypothesized (Ullman, 2006). If such a loop does in fact exist, then the results of Goschke et al. would suggest that the loop may be part of the mechanism for learning a sequence of letters. Additionally, if the sequential contingencies that can be learned in a sequence of letters are similar to those that can be learned in a sequence of responses, this would suggest that the two mechanisms have a similar architecture and support the idea that the basal ganglia are the source of this common architecture.

References

Abrahamse, E. L., van der Lubbe, R. H. J., & Verwey, W. B. (2009). Sensory information in perceptual-motor sequence learning: Visual and/or tactile stimuli. *Experimental Brain Research, 197*, 175–183.

Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: Links, causes and implications for spatial attention. *Trends in Cognitive Sciences, 10*, 124–130.

Badgaiyan, R. D., Fischman, A. J., & Alpert, N. M. (2007). Striatal dopamine release in sequential learning. *Neuroimage, 38*, 549–556.

Belopolsky, A. V., & Theeuwes, J. (2009). When are attention and saccade preparation dissociated? *Psychological Science, 20*, 1340–1347.

Bird, G., Osman, M., Saggerson, A., & Heyes, C. (2005). Sequence learning by action, observation, and action observation. *British Journal of Psychology, 96*, 371–388.

Bischoff-Grethe, A., Goedert, K. M., Willingham, D. T., & Grafton, S. T. (2004). Neural substrates of response-based sequence learning using fMRI. *Journal of Cognitive Neuroscience, 16*, 127–138.

Bischoff-Grethe, A., Martin, M., Mao, H., & Berns, G. S. (2001). The context of uncertainty modulates the subcortical response to predictability. *Journal of Cognitive Neuroscience, 13*, 986–993.

Clegg, B. A. (2005). Stimulus-specific sequence representation in serial reaction time tasks. *The Quarterly Journal of Experimental Psychology, 58*, 1087–1101.

Craighero, L., Nascimben, M., & Fadiga, L. (2004). Eye position affects orienting of visuospatial attention. *Current Biology, 14*, 331–333.

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.

Daselaar, S. M., Rombouts, S. A. R. B., Veltman, D. J., Raaijmakers, J. G. W., & Jonkers, C. (2003). Similar network activated by young and old adults during the acquisition of a motor sequence. *Neurobiology of Aging, 24*, 1013–1019.

de Haan, B., Morgan, P. S., & Rorden, C. (2008). Covert orienting of attention and overt eye movements activate identical brain regions. *Brain Research, 1204*, 102–111.

Deroost, N., Coomans, D., & Soetens, E. (2009). Perceptual load improves the expression but not learning of relevant sequence information. *Experimental Psychology, 56*, 84–91.

Deroost, N., & Soetens, E. (2006a). The role of response selection in sequence learning. *The Quarterly Journal of Experimental Psychology, 59*, 449–456.

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

Deroost, N., & Soetens, E. (2006c). Spatial processing and perceptual sequence learning in SRT tasks. *Experimental Psychology, 53*, 16–30.

Gagnon, D., O'Driscoll, G. A., Petrides, M., & Pike, G. B. (2002). The effect of spatial and temporal information on saccades and neural activity in oculomotor structures. *Brain, 125*, 123–139.

Goschke, T., Friederici, A. D., Kotz, S. A., & van Kampen, A. (2001). Procedural learning in Broca's aphasia: Dissociation between the implicit acquisition of spatio-motor and phoneme sequences. *Journal of Cognitive Neuroscience, 13*, 370–388.

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7*, 497–510.

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor-sequence encoding: A PET study. *Brain, 120*, 123–140.

Helmuth, L. L., Mayr, U., & Daum, I. (2000). Sequence learning in Parkinson's disease: A comparison of spatial-attention and number-response sequences. *Neuropsychologia, 38*, 1443–1451.

Ikkai, A., & Curtis, C. E. (2008). Cortical activity time locked to the shift and maintenance of spatial attention. *Cerebral Cortex, 18*, 1384–1394.

Jimenez, L., & Vazquez, G. A. (2008). Implicit sequence learning in a search task. *The Quarterly Journal of Experimental Psychology, 61*, 1650–1657.

Keele, S. W., Jennings, P., Jones, S., Caulton, D., & Cohen, A. (1995). On the modularity of sequence representation. *Journal of Motor Behavior, 27*, 17–30.

Kelly, S. W., Burton, A. M., Riedel, B., & Lynch, E. (2003). Sequence learning by action and observation: Evidence for separate mechanisms. *British Journal of Psychology, 94*, 355–372.

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, 22*, 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.

Middleton, F. A., & Strick, P. L. (2001). A revised neuroanatomy of frontal-subcortical circuits. In D. G. Lichter & J. L. Cummings (Eds.), *Frontal-subcortical circuits in psychiatric and neurological disorders* (pp. 44–58). New York, NY: Guilford Publications.

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

Peigneux, P., Maquet, P., Meulemans, T., Destrebecqz, A., Laureys, S., et al. (2000). Striatum forever, despite sequence learning variability: A random effect analysis of PET data. *Human Brain Mapping, 10*, 179–194.

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.

Price, A., & Shin, J. C. (2009). The impact of Parkinson's disease on sequence learning: Perceptual pattern learning and executive function. *Brain and Cognition, 69*, 252–261.

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., et al. (1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Human Brain Mapping, 5*, 124–132.

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.

Reiss, J. P., Campbell, D. W., Leslie, W. D., Paulus, M. P., Stroman, P. W., et al. (2005). The role of the striatum in implicit learning: A functional magnetic resonance imaging study. *NeuroReport, 16*, 1291–1295.

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

Remillard, G. (2008a). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *The 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. (2009). Pure perceptual-based sequence learning: A role for visuospatial attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 528–541.

Remillard, G. (2010). Implicit learning of fifth- and sixth-order sequential probabilities. *Memory & Cognition* (in press).

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.

Rowland, L. A., & Shanks, D. R. (2006a). Sequence learning and selection difficulty. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 287–299.

Rowland, L. A., & Shanks, D. R. (2006b). Attention modulates the learning of multiple contingencies. *Psychonomic Bulletin & Review, 13*, 643–648.

Russeler, J., Hennighausen, E., & Rosler, F. (2001). Response anticipation processes in the learning of a sensorimotor sequence. *Journal of Psychophysiology, 15*, 95–105.

Schumacher, E. H., & Schwarb, H. (2009). Parallel response selection disrupts sequence learning under dual-task conditions. *Journal of Experimental Psychology: General, 138*, 270–290.

Schwarb, H., & Schumacher, E. H. (2009). Neural evidence of a role for spatial response selection in the learning of spatial sequences. *Brain Research, 1247*, 114–125.

Seidler, R. D., Purushotham, A., Kim, S. G., Ugurbil, K., Willingham, D., et al. (2005). Neural correlates of encoding and expression in implicit sequence learning. *Experimental Brain Research, 165*, 114–124.

Smith, D. T., Ball, K., Ellison, A., & Schenk, T. (2010). Deficits of reflexive attention induced by abduction of the eye. *Neuropsychologia, 48*, 1269–1276.

Song, S., Howard, J. H., Jr., & Howard, D. V. (2008). Perceptual sequence learning in a serial reaction time task. *Experimental Brain Research, 189*, 145–158.

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.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231–270.

Ullman, M. T. (2006). Is Broca's area part of a basal ganglia thalamocortical circuit? *Cortex, 42*, 480–485.

Vakil, E., Shimon, K., Huberman, M., & Osimani, A. (2000). Motor and non-motor sequence learning in patients with basal ganglia lesions: The case of the serial reaction time (SRT). *Neuropsychologia, 38*, 1–10.

Van der Stigchel, S., & Theeuwes, J. (2007). The relationship between covert and overt attention in endogenous cuing. *Perception & Psychophysics, 69*, 719–731.

Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Memory & Cognition, 27*, 561–572.

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