Pure Perceptual-Based Sequence Learning

Gilbert Remillard
University of Winnipeg

Learning a sequence of target locations when the sequence is uncorrelated with a sequence of responses and target location is not the response dimension (pure perceptual-based sequence learning) was examined. Using probabilistic sequences of target locations, the author shows that such learning can be implicit, is unaffected by distance between target locations, and is mostly limited to first-order transition probabilities. Moreover, the mechanism underlying learning affords processing of information at anticipated target locations and appears to be attention based. Implications for hypotheses of implicit sequence learning are discussed.

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 most cases, the sequence of target locations is deterministic. Sequence learning occurs when the repeating sequence of target locations elicits shorter reaction times (RTs) than does a random sequence of target locations. In other cases, the sequence of target locations is probabilistic. Sequence learning occurs when, given previous target locations, more probable succeeding locations elicit shorter RTs than do less probable succeeding locations.

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

A major criticism of SRT task studies has been that they fail to identify which of a number of sequential constraints have been learned, and therefore one cannot be certain that measures of sequence awareness assessed awareness of the information learned (Jackson & Jackson, 1995; Perruchet, Gallego, & Savy, 1990; Shanks, Green, & Kolodny, 1994; Shanks & St. John, 1994). Recently, Remillard and Clark (2001) addressed this criticism by using highly controlled probabilistic sequences of target locations and showed that people can implicitly learn first-, second-, and third-order transition probabilities.[1]

Another approach to establishing implicit sequence learning is to manipulate the availability of conscious processes. If sequence learning is explicit, then manipulations that affect the availability of conscious processes to the SRT task should also affect sequence learning. Thus, if such manipulations do not affect sequence learning, this would suggest that sequence learning is implicit.

Availability of conscious processes appears to be important for sequence learning when the sequence of target locations is deterministic but not when it is probabilistic. For example, the imposition of a secondary task on the SRT task, which presumably reduces the availability of conscious processes to the SRT task, has (a) no effect on sequence learning when the sequence of target locations is probabilistic (Cleeremans & Jimenez, 1998; Jimenez & Mendez, 1999; Schvaneveldt & Gomez, 1998) and (b) a negative effect when the sequence of target locations is deterministic (Cleeremans & Jimenez, 1998; Frensch, Buchner, & Lin, 1994; Frensch & Miner, 1994; Heuer & Schmidtke, 1996; Hsiao & Reber, 2001; Shanks & Channon, 2002; Stadler, 1995). Similarly, orienting participants to the sequential structure of target locations, which presumably increases the availability of conscious processes to the SRT task, has (a) no effect on sequence learning when the sequence of target locations is probabilistic (Cleeremans & Jimenez, 1998; Jimenez, Mendez, & Cleeremans, 1996; see also D. V. Howard & Howard, 2001) and (b) a positive effect when the sequence of target locations is deterministic (Cleeremans & Jimenez, 1998; Curran & Keele, 1993; Frensch & Miner, 1994). Finally, advance cuing of the next target location, which presumably draws conscious processes toward the cue and away from learning the sequence of target locations, has no effect on sequence learning when the sequence of target locations is probabilistic (Cleeremans, 1997; Jimenez & Mendez, 2001).

Overall, the evidence suggests that learning probabilistic sequences of target locations is implicit, whereas learning deterministic sequences is to some extent explicit (for further support of this conclusion using a digit-sequence entry task, see Marsolek &

Correspondence concerning this article should be addressed to Gilbert Remillard, Department of Psychology, University of Winnipeg, 515 Portage Avenue, Winnipeg, Manitoba R3B 2E9, Canada. E-mail: gremilla@uwinnipeg.ca

[1] An nth-order transition probability, $P(E|A_n\text{-} \ldots \text{-}A_2\text{-}A_1)$, is the probability of an event E occurring on trial t given the occurrence of events $A_n, \ldots, A_2, A_1$ on trials $t - n, \ldots, t - 2, t - 1$, respectively, and is defined as the number of times that E follows the run $A_n\text{-} \ldots \text{-}A_2\text{-}A_1$ divided by the total number of times that $A_n\text{-} \ldots \text{-}A_2\text{-}A_1$ occurs.

Field, 1999). Cleeremans and Jimenez (1998) have proposed that implicit and explicit learning of deterministic sequences can occur simultaneously and involve, respectively, learning of transition probabilities (see also Stadler, 1992; Stadler & Neely, 1997) and learning of serial position information (i.e., the order of target locations). Thus it would seem that probabilistic sequences are better suited than deterministic sequences for studying implicit sequence learning.

Pure Perceptual-Based Sequence Learning

In most SRT task studies, the sequence of target locations is correlated with the sequence of responses because the key corresponding to the location of the target must be pressed. Thus, learning may involve a sequence of target locations (perceptual-based learning), a sequence of response (e.g., key) locations (response-based learning), or a sequence of effector (e.g., finger) movements (effector-based learning). Studies agree that sequence learning is not effector based (Cohen, Ivry, & Keele, 1990; Keele, Jennings, Jones, Caulton, & Cohen, 1995; Stadler, 1989; Willingham, Wells, Farrell, & Stemwedel, 2000) but disagree as to whether sequence learning is primarily response based (Willingham, 1999; Willingham et al., 2000) or is to some extent perceptual based (Keele et al., 1995; Stadler, 1989).

Some studies have examined whether pure perceptual-based learning is possible—that is, whether perceptual-based learning is possible when target location is not the response dimension and the sequence of target locations is uncorrelated with the sequence of responses. For example, Willingham et al. (1989) had participants respond to the color of the target rather than to its location. The sequence of target colors was random, whereas the sequence of target locations repeated. Thus, target location was not the response dimension, and the sequence of target locations was uncorrelated with the sequence of responses. RTs revealed no learning of the sequence of target locations, suggesting that pure perceptual-based learning is not possible.

A number of studies have shown that people can learn a repeating sequence of target locations by observing the sequence and not making any kind of response, suggesting that pure perceptual-based learning is possible (Heyes & Foster, 2002; J. H. Howard, Mutter, & Howard, 1992; Seger 1996, 1997; Willingham, 1999; but see Kelly & Burton, 2001). However, sequence awareness was elevated in those studies, indicating that sequence learning may have been explicit rather than implicit.

Finally, in a conceptual replication of Willingham et al.'s (1989) target color experiment, Mayr (1996; see also Helmuth, Mayr, & Daum, 2000) used widely separated target locations (the four corners of an imaginary square with a side length of 22.6° of visual angle) and obtained evidence for pure perceptual-based learning. Mayr reasoned that the four target locations in the Willingham et al. (1989) study, which were horizontally arranged with adjacent locations separated by 4.7°, may have been too narrowly separated, so that shifts of visuospatial attention or eye movements were not large enough to permit perceptual-based learning to develop or to be detected if it did develop. However, participants in Mayr's study had, on average, significant awareness of the sequence of target locations, suggesting that sequence learning may have been to some extent explicit.

Showing that pure perceptual-based learning can be implicit has important implications for hypotheses of implicit sequence learning. According to some researchers, responding motorically to events in a sequence is critical for learning the sequence of events. For example, Ziessler (1994, 1998; Ziessler & Nattkemper, 2001) argued that response-effect learning is the major component of sequence learning. When exposed to a sequence of events, people learn to associate the response to the current event with the next event in the sequence because the next event is the "effect" of responding to the current event. Hoffman, Sebald, and Stocker (2001) added that people will also learn to associate events in the sequence if distinct events are the "effect" of distinct responses. Nattkemper and Prinz (1997; see also Russeler, Hennighausen, & Rosler, 2001; Russeler & Rosler, 2000) suggested that learning a sequence of events occurs at a motor level, not at a perceptual level. In contrast, Kelly and Burton (2001, Experiment 2) presented data suggesting that learning a sequence of events can occur at a perceptual level provided that events are responded to motorically. Finally, Willingham (1998, 1999; Willingham et al., 2000) proposed that learning a sequence of spatial locations involves the creation of a representation of the sequence in egocentric space and that this is possible only if motor responses are directed to the spatial locations.

Other researchers have suggested that effortful processing of events in a sequence is critical for learning the sequence of events (Baldwin & Kutas, 1997; Hartman, Knopman, & Nissen, 1989; Jimenez & Mendez, 1999; see also Jiang & Chun, 2001). For example, Hartman et al. (1989) observed learning of a repeating sequence of words when the words had to be semantically categorized (an effortful task) but not when they simply had to be read (an automatic task).

Contrary to the preceding hypotheses, showing that pure perceptual-based learning can be implicit would suggest that motoric responding to or effortful processing of events in a sequence is not necessary for implicitly learning the sequence of events. It would show that a sequence of target locations could be learned implicitly when target location is not the response dimension, that is, when target location is not responded to motorically or processed effortfully.

One could argue, however, that pure perceptual-based learning involves oculomotor programming so that target location is responded to motorically. This is difficult to discount because the sudden appearance of a stimulus in the visual field, as occurs in the SRT task, automatically programs an eye movement toward the stimulus (Guitton, Buchtel, & Douglas, 1985; Ladavas, Zeloni, Zaccara, & Gangemi, 1997; Rafal, Calabresi, Brennan, & Sciolto, 1989; Rizzolatti, Riggio, & Sheliga, 1994; Theeuwes, Kramer, Hahn, & Irwin, 1998; Theeuwes, Kramer, Hahn, Irwin, & Zelinksy, 1999; Todd & Van Gelder, 1979). Thus, preventing oculomotor programming in the SRT task is difficult.

However, the attention–oculomotor literature suggests pure perceptual-based learning is more likely to be associated with programming shifts of visuospatial attention than with oculomotor programming. First, eye movements are normally preceded by shifts of attention (Chelazzi et al., 1995; Deubel & Schneider, 1996; Godijn & Pratt, 2002; Hoffman & Subramaniam, 1995; Kowler, Anderson, Dosher, & Blaser, 1995; Posner, 1980; Remington, 1980; Stelmach, Campsall, & Herdman, 1997, Experiment 2). Second, the mechanism for programming shifts of attention is

independent of that for programming eye movements (Abrams & Pratt, 2000; Deubel & Schneider, 1996; Deubel, Schneider, & Paprotta, 1998; Klein & Pontefact, 1994; Ladavas et al., 1997; Posner, 1980; Rafal et al., 1989; Remington, 1980; Reuter-Lorenz & Fendrich, 1992; Shulman, 1984; Stelmach et al., 1997; for a contradictory view, however, see Rizzolatti & Craighero, 1998; Rizzolatti et al., 1994). Finally, the sudden appearance of a stimulus in the visual field, as occurs in the SRT task, automatically captures attention (Folk, Remington, & Wright, 1994; Jonides, 1981; Jonides & Yantis, 1988; McCormick, 1997; Muller & Rabbitt, 1989; Remington, Johnston, & Yantis, 1992). Thus, it is reasonable for one to assume that oculomotor programming is not necessary for pure perceptual-based learning and that automatic orienting of visuospatial attention is sufficient.

The Present Study

The present study had five goals, each designed to characterize pure perceptual-based learning. The first goal was to show that such learning can be implicit. As discussed above, this has important implications for hypotheses of implicit sequence learning. All of the studies that have yielded pure perceptual-based learning have used deterministic sequences of target locations, and most have produced significant awareness of the sequences. This is consistent with the evidence reviewed earlier that learning deterministic sequences may be to some extent explicit. The present study used probabilistic sequences of target locations. The evidence reviewed earlier suggests that learning probabilistic sequences is implicit. Also, probabilistic sequences permit tight control over the constraints that are learned, which is important for accurately assessing awareness of the information learned (Remillard & Clark, 2001). Thus observing pure perceptual-based learning with probabilistic sequences of target locations and no awareness of the information learned would be strong evidence that such learning can be implicit.

The second goal was to show that pure perceptual-based learning is possible when target locations are horizontally arranged and separated by relatively narrow distances. Willingham et al. (1989) used four horizontally arranged target locations with adjacent locations separated by 4.7° of visual angle and found no evidence for pure perceptual-based learning. In a conceptual replication, Mayr (1996) used four target locations forming the corners of an imaginary square with side length of 22.6° and obtained evidence for pure perceptual-based learning. Mayr suggested that the four target locations in the Willingham et al. (1989) study may have been too narrowly separated so that shifts of visuospatial attention or eye movements were not large enough to permit perceptual-based learning to develop or to be detected if it did develop. The present study used displays that were narrower than the 14.2° separation between the leftmost and rightmost target locations in the Willingham et al. (1989) study.

The third goal was to determine whether pure perceptual-based learning is affected by distance between target locations. Across experiments, the width of the display was narrowed. If learning does not differ across experiments, this would suggest that it is unaffected by distance between target locations.

The fourth goal was to determine the complexity of the information that can be learned perceptually. Pure perceptual-based learning of first- (Experiments 1–3) and second-order (Experiment 4) transition probabilities was examined. First-order probability information is less complex than second-order probability information, because in the former only the preceding target location is needed to differentially predict the next target location, whereas in the latter the preceding two target locations are required. Mayr (1996, Experiment 2) observed robust learning of a second-order conditional sequence, suggesting that pure perceptual-based learning of relatively complex information is possible. However, the sequence of target locations used by Mayr was deterministic, and participants had, on average, significant awareness of the sequence. Thus, the complexity of the information that can be learned perceptually and implicitly is currently unknown.

The present study used six horizontally arranged target locations. With the exception of Experiment 3, targets were the bigrams *xo* and *ox*, which required left and right key responses, respectively. Thus, target location was not the response dimension. The sequence of targets, and hence responses, was unstructured and independent of the sequence of target locations, which was probabilistically structured. Given the preceding target locations, there was one high- and one low-probability transition. Shorter RTs on high- than low-probability transitions would indicate pure perceptual-based learning of the transition probabilities.

Because pure perceptual-based learning might be difficult to detect with narrowly separated target locations (Mayr, 1996), sensitivity to learning of the transition probabilities was enhanced by allowing knowledge of the transition probabilities to produce substantial RT benefits and costs on high- and low-probability transitions, respectively.[2] On a trial, each of the six locations was marked with a bigram *xo* or *ox* chosen randomly with the constraint that Locations 1 versus 6, 2 versus 5, and 3 versus 4 were marked with different bigrams (e.g., see Figure 1, left panel, row 2). This ensured that there were three of each bigram and that low- and high-probability transitions were marked with different bigrams. 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 preceding trials, if one location (e.g., Location 3) was the low-probability transition, then its complement (e.g., Location 4) was the high-probability transition. After a 400-ms delay, a line appeared below the bigram marking the next location in the sequence of target locations. Immediately following a response to the underlined target (left key for *xo* and right key for *ox*), the next trial began. Because the bigrams marking each of the

---

[2] The present method for enhancing sensitivity to learning of the transition probabilities was based on an approach used by Goschke (1998, pp. 416–419; Goschke, Friederici, Kotz, & Kampen, 2001) for studying implicit learning of a repeating sequence of auditorily presented letters. On each trial in those studies, the four possible letters were displayed as a random letter string on a monitor (e.g., *CDBA*) and 500 ms later, a target letter was presented auditorily (e.g., "D"). Immediately after the key corresponding to the location of the target letter in the letter string was pressed (e.g., the key for Location 2), the next trial began with another random letter string. RTs decreased with training and increased when the sequence of auditorily presented letters became random. According to Goschke (1998), participants located the anticipated target letter in the letter string during the 500-ms interval and prepared the response corresponding to its location in the string. This produced RT benefits which, when the sequence became random, disappeared and perhaps turned to costs as incorrect responses were prepared.

```
        Present                                                    Absent

ox   xo   xo   ox   ox   xo        until           ox   mn   nm   mn   nm   mn
--                                 response        --

xo   xo   ox   xo   ox   ox        400 ms          mn   mn   mn   nm   nm   nm

xo   xo   ox   xo   ox   ox        until           mn   mn   mn   xo   nm   nm
               --                  response                       --

ox   xo   ox   xo   ox   xo        400 ms          nm   mn   mn   nm   nm   mn

ox   xo   ox   xo   ox   xo        until           nm   xo   mn   nm   nm   mn
     --                            response             --
```

Figure 1. An example sequence of events in the present and absent conditions corresponding to the sequence of target locations 1–4–2 and the sequence of targets *ox–xo–xo*. Location 1 could be followed by Locations 3 or 4, which are marked by different bigrams (row 2), and Location 4 could be followed by Locations 2 or 5, which are marked by different bigrams (row 4).

six locations were chosen pseudorandomly on each trial, the bigrams *xo* and *ox* were equally likely to mark each location.

If participants learn that given the preceding target locations, location A (e.g., location 4) is a more likely transition than location B (e.g., location 3) then they might process, during the 400-ms interval, the bigram marking location A and prepare the corresponding response. This should produce an RT benefit if location A is underlined and a cost if location B, which is marked with a different bigram requiring a different response, is underlined.

Enhanced sensitivity to learning of the transition probabilities is contingent on processing the bigram marking the high-probability transition during the 400-ms interval. This raises an interesting question: Does the mechanism underlying pure perceptual-based learning afford processing of information at an anticipated target location before the cue (i.e., an underline) is presented at the location, or does it afford processing only after the cue is presented but with increased efficiency? For example, if the mechanism is assumed to be a program for successive orientations of visuospatial attention (Posner & Rothbart, 1992), then the question might be whether an attentional shift to the anticipated target location is programmed and executed before the cue or programmed before the cue but executed only after the cue. Mayr (1996) raised a similar question when he noted that "presumably, implicit learning of the spatial sequence either allowed participants to make anticipatory eye [or attentional] movements to correct locations or reduced the threshold for correct eye [or attentional] movements once the object appeared on the screen" (p. 359).

The final goal was to show that the mechanism underlying pure perceptual-based learning affords processing of information at an anticipated target location before the cue is presented. To this end, a second condition was introduced in which location markers were the bigrams *mn* and *nm*. On a trial, each of the six locations was marked with a bigram *mn* or *nm* chosen randomly with the constraint that Locations 1 versus 6, 2 versus 5, and 3 versus 4 were marked with different bigrams (e.g., see Figure 1, right panel, row 2). After a 400-ms delay, the bigram marking the next location in

the sequence of target locations was replaced with an underlined target *xo* or *ox*. Thus, processing of a potential target during the 400-ms interval was not possible in this condition. Immediately following a response to the underlined target, the next trial began.

The conditions in which the location markers were *xo–ox* and *mn–nm* were named the *present* and *absent* conditions, respectively, because the target was either present or absent during the 400-ms interval. The present and absent conditions were identical in all other respects. It is reasonable, therefore, to assume that learning of the transition probabilities would be equivalent in the two conditions. Thus, a larger RT difference between low- and high-probability transitions in the present than in the absent condition would have to be attributed to the former's greater sensitivity to learning because of processing, during the 400-ms interval, of the bigram marking the high-probability transition.

Experiment 1

The SRT task consisted of six target locations, two targets, and two response keys. The left and right keys were pressed in response to the targets *xo* and *ox*, respectively. The sequence of responses was unstructured in that first-order probabilities were .50. For example, if the left key was pressed on trial $t - 1$, then the probabilities of a left and right key response on trial t were each .50. In contrast, the sequence of target locations was structured with first-order probabilities of .33 and .67. For example, if Location 1 was the target location on trial $t - 1$, then the probabilities of Locations 3 and 4 being target locations on trial t might be .33 and .67, respectively. Thus shorter RTs on high- than low-probability transitions would be evidence for pure perceptual-based learning of first-order probabilities. The present and absent conditions were as described earlier.

RTs on low- and high- probability transitions were calculated as a function of type of run completed. Four types of five-element runs were identified in the probabilistic sequences of target locations on the basis of the first and second elements being equal (E)

or unequal (U) to the fourth and fifth elements, respectively (see Table 1). For example, 1–4–2–1–3 is an EU run because the first and fourth elements are equal and the second and fifth elements are unequal. RT to the last element may be shorter for EE runs, in which repetition of a bigram is correctly primed (e.g., 1–3–2–1 primes 3, and 3 occurs), than for UE runs. Likewise, RT to the last element may be longer for EU runs, in which repetition of a bigram is incorrectly primed (e.g., 1–4–2–1 primes 4, but 3 occurs), than for UU runs (Remillard & Clark, 2001). A greater proportion of high- than low-probability transitions in the present study completed the faster EE and UU runs, and therefore type of run completed was a confound. By calculating RT as a function of run and averaging across runs, RTs on low- and high-probability transitions are equally affected by the different runs. Finally, greater RT differences between runs in the present than absent condition would suggest that the priming mechanism affords processing of information at a primed target location before the cue (i.e., an underline) is presented.

Method

Participants. The participants were 24 University of Winnipeg (Winnipeg, Manitoba, Canada) undergraduates ranging in age from 18 to 22 years.

SRT task. The SRT task was run on a personal computer with standard monitor and keyboard. Millisecond timing was implemented using Bovens and Brysbaert's (1990) routine. The six target locations were horizontally arranged and marked with the bigrams *xo* and *ox* in the present condition and *mn* and *nm* in the absent condition. At a viewing distance of approximately 55 cm, each bigram was 0.62° of visual angle in width and 0.42° in height, and the centers of adjacent bigrams were separated by 2.8°. The centers of the bigrams marking the leftmost and rightmost target locations were separated by 14.0° (the width of the display). The red-stickered *V* and *M* response keys, on which participants placed their left and right index fingers, corresponded to the targets *xo* and *ox*, respectively.

There were three sessions, one on each of 3 consecutive days. Each session was composed of 16 blocks of trials with 110 trials per block. The nature of a trial is described in the Experimental Conditions section below. Session 1 began with a practice block of 99 trials.

A performance history was provided at the end of each block. The numbers 1–16 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.

Table 1
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

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

Structure of the sequences of target locations. Letting the Numbers 1–6 represent the six target locations from left to right, respectively, Table 2 presents the third-order probabilities and frequencies that were inherent in the sequences of target locations across every two blocks of trials. For example, row 1 indicates that the sequence 3–2–1 was followed 4 times by Location 3 and 8 times by Location 4; that is, $P(3|3–2–1) = .33$ (low-probability transition; L), and $P(4|3–2–1) = .67$ (high-probability transition; H). Row 9 indicates that the sequence 1–3–2 was followed by 1 four times and by 6four times; that is, $P(1|1–3–2) = .50$, and $P(6|1–3–2) = .50$ (medium-probability transitions; M). Set 2 transitions (L2, H2) immediately followed Set 1 transitions (L1, H1) in the sequences of target locations. Of interest were the first-order probabilities, which were .33, .50, and .67. For example, rows 1–4 indicate that Location 1 was followed 12 times by Location 3 and 24 times by Location 4; that is, $P(3|1) = .33$, and $P(4|1) = .67$.

The sequential structure was controlled so that certain types of information were not confounded with first-order probability. Each location was a target location equally often (i.e., $P[1] = P[2] = \ldots = P[6] = .17$), Lag 3 probabilities[3] were 0.50 (e.g., $P[4|3–x–x] = 0.50$), and Lag 2 probabilities and probabilities of the form $P[E|A_2–A_1–x]$ were 0.44, 0.50, or 0.56 (e.g., $P[3|2–x] = 0.50$, and $P[5|2–1–x] = 0.56$). Thus shorter RTs on H than L transitions would be evidence for learning of the first-order probabilities, although learning of second- or third-order probabilities cannot be ruled out, because these were completely confounded with first-order probability (e.g., $P[3|1] = P[3|2–1] = P[3|3–2–1] = .33$). The confound is addressed in Experiment 4.

For each participant and successive pair of trial blocks, the sequence of target locations was generated by submitting the frequencies in Table 2 to a sequence-generation algorithm that randomly generated a 219-element sequence with the specified frequencies (Remillard & Clark, 1999). Elements 1–110 and 110–219 each constituted a block of 110 trials. For the practice block of 99 trials at the beginning of Session 1, the frequencies in Table 2 were replaced with the number 2. Thus the sequence of target locations in the practice block was unstructured, in that first-, second-, and third-order probabilities were .50.

There were six versions of Table 2. Version 1 was Table 2 itself. Version 2 was formed from Table 2 by exchanging L and H transitions. Version 3 was created by having the top, middle, and bottom thirds of Table 2 describe M, L2–H2, and L1–H1 transitions, respectively. Version 4 was formed from Version 3 by exchanging L and H transitions. Version 5 was created by having the top, middle, and bottom thirds of Table 2 describe L2–H2, L1–H1, and M transitions, respectively. Version 6 was formed from Version 5 by exchanging L and H transitions. The frequencies for each version appear in Appendix A.

Each five-element run was classified into one of four types on the basis of the first and second elements being equal or unequal to the fourth and fifth elements, respectively (see Table 1). A First (E, U) × Last (E, U) interaction with RT to the last element being shorter for EE than UE runs and longer for EU than UU runs would indicate the presence of priming effects.

Structure of the sequences of targets and responses. The sequences of targets, and hence left and right key responses, were unstructured and independent of the sequences of target locations. For each participant and successive pair of trial blocks, the sequence of targets was generated by submitting the frequencies in Table 3 to a sequence-generation algorithm that randomly generated a 227-element sequence with the specified frequencies (Remillard & Clark, 1999). For example, the sequence 1–1–1 was followed 14 times by Target 1 and 14 times by Target 2. Elements 1–110

---

[3] A lag *n* probability, $P(E|A\text{-}x\text{-}\ldots\text{-}x)$, where the number of *x*s is $n − 1$, is the probability of an event *E* occurring on trial *t* given the occurrence of event *A* on trial $t − n$, and is defined as the number of times that *E* occurs *n* trials ahead of *A* divided by the total number of times that *A* occurs.

Table 2
Third-Order Probabilities (and Frequencies) Inherent in the Sequences of Target Locations Across Every Two Blocks of Trials (Experiments 1–3)

Previous target locations	Next target location					
	1	2	3	4	5	6
3–2–1	—	—	L1 (4)	H1 (8)	—	—
4–2–1	—	—	L1 (2)	H1 (4)	—	—
3–5–1	—	—	L1 (2)	H1 (4)	—	—
4–5–1	—	—	L1 (4)	H1 (8)	—	—
3–2–6	—	—	H1 (8)	L1 (4)	—	—
4–2–6	—	—	H1 (4)	L1 (2)	—	—
3–5–6	—	—	H1 (4)	L1 (2)	—	—
4–5–6	—	—	H1 (8)	L1 (4)	—	—
1–3–2	M (4)	—	—	—	—	M (4)
6–3–2	M (8)	—	—	—	—	M (8)
1–4–2	M (4)	—	—	—	—	M (4)
6–4–2	M (2)	—	—	—	—	M (2)
1–3–5	M (2)	—	—	—	—	M (2)
6–3–5	M (4)	—	—	—	—	M (4)
1–4–5	M (8)	—	—	—	—	M (8)
6–4–5	M (4)	—	—	—	—	M (4)
2–1–3	—	H2 (4)	—	—	L2 (2)	—
5–1–3	—	H2 (4)	—	—	L2 (2)	—
2–6–3	—	H2 (8)	—	—	L2 (4)	—
5–6–3	—	H2 (8)	—	—	L2 (4)	—
2–1–4	—	L2 (4)	—	—	H2 (8)	—
5–1–4	—	L2 (4)	—	—	H2 (8)	—
2–6–4	—	L2 (2)	—	—	H2 (4)	—
5–6–4	—	L2 (2)	—	—	H2 (4)	—

Note. Dashes indicate that transitions did not occur. Probabilities: L = .33; M = .50; H = .67. L1 = low-probability transition from Set 1; H1 = high-probability transition from Set 1; M = medium-probability transitions; L2 = low-probability transition from Set 2; H2 = high-probability transition from Set 2.

and 111–220 each constituted a block of 110 trials. Elements 221–227 were excluded, and therefore the frequencies in Table 3 were not exact across the two blocks of trials. Thus, across every two blocks of trials, first-, second-, and third-order probabilities in the sequences of targets were approximately .50. For the practice block of 99 trials at the beginning of Session 1, the frequencies in Table 3 were replaced with the number 6.

Experimental conditions. In the present condition, location markers were the bigrams *xo* and *ox*. On trial *t*, a double dash (—) appeared below a bigram marking one of the locations. Participants pressed the

Table 3
Frequencies for the Sequences of Targets Across Every Two Blocks of Trials (Experiments 1–4)

Previous targets	Next target	
	1	2
1–1–1	14	14
1–1–2	14	14
1–2–1	14	14
1–2–2	14	14
2–1–1	14	14
2–1–2	14	14
2–2–1	14	14
2–2–2	14	14

Note. Targets *xo* and *ox* required left and right key responses, respectively. Target 1 = *xo*; Target 2 = *ox*.

key corresponding to the underlined target. Immediately after a correct response, the line was erased, and the location markers were changed as follows: If the target location on trial *t* + 1 was Location A and *xo* (*ox*) was the target on trial *t* + 1, then bigram *xo* (*ox*) marked Location A. Bigrams 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 bigrams. This ensured that L versus H transitions were marked with different bigrams. After a 400-ms delay, trial *t* + 1 began with the line appearing below the bigram marking Location A (see Figure 1).

In the absent condition, location markers were the bigrams *mn* and *nm*. On trial *t*, a target *xo* or *ox* replaced a bigram *mn* or *nm* marking one of the locations, and a double dash appeared below the target. Participants pressed the key corresponding to the underlined target. Immediately after a correct response, the target and line were erased, and the location markers were changed at random with the constraint that Locations 1 versus 6, 2 versus 5, and 3 versus 4 were marked with different bigrams. After a 400-ms delay, trial *t* + 1 began. It is important to note that during the 400-ms interval, the identity of the marker at any given location was not predictive of the upcoming target.

The present and absent conditions were identical except for the bigrams marking the six locations. Thus it is reasonable to assume that learning of the first-order probabilities would be equivalent in the two conditions. A larger RT difference between L and H transitions in the present than absent condition would thus have to be attributed to the former's greater sensitivity to learning because of processing, during the 400-ms interval, of the bigram marking the high-probability transition and preparation of the corresponding response. Such preparation should produce RT benefits if the line appears below the bigram marking the high-probability transition

or RT costs if it appears below the different bigram marking the low-probability transition.

Assuming that priming effects associated with the different types of runs were equivalent in the present and absent conditions, a larger RT difference between UE and EE runs and between EU and UU runs in the present condition than in the absent condition would have to be attributed to the former's greater sensitivity to priming effects due to processing, during the 400-ms interval, of the bigram marking the primed location (in the case of EE and EU runs) and preparation of the corresponding response. Such preparation should produce RT benefits if the line appears below the bigram marking the primed location (EE runs) or RT costs if it appears below the different bigram marking the unprimed location (EU runs).

Awareness questionnaire. The questionnaire to assess awareness of the first-order probabilities consisted of six items with two options per item. The items were *1→ 3 4, 6 → 3 4, 2 → 1 6, 5 →1 6, 3 → 2 5,* and *4 → 2 5*. For each item, numbers represented target locations, and participants had to choose the high-probability transition. For example, the first item required an indication of whether the double dash, after appearing in Location 1, was more likely to appear in Location 3 or Location 4 next. Four items pertained to L–H transitions and two items pertained to M transitions. Scores greater than 50% correct (random guessing performance) on the four items pertaining to L–H transitions indicated awareness of the first-order probabilities. For participants' reference while completing items, each of the six locations was marked with the bigram *xx* and *nn* in the present and absent conditions, respectively.

Procedure. Two participants were randomly assigned to each of the 12 cells created by crossing condition (present, absent) and version (1–6) of Table 2. At the beginning of Session 1, the SRT task was described, and participants were instructed to try to improve their RT with practice while keeping their error rate below 6%. The structure underlying the sequence of target locations was not mentioned. Immediately following the last block of Session 3, the awareness questionnaire was administered.

Results and Discussion

For each participant, the median RT of correct responses was determined as a function of transition (L, H), set (1, 2), run completed (EE, UE, EU, UU), and session (1, 2, 3). The results appear in the left panel of Figure 2.
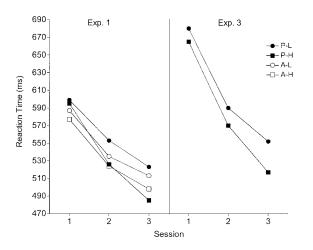

Figure 2. Reaction time, averaged across set (1, 2) and run (EE, UE, EU, UU), as a function of transition (L, H), session (1, 2, 3), condition (present, absent), and experiment (1, 3). There was no absent condition in Experiment 3. P = present condition; A = absent condition; L = low-probability transition; H = high-probability transition; Exp. = Experiment.

Analyses of variance (ANOVAs) with transition (L, H), set (1, 2), first (E, U), last (E, U), and session (1, 2, 3) as within-subject factors were performed on the RT data from the present and absent conditions. When comparing the two conditions, condition (present, absent) and version (1–6) were introduced as between-subjects factors.[4] The session factor was broken into its linear (Session L) and quadratic (Session Q) components. None of the effects involving Session Q approached significance, and therefore only effects involving Session L are reported. Tests for the effect of transition and the Transition × Condition interaction were one-tailed. Shorter RTs on H than L transitions and a greater difference in the present than absent condition were expected. Tests for all other effects were two-tailed. Alpha was .05.[5]

Learning of first-order probabilities. In the present condition, the effect of transition, $F(1, 11) = 5.86$, $MSE = 13,275.09$, $p = .017$, and the Transition × Session L interaction, $F(1, 11) = 16.24$, $MSE = 1,629.83$, $p = .002$, were significant. Thus RT was shorter on H than L transitions, and the difference increased across sessions. This clearly indicates learning of the first-order probabilities.

In the absent condition, the effect of transition was significant, $F(1, 11) = 3.55$, $MSE = 6,138.25$, $p = .043$, and the Transition × Session L interaction was not, $F(1, 11) = 1.44$, $MSE = 539.77$, $p = .256$. Thus RT was shorter on H than L transitions, indicating learning of the first-order probabilities.

When comparing the present and absent conditions, the Transition × Condition interaction approached significance, $F(1, 12) = 3.13$, $MSE = 2,746.65$, $p = .051$, and the Transition × Condition × Session L interaction was significant, $F(1, 12) = 8.94$, $MSE = 1,016.76$, $p = .011$. Thus the RT difference between L and H transitions increased at a faster rate across sessions in the present than absent condition. In Session 3, the Transition × Condition interaction was significant, $F(1, 12) = 9.04$, $MSE = 1,334.04$, $p = .006$. Thus the RT difference between L and H transitions was greater in the present than absent condition. Assuming that learning of the first-order probabilities was equivalent in the present and absent conditions, the results suggest that in the present condition, participants processed, during the 400-ms interval, the bigram marking the high-probability transition.

Priming effects. Averaging across transition (L, H), set (1, 2), and session (1, 2, 3), RTs on EE, UE, EU, and UU runs were, respectively, 547, 553, 548, and 540 ms in the present condition and 540, 540, 539, and 536 ms in the absent condition. In the present condition, the First × Last interaction was significant, $F(1, 11) = 14.41$, $MSE = 478.70$, $p = .003$, reflecting the shorter RTs

---

[4] There was considerable variability in RT differences between L and H transitions across the six versions of Table 2. To remove this variability from the error terms and increase the sensitivity of tests of Transition × Condition interactions, version (1–6) was introduced as a between-subjects factor.

[5] Error rates were also examined. Across experiments, most results did not approach significance (ps > .10). The exceptions generally reflected a pattern of error rates that paralleled the pattern of RTs (e.g., a higher error rate on L than on H transitions, or a larger error rate difference between L and H transitions in the present than in the absent condition). Thus there was no evidence that RT differences between L and H transitions and between runs were due to speed–accuracy tradeoffs.

on EE than UE runs and longer RTs on EU than UU runs. Thus, priming effects were present. The First × Last × Session L interaction was not significant, $F(1, 11) = 2.53$, $MSE = 439.70$, $p = .140$.

In the absent condition, the First × Last interaction, $F(1, 11) = 2.47$, $MSE = 104.28$, $p = .145$, and the First × Last × Session L interaction, $F(1, 11) = 2.68$, $MSE = 235.23$, $p = .130$, were not significant. Thus, there was no evidence for priming effects.

When comparing the present and absent conditions, the First × Last × Condition interaction was significant, $F(1, 12) = 20.92$, $MSE = 107.30$, $p = .001$, reflecting the greater RT difference between UE and EE runs and between EU and UU runs in the present than absent condition. Assuming that priming effects were equivalent in the two conditions, the result suggests that in the present condition participants processed, during the 400-ms interval, the bigram marking the primed target location (in the case of EE and EU runs). The First × Last × Condition × Session L interaction was not significant, $F(1, 12) < 1$.

The above results were generally replicated in the various groups of Experiment 2 as well as in Experiments 3 and 4. Thus, priming effects will not be discussed further, in order to focus on the more important issue of learning-transition probabilities.

Overall RTs. Overall RTs (i.e., RT averaged across L and H transitions) were similar in the present and absent conditions. The effect of condition, $F(1, 12) < 1$, and the Condition × Session L interaction, $F(1, 12) = 2.17$, $MSE = 6473.59$, $p = .166$, were not significant. Thus the larger RT differences between L and H transitions and between runs in the present than absent condition were not an artifact of overall RT differences between conditions (e.g., see Chapman, Chapman, Curran, & Miller, 1994; Curran, 1997, p. 27) nor the result of differences in difficulty (e.g., greater learning of the first-order probabilities or greater priming effects in the present than absent condition because the former was more or less difficult than the latter).

Awareness of first-order probabilities. On the awareness questionnaire, the percentage of the four items pertaining to L–H transitions receiving correct responses (i.e., for which H transitions were chosen) was determined for each participant. In both the present and absent conditions, scores were 56.25%, which did not differ significantly from what would be expected by random guessing (50%), both $Fs(1, 11) < 1$. Thus, there was no evidence for awareness of the first-order probabilities.

Experiment 2

Results from the present condition in Experiment 1 show that (a) pure perceptual-based learning of first-order probabilities is possible when target locations are horizontally arranged and separated by relatively narrow distances and (b) such learning can be implicit. Moreover, the present–absent differences in performance suggest that the mechanism underlying pure perceptual-based learning affords processing of information at an anticipated target location before the cue (i.e., an underline) is presented.

Alternatively, the present–absent differences in performance could have been due to greater learning of the first-order probabilities in the present than in the absent condition rather than to processing bigrams marking anticipated target locations in the present condition. Although this is unlikely, given that the two conditions were identical except for the bigrams marking the six locations and were equally difficult, as indicated by similar overall RTs, it cannot be ruled out. In Experiment 2, participants alternated between the two conditions. I assumed that under those circumstances, knowledge of the first-order probabilities in the present condition would fully transfer to the absent condition. A replication of the present–absent differences in Experiment 1 could then not be attributed to differences in knowledge of the first-order probabilities.

Complete transfer of first-order probability knowledge from the present condition to absent condition is a strong assumption and may be difficult to verify empirically. However, I tested for some transfer by comparing performance in two groups. In the consistent group, L–H transitions in the absent condition were the same as those in the present condition; in the inconsistent group, L–H transitions in the absent condition were opposite those in the present condition. For example, if 1–3 was an L transition in the present condition, then in the absent condition it was an L transition in the consistent group and an H transition in the inconsistent group. Likewise, if 1–3 was an H transition in the present condition, then in the absent condition it was an H transition in the consistent group and an L transition in the inconsistent group.

If there is no transfer of first-order probability knowledge from the present condition to the absent condition, then RT differences between L and H transitions in the absent condition should have been similar in the consistent and inconsistent groups. Conversely, if there was some transfer of knowledge, then RT differences between L and H transitions in the absent condition should have been smaller in the inconsistent group than in the consistent group.

The second purpose of Experiment 2 was to use a shorter distance between adjacent target locations than that used in Experiment 1. The distance between the centers of adjacent bigrams was reduced from 2.8° of visual angle to 1.8°. Thus, the distance between the centers of the bigrams marking the leftmost and rightmost target locations was reduced from 14.0° to 8.8°.

Method

Participants. The participants were 24 University of Winnipeg undergraduates ranging in age from 17 to 34 years.

SRT task. The SRT task was identical to that in Experiment 1 except that adjacent target locations were more narrowly separated (see above), and in each session training alternated between the present and absent conditions. In one order, Blocks 1–2, 5–6, 9–10, and 13–14 were the present condition and Blocks 3–4, 7–8, 11–12, and 15–16 were the absent condition. In the other order, the assignments were reversed. Session 1 began with a practice block of 99 trials performed under the present condition.

Structure of the sequences of target locations. The sequences of target locations were structured and generated as in Experiment 1. L–H transitions in the absent condition were the same as those in the present condition in the consistent group and opposite those in the present condition in the inconsistent group. Specifically, if the sequential structure in the present condition was Version 1, 2, 3, 4, 5, or 6 of Table 2, then the structure in the absent condition was, respectively, Versions 1, 2, 3, 4, 5, and 6 in the consistent group and Versions 2, 1, 4, 3, 6, and 5 in the inconsistent group.

Structure of the sequences of targets and responses. The sequences of targets, and hence left and right key responses, were generated as in Experiment 1.

Experimental conditions. The present and absent conditions were as in Experiment 1.

Awareness questionnaire. The awareness questionnaire was identical to that used in Experiment 1. For each item, participants had to choose the high-probability transition. There was no mention of whether they should refer to the present or absent condition. For participants' reference while completing items, each of the six locations was marked with the bigram *xx*.

Procedure. The consistent and inconsistent groups were each run in a separate experiment with 12 participants per group. Within each group, 1 participant was randomly assigned to each of the 12 cells created by crossing order of present–absent blocks (one, two) and version (1–6) for the present condition. In all other respects, the procedure followed that of Experiment 1.

Results and Discussion

For each participant, the median RT of correct responses was determined as a function of transition (L, H), set (1, 2), run completed (EE, UE, EU, UU), session (1, 2, 3), and condition (present, absent). The results appear in Figure 3.

ANOVAs were as in Experiment 1, except that condition (present, absent) was a within-subject factor. When comparing the consistent and inconsistent groups, group (consistent or inconsistent) and version (1–6) were introduced as between-subjects factors.

Present condition. In the consistent group, the effect of transition, $F(1, 11) = 7.72$, $MSE = 11,476.34$, $p = .009$, and the Transition × Session L interaction, $F(1, 11) = 10.47$, $MSE = 1,757.16$, $p = .008$, were significant. Thus, RT was shorter on H than on L transitions, and the difference increased across sessions. This indicates learning of the first-order probabilities.

In the inconsistent group, the effect of transition was significant, $F(1, 11) = 5.03$, $MSE = 7,076.85$, $p = .023$, and the Transition × Session L interaction was not, $F(1, 11) = 1.72$, $MSE = 3,137.72$, $p = .217$. Thus RT was shorter on H than on L transitions, and the difference did not change significantly across sessions.

When comparing the consistent and inconsistent groups, the Transition × Group interaction, $F(1, 12) = 2.08$, $MSE = 2,850.26$, $p = .175$, and the Transition × Group × Session L interaction, $F(1, 12) = 1.19$, $MSE = 1,623.81$, $p = .296$, were not significant.

Although RT differences between L and H transitions did not differ significantly across the two groups, there is some evidence that performance in the inconsistent group was impaired by the reversal of L and H transitions across the present and absent conditions. The RT difference between L and H transitions did not increase significantly across sessions in the inconsistent group, but it did do so in (a) the consistent group, (b) Experiment 1, (c) two other "consistent" groups to be described in the *Overall RTs (consistent group)* section below, and (d) the next experiment.[6]

Finally, RT differences between L and H transitions in the consistent group were nearly identical to those in the present condition of Experiment 1 even though the consistent group received half the training in the present condition (24 blocks vs. 48 blocks across three sessions). This, together with the evidence for impaired performance in the inconsistent group, suggests that there was some transfer of first-order probability knowledge from the absent condition to the present condition.

Absent condition. In the consistent group, the effect of transition was not significant, $F(1, 11) = 2.37$, $MSE = 5,155.73$, $p = .076$, and the Transition × Session L interaction was significant, $F(1, 11) = 15.55$, $MSE = 641.47$, $p = .002$. Thus, the RT difference between L and H transition increased across sessions. In Session 3, RT was shorter on H than L transitions, $F(1, 11) = 5.19$, $MSE = 3,169.62$, $p = .022$.

In the inconsistent group, both the effect of transition, $F(1, 11) < 1$, and the Transition × Session L interaction, $F(1, 11) = 1.65$, $MSE = 845.56$, $p = .226$, were not significant. Thus, there was no evidence for shorter RTs on H than on L transitions.

When comparing the consistent and inconsistent groups, the Transition × Group interaction was not significant, $F(1, 12) = 2.40$, $MSE = 1,049.68$, $p = .147$, and the Transition × Group × Session L interaction was significant, $F(1, 12) = 19.85$, $MSE = 473.93$, $p = .001$, reflecting the increasing RT difference between L and H transitions across sessions in the consistent but not the inconsistent group. In Session 3, the Transition × Group interaction was significant, $F(1, 12) = 20.02$, $MSE = 432.29$, $p = .001$. Thus, the RT difference between L and H transitions was greater in the consistent than in the inconsistent group. The preceding results strongly suggest that there was some transfer of first-order probability knowledge from the present condition to the absent condition. Otherwise, the pattern of RT differences between L and H transitions would have been similar in the consistent and inconsistent groups.
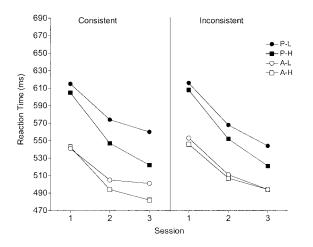

Figure 3. Reaction time, averaged across set (1, 2) and run (EE, UE, EU, UU) as a function of transition (L, H), session (1, 2, 3), condition (present, absent), and group (consistent, inconsistent) in Experiment 2. P = present condition; A = absent condition; L = low-probability transition; H = high-probability transition.

---

[6] Further evidence for impaired performance in the inconsistent group comes from the first block of the present condition after switching from the absent condition. Combining the data from Sessions 2 and 3 to obtain a sufficient number of observations, the median RT of correct responses in the present condition was determined as a function of transition (L, H), set (1, 2), run completed (EE, UE, EU, UU), and block after a switch (first, second) for each participant. The RT difference between L and H transitions was significantly greater in the consistent than in the inconsistent group in the first block (35 vs. 15 ms; $p = .018$), but not in the second block (34 vs. 23 ms; $p = .249$). In the inconsistent group, the RT difference between L and H transitions was significant in the second block ($p = .022$, one-tailed), but not in the first block ($p = .067$, one-tailed). Thus, impairment in the inconsistent group was limited to the first block after switching from the absent condition.

Present versus absent condition (consistent group). In the consistent group, the Transition × Condition interaction was significant, $F(1, 11) = 9.13$, $MSE = 1,914.96$, $p = .006$, and the Transition × Condition × Session L interaction was not, $F(1, 11) = 1.57$, $MSE = 408.04$, $p = .236$. Thus the RT difference between L and H transitions was greater in the present than in the absent condition.

If one assumes that knowledge of the first-order probabilities transferred fully from the present to the absent condition, the preceding result cannot be attributed to greater first-order probability knowledge in the present than absent condition and therefore suggests that in the present condition participants processed, during the 400-ms interval, the bigram marking the high-probability transition. Although the assumption of complete transfer of first-order probability knowledge from the present to the absent condition may be difficult to test empirically, performance differences between the consistent and inconsistent groups in the absent condition suggest that there was some transfer.

Overall RTs (consistent group). In the consistent group, both the effect of condition, $F(1, 11) = 62.08$, $MSE = 16,231.41$, $p < .001$, and the Condition × Session L interaction, $F(1, 11) = 5.61$, $MSE = 2,996.60$, $p = .037$, were significant. Thus, overall RT was longer in the present than in the absent condition. This raises the possibility that the larger RT difference between L and H transitions in the present than absent condition was not the result of processing bigrams marking high-probability transitions in the present condition but rather was an artifact of overall RT differences between conditions (e.g., see Chapman et al., 1994) or the result of differences in difficulty (e.g., greater use of first-order probability knowledge in the present than in the absent condition because the former was more difficult than the latter). To discount this possibility, overall RT difference between the present and absent conditions was manipulated by running two other groups that were identical to the consistent group except that location markers in the absent condition were the bigrams *xx* and *oo* in the difficult group, and short lines above which a target appeared in the easy group.[7]

As expected, overall RT difference between the present and absent conditions varied significantly across the three groups, with the differences being 15, 60, and 101 ms in the difficult, consistent, and easy groups, respectively. In contrast, the pattern of RT differences between L and H transitions in the present and absent conditions did not vary significantly across the three groups.[8] Thus overall RT difference between conditions is likely not responsible for the larger RT difference between L and H transitions in the present than in the absent condition.

Awareness of first-order probabilities. On the awareness questionnaire, the percentage of the four items pertaining to L–H transitions receiving correct responses (i.e., for which H transitions were chosen) was determined for each participant. In the consistent group, the score was 37.5%, which did not differ significantly from what would be expected by random guessing (50%), $F(1, 11) = 3.67$, $MSE = 511.36$, $p = .082$. Thus there was no evidence for awareness of the first-order probabilities. There was actually a tendency for choosing L transitions over H transitions.

In the inconsistent group, scores were 50.0% when the questionnaire was scored with respect to both the present and absent conditions. These values were identical to what would be expected

by random guessing. Again there was no evidence for awareness of the first-order probabilities.

Experiment 3

Session 3 RT differences between L and H transitions were 38 and 15 ms, respectively, in the present and absent conditions of Experiment 1, and 38 and 19 ms, respectively, in the present and absent conditions of Experiment 2 (consistent group). RT differences were similar in the two experiments despite the display being narrower in the latter than former (8.8° vs. 14.0° of visual angle).[9] This suggests that pure perceptual-based learning is unaffected by the distance between target locations. As a stronger test of this hypothesis, Experiment 3 sought to determine whether RT differences between L and H transitions would be similar to those in Experiments 1 and 2 with much more narrowly separated target locations.

In an initial attempt, a study was conducted that was identical to the consistent group of Experiment 2 except that the bigrams *xo*, *ox*, *mn*, and *nm* were replaced with the letters *C*, *O*, *M*, and *N*, respectively, and adjacent target locations were more narrowly separated. Specifically, targets were the letters *C* (left key response) and *O* (right key response), and the six locations were marked with the letters *C* and *O* in the present condition and *M* and *N* in the absent condition. Each letter was 0.31° in width and 0.52° in height, and the centers of adjacent letters were separated by 0.73°. The centers of the letters marking the leftmost and rightmost target locations were separated by 3.6° (the width of the display). Unexpectedly, the study produced no evidence for first-order probability learning.

One explanation is that with the narrowly separated target locations in the present condition, the gaps in a display created by

---

[7] Another explanation for the larger RT difference between L and H transitions in the present than in the absent condition is forward masking of the targets *xo* and *ox* by the bigrams *mn* and *nm* in the absent condition. Forward masking may have slowed responding to targets at anticipated target locations where attention might have been focused. If performance in the easy group, in which there is no forward masking, is similar to that in the consistent and difficult groups, in which forward masking is possible, this would rule out the forward-masking hypothesis.

[8] When examining performance within each group, ANOVAs were as in Experiment 1, except that condition (present, absent) was a within-subject factor. When comparing groups, group (consistent, difficult, easy) and version (1–6) were introduced as between-subjects factors. The Condition × Group interaction was significant, indicating that overall RT difference between the present and absent conditions varied across the three groups. In the present condition, both the transition effect and the Transition × Session L effects were significant in the difficult and easy groups and did not interact significantly with group (both $ps > .280$). In the absent condition, the transition effect was significant in the difficult group and the Transition × Session L effect was significant in the easy group. These effects did not interact significantly with group (both $ps > .176$). Finally, the Transition × Condition interaction was significant in the difficult and easy groups and did not interact significantly with group ($p = .200$).

[9] The Session 3 RT difference between L and H transitions was 16 ms in a pilot study similar to the absent condition of Experiment 1 except that (a) targets were the letters *c* and *o*, (b) location markers were short lines above which the targets appeared, and (c) the display subtended 17.9°. This is very similar to the RT differences in the absent conditions of Experiments 1 and 2 in spite of an even wider display.

C (e.g., *O O C O C C*) may have been salient. Thus, when the letters were reordered immediately after a response, movement of the gaps may have functioned as an abrupt visual change that automatically captured visuospatial attention (e.g., Folk et al., 1994; Jonides, 1981; Jonides & Yantis, 1988; McCormick, 1997; Muller & Rabbitt, 1989; Remington et al., 1992). The capture of attention may have prevented the processing of information at an anticipated target location, thereby eliminating the present condition's sensitivity to first-order probability learning. Alternatively, the capture of attention may have hindered learning by creating a more complex sequence of target locations (or attentional shifts) in which every other element was random.

In a second attempt, stimuli were used that did not create salient features in the display and the absent condition was eliminated so that participants would be constantly exposed to the present condition. It was hoped that these changes would increase the likelihood of observing first-order probability learning. The changes were successful and are described below.

Method

The participants were 12 University of Winnipeg undergraduates ranging in age from 18 to 32 years. The experiment was identical to the present condition in Experiment 1 except that the stimuli were no longer the bigrams *xo* and *ox*, and the distance between adjacent target locations was much more narrow. The bigram *xo* was replaced by two vertical lines, each 0.47° in height and 0.05° in width, with the two lines separated by 0.16°. The left line had a 0.05° gap halfway up. For the bigram *ox*, only the right line had a gap. The centers of adjacent pairs of vertical lines were separated by 0.78°. Thus, the centers of the pairs of vertical lines marking the leftmost and rightmost target locations were separated by 3.9°. For participants' reference while they completed the awareness questionnaire, each of the six locations was marked with a pair of solid (i.e., gap-free) vertical lines.

Results and Discussion

For each participant, the median RT of correct responses was determined as a function of transition (L, H), set (1, 2), run completed (EE, UE, EU, UU), and session (1, 2, 3). The results appear in the right panel of Figure 2.

ANOVAs were as in Experiment 1, except that there was no absent condition. When comparing the present conditions of Experiments 1, 2 (consistent group), and 3, experiment (1, 2, or 3) and version (1–6) were introduced as between-subjects factors.

Learning of first-order probabilities. The effect of transition, $F(1, 11) = 4.65$, $MSE = 16,964.34$, $p = .027$, and the Transition × Session L interaction, $F(1, 11) = 7.47$, $MSE = 1,267.57$, $p = .019$, were significant. Thus, RT was shorter on H than on L transitions, and the difference increased across sessions. This indicates learning of the first-order probabilities.

RT differences between L and H transitions in Experiment 3 were similar to those in Experiments 1 and 2. The Transition × Experiment and Transition × Session L × Experiment interactions were not significant, both $Fs(2, 18) < 1$. The Session 3 RT difference between L and H transitions in Experiment 3 was 35 ms, which is almost identical to the 38 ms in Experiments 1 and 2. Thus, it appears that pure perceptual-based learning is unaffected by distance between target locations, at least within the range of distances used in the present study.

One possible concern is that overall RTs appeared to be longer in Experiment 3 than in Experiments 1 and 2. However, this was the case only in Session 1, in which the effect of experiment was significant, $F(2, 18) = 7.08$, $MSE = 5,571.64$, $p = .005$. The effect of experiment was not significant in Sessions 2 and 3 (both $ps > .133$). Thus, it is unlikely that relative to Experiments 1 and 2, RT differences between L and H transitions in Experiment 3 were disproportionately influenced by overall RT (e.g., Chapman et al., 1994) or task difficulty. Moreover, results from Experiment 2 (see *Overall RTs (consistent group)* section) suggest that overall RT has little influence on RT differences between L and H transitions.

Awareness of first-order probabilities. On the awareness questionnaire, the percentage of the four items pertaining to L–H transitions receiving correct responses (i.e., for which H transitions were chosen) was determined for each participant. The score of 45.8% did not differ significantly from what would be expected by random guessing (50%), $F(1, 11) < 1$. Thus, there was no evidence for awareness of the first-order probabilities.

Performance on the awareness questionnaire in Experiments 1–3 did not differ significantly from 50%. This suggests participants were guessing and hence unaware of first-order probabilities. As a further test of this assertion, participants scoring lower on the questionnaire were compared to those scoring higher. If participants were guessing, then RT differences between L and H transitions should not differ between low and high scorers. To obtain a sufficient number of participants, Experiments 1–3 were combined, resulting in 36 participants (12 from the present condition in Experiment 1, 12 from the consistent group in Experiment 2, and 12 from Experiment 3). In each experiment, a pair of participants had been assigned to each of the six versions of Table 2. For each of the 18 pairs (6 from each experiment), the member scoring lower on the questionnaire was assigned to the low-score group, and the other member was assigned to the high-score group. If each member had the same score, then they were randomly assigned to groups.

The mean score was significantly less than 50% in the low-score group (31.9%), $F(1, 17) = 7.39$, $MSE = 794.54$, $p = .015$, and significantly greater than 50% in the high-score group (61.1%), $F(1, 17) = 5.79$, $MSE = 383.99$, $p = .028$. In spite of the group difference in performance on the awareness questionnaire, RT differences between L and H transitions were nearly identical in the two groups. Averaging across run (EE, UE, EU, UU) and set (1, 2), RT differences between L and H transitions in Sessions 1, 2, and 3 were, respectively, 10, 26, and 37 ms in the low-score group, and 10, 24, and 36 ms in the high-score group. The Transition × Group and Transition × Session L × Group interactions were not significant, both $Fs(1, 24) < 1$ (version [1–6] was a between-subjects factor in these analyses). Thus, RT differences between L and H transitions were not correlated with performance on the awareness questionnaire, suggesting that participants were guessing on the questionnaire and hence unaware of first-order probabilities.

Experiment 4

Experiment 4 examined pure perceptual-based learning of second-order probabilities. The participants were 12 University of Winnipeg undergraduates ranging in age from 17 to 27 years. The experiment was identical to the experience of the consistent group

in Experiment 2, except that there were four sessions of training, and first-order probabilities were exclusively .50. Thus learning of first-order probabilities was not possible. Second- and third-order probabilities were still .33 (L), .50 (M), and .67 (H). Appendix B lists the frequencies (across every two blocks of trials) associated with each of the six versions of the sequential structure of target locations.

The data were analyzed in a manner analogous to that for the consistent group in Experiment 2, except that session had 4 levels. In the absent condition, RT differences between L and H transitions were 1, 3, 1, and −2 ms in Sessions 1–4, respectively. Both the effect of transition and the Transition × Session L interaction were not significant. In the present condition, RT differences between L and H transitions were 5, 2, 10, and 1 ms in sessions 1–4, respectively. The effect of transition was significant ($p = .044$) and the Transition × Session L interaction was not ($F < 1$).

Unfortunately, the pattern of RT differences across sessions in the present condition makes it difficult to interpret the significant effect of transition. The RT difference between L and H transitions was significant in Session 1 ($p = .049$), marginally significant in Session 3 ($p = .055$), and nonsignificant in Sessions 2 and 4 (both $Fs < 1$). Such an inconsistent pattern could reflect weak learning of second-order probabilities, or the RT differences in Sessions 1 and 3 could be an artifact. In any case, the results suggest that pure perceptual-based learning of second-order probabilities is weak at best. The results also make it clear that RT differences between L and H transitions in Experiments 1–3 reflected primarily first-order probability learning and not learning of second- or third-order probabilities, which were confounded with first-order probability.

General Discussion

There were a number of important results in the present study. First, pure perceptual-based learning of first-order probabilities was implicit. This is the first solid demonstration that such learning can be implicit. Second, learning was unaffected by distance between target locations. This suggests eye movements were not necessary for learning. Assuming eye movements were executed less often in Experiment 3 (narrow display) than in Experiments 1 and 2 (wider displays), there should have been less learning in Experiment 3 if eye movements were necessary for learning. However, learning was equivalent across experiments. Further evidence against the necessity of eye movements for learning comes from conceptually similar studies showing that people can implicitly learn the relationship between the form of a cue and the location of a subsequent target when target location is not the response dimension and people are required not to move their eyes (Lambert, Naikar, McLachlan, & Aitken, 1999; Lambert & Sumich, 1996; see also Olson & Chun, 2001, Experiment 3).

Third, the mechanism underlying pure perceptual-based learning of first-order probabilities afforded processing of information at an anticipated target location before the cue (i.e., an underline) was presented. Thus there was orienting to anticipated target locations. Orienting undoubtedly involved shifts of visuospatial attention and may or may not have involved eye movements. As noted in the introduction, shifts of attention normally precede eye movements. Thus, orienting the eyes implies orienting of attention. However, the converse is not necessarily true. It is possible to orient attention while keeping the eyes fixed (e.g., Posner, 1980). The orienting of attention to anticipated target locations indicates that the attention system had knowledge of first-order probabilities. This is consistent with learning being attention based. It could be argued, however, that learning was oculomotor based and that the resulting probability knowledge was accessible to the attention system. This seems unlikely, though, given that the mechanism for programming shifts of attention is independent of that for programming eye movements (see the introduction) and given the suggestion by some that information flows from the attention system to the oculomotor system and not vice versa (Deubel & Schneider, 1996; Deubel et al., 1998; Kowler et al., 1995).

Fourth, there was no evidence for pure perceptual-based learning of first-order probabilities in Experiment 3 when target locations were narrowly separated and location markers in the present condition were the letters *C* and *O*. One explanation is that the salient gaps in a display created by *C* produced an abrupt visual change that automatically captured visuospatial attention when the letters were reordered after a response. The capture of attention then interfered with learning or with the expression of learning. Consistent with this explanation, there was robust learning when the location markers were changed to eliminate salient features in the display. These results are consistent with learning being attention based.

Further evidence that learning in the present study was attention based are the parallels between learning effects (i.e., RT differences between L and H transitions) in the present study and cuing effects in attention-cuing studies. In such studies, visuospatial attention is cued to a location while the eyes remain fixed. RT to a target is shorter when the target appears in the cued (expected) location than in an uncued (unexpected) location—the difference in RT being the cuing effect. Like the learning effect in the present study, the cuing effect is unaffected by distance between target locations (Posner, 1978, pp. 197–203; Remington & Pierce, 1984). Moreover, when responding to the target involves making a two-choice discrimination (as in the present study), the cuing effect is unaffected by overall RT (e.g., Johnston, McCann, & Remington, 1996, Experiment 2; Posner, Snyder, & Davidson, 1980, Experiment 4) and ranges from 15 to 30 ms across studies (e.g., Johnston et al., 1996, Experiment 2; Lupianez & Milliken, 1999, Experiment 2; Posner et al., 1980, Experiment 4; Tipples, 2002). Session 3 learning effects in the absent conditions of the present study fell in this range. Thus, learning effects in the present study behaved like cuing effects, suggesting that learning was attention based.

The fact that pure perceptual-based learning of first-order probabilities was implicit and probably attention based, together with the fact that the sudden appearance of a stimulus in the visual field, as occurs in the SRT task, automatically captures visuospatial attention, suggests that automatic orienting of attention was sufficient for implicit learning. This is contrary to current hypotheses of implicit sequence learning, which suggest that responding motorically or effortfully to events in a sequence is necessary for learning the sequence of events. However, it is possible that responding motorically or effortfully to events in a sequence, although not necessary for learning the sequence of events, may enhance learning.

A fifth important result was that in contrast to robust learning of first-order probabilities, pure perceptual-based learning of second-

order probabilities in Experiment 4 was weak at best. Thus, the mechanism underlying pure perceptual-based learning had a memory that was mostly limited to the preceding target location.

Finally, the absent conditions in Experiments 1 and 2 were conceptually similar to Willingham et al.'s (1989) target color experiment, and pure perceptual-based learning was observed in the former but not in the latter. One explanation is that the target stimuli in the present study were less discriminable than the target colors in the Willingham et al. (1989) study, and therefore eye movements were executed more often in the former. The plausibility of this explanation rests on the assumption that eye movements are necessary for learning, and as argued above, learning is probably attention based and not oculomotor based. A better explanation is that participants in Willingham et al.'s (1989) study had limited practice with the sequence of target locations (40 repetitions of the 10-element sequence), and learning was assessed using a between-subjects measure (RT differences between the repeating sequence group and a random sequence group) instead of a within-subject measure (the change in RT when the repeating sequence becomes random). The former measure sometimes reveals no sequence learning when the latter does (e.g., Keele et al., 1995, Experiment 2; McDowall et al., 1995; Stadler, 1992). Thus, limited practice and a relatively insensitive between-subjects measure of learning may have made it very difficult to obtain evidence of pure perceptual-based learning.

Using a novel procedure (the present condition) and probabilistic sequences of target locations, the present study has shown that pure perceptual-based learning can be implicit and is unaffected by distance between target locations. Moreover, the mechanism underlying learning affords processing of information at an anticipated target location, appears to be attention based, and has a memory mostly limited to the preceding target location. Future research should further characterize the mechanism(s) underlying pure perceptual-based learning.

References

Abrams, R. A., & Pratt, J. (2000). Oculocentric coding of inhibited eye movements to recently attended locations. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 776–788.

Baldwin, K. B., & Kutas, M. (1997). An ERP analysis of implicit structured sequence learning. *Psychophysiology, 34,* 74–86.

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

Chapman, L. J., Chapman, J. P., Curran, T. E., & Miller, M. B. (1994). Do children and the elderly show heightened semantic priming? How to answer the question. *Developmental Review, 14,* 159–185.

Chelazzi, L., Biscaldi, M., Corbetta, M., Peru, A., Tassinari, G., & Berlucchi, G. (1995). Oculomotor activity and visual attention. *Behavioural Brain Research, 71,* 81–88.

Cleeremans, A. (1997). Sequence learning in a dual-stimulus setting. *Psychological Research, 60,* 72–86.

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

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

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30.

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.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827–1837.

Deubel, H., Schneider, W. X., & Paprotta, I. (1998). Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition, 5,* 81–107.

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 317–329.

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, and 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.

Godijn, R., & Pratt, J. (2002). Endogenous saccades are preceded by shifts of visual attention: Evidence from cross-saccadic priming effects. *Acta Psychologica, 110,* 83–102.

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

Goschke, T., Friederici, A. D., Kotz, S. A., & 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.

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research, 58,* 455–472.

Hartman, M., Knopman, D. S., & Nissen, M. J. (1989). Implicit learning of new verbal associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1070–1082.

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.

Heuer, H., & Schmidtke, V. (1996). Secondary-task effects on sequence learning. *Psychological Research, 59,* 119–133.

Heyes, C. M., & Foster, C. L. (2002). Motor learning by observation: Evidence from a serial reaction time task. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55*(A), 593–607.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57,* 787–795.

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

Howard, D. V., & Howard, J. H., Jr. (2001). When it does hurt to try: Adult age differences in the effects of instructions on implicit pattern learning. *Psychonomic Bulletin & Review, 8,* 798–805.

Howard, J. H., Jr., Mutter, S. A., & Howard, D. V. (1992). Serial pattern learning by event observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1029–1039.

Hsiao, A. T., & Reber, A. S. (2001). The dual-task SRT procedure: Fine-tuning the timing. *Psychonomic Bulletin & Review, 8,* 336–342.

Jackson, G. M., & Jackson, S. R. (1995, December). Do measures of explicit learning actually measure what is being learnt in the serial reaction time task? A critique of current methods. *Psyche, 2.* Retrieved August, 29, 2001 from http://psyche.cs.monash.edu.au

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54*(A), 1105–1124.

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

Jimenez, L., & Mendez, C. (2001). Implicit sequence learning with competing explicit cues. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54*(A), 345–369.

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

Johnston, J. C., McCann, R. S., & Remington, R. W. (1996). Selective attention operates at two processing loci. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 439–458). Washington, DC: American Psychological Association.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43,* 346–354.

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, M. A. (2001). Learning complex sequences: No role for observation? *Psychological Research, 65,* 15–23.

Klein, R. M., & Pontefact, A. (1994). Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 333–350). Cambridge, MA: MIT Press.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35,* 1897–1916.

Ladavas, E., Zeloni, G., Zaccara, G., & Gangemi, P. (1997). Eye movements and orienting of attention in patients with visual neglect. *Journal of Cognitive Neuroscience, 9,* 67–74.

Lambert, A., Naikar, N., McLachlan, K., & Aitken, V. (1999). A new component of visual orienting: Implicit effects of peripheral information and subthreshold cues on covert attention. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 321–340.

Lambert, A. J., & Sumich, A. L. (1996). Spatial orienting controlled without awareness: A semantically based implicit learning effect. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 490–518.

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.

Lupianez, J., & Milliken, B. (1999). Inhibition of return and the attentional set for integrating versus differentiating information. *Journal of General Psychology, 126,* 392–418.

Marsolek, C. J., & Field, J. E. (1999). Perceptual-motor sequence learning of general regularities and specific sequences. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 815–836.

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.

McCormick, P. A. (1997). Orienting attention without awareness. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 168–180.

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.

Muller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 315–330.

Nattkemper, D., & Prinz, W. (1997). Stimulus and response anticipation in a serial reaction task. *Psychological Research, 60,* 98–112.

Olson, I. R., & Chun, M. M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1299–1313.

Perruchet, P., Gallego, J., & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology, 22,* 493–516.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Posner, M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 91–111). San Diego, CA: Academic Press.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160–174.

Rafal, R. D., Calabresi, P. A., Brennan, C. W., & Sciolto, T. K. (1989). Saccade preparation inhibits reorienting to recently attended locations. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 673–685.

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

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

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

Remington, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 726–744.

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics, 51,* 279–290.

Remington, R., & Pierce, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception & Psychophysics, 35,* 393–399.

Reuter-Lorenz, P. A., & Fendrich, R. (1992). Oculomotor readiness and covert orienting: Differences between central and peripheral cues. *Perception & Psychophysics, 52,* 336–344.

Rizzolatti, G., & Craighero, L. (1998). Spatial attention: Mechanisms and theories. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in psychological science* (Vol. 2, pp. 171–198). East Sussex, UK: Psychology Press.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 231–265). Cambridge, MA: MIT Press.

Russeler, J., Hennighausen, E., & Rosler, F. (2001). Response anticipation

processes in the learning of a sensorimotor sequence. *Journal of Psychophysiology, 15,* 95–105.

Russeler, J., & Rosler, F. (2000). Implicit and explicit learning of event sequences: Evidence for distinct coding of perceptual and motor representations. *Acta Psychologica, 104,* 45–67.

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

Seger, C. A. (1996). *Implicit learning through observation on the serial reaction time task.* Unpublished manuscript.

Seger, C. A. (1997). Two forms of sequential implicit learning. *Consciousness and Cognition, 6,* 108–131.

Shanks, D. R., & Channon, S. (2002). Effects of a secondary task on "implicit" sequence learning: Learning or performance? *Psychological Research, 66,* 99–109.

Shanks, D. R., Green, R. E. A., & Kolodny, J. A. (1994). A critical examination of the evidence for unconscious (implicit) learning. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing*(pp. 837–860). Cambridge, MA: MIT Press.

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

Shulman, G. L. (1984). An asymmetry in the control of eye movements and shifts of attention. *Acta Psychologica, 55,* 53–69.

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

Stadler, M. A. (1992). Statistical structure and implicit serial learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 318–327.

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.

Stadler, M. A., & Neely, C. B. (1997). Effects of sequence length and structure on implicit serial learning. *Psychological Research, 60,* 14–23.

Stelmach, L. B., Campsall, J. M., & Herdman, C. M. (1997). Attentional and ocular movements. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 823–844.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science, 9,* 379–385.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1595–1608.

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review, 9,* 314–318.

Todd, J. T., & Van Gelder, P. (1979). Implications of a transient-sustained dichotomy for the measurement of human performance. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 625–638.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105,* 558–584.

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

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

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.

Ziessler, M. (1994). The impact of motor responses on serial-pattern learning. *Psychological Research, 57,* 30–41.

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

Ziessler, M., & Nattkemper, D. (2001). Learning of event sequences is based on response-effect learning: Further evidence from a serial reaction task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 595–613.

(*Appendixes follow*)

Appendix A

Frequencies for the Six Versions of Table 2

Previous target locations	Next target location					
	1	2	3	4	5	6
3–2–1	—	—	444444	824428	—	—
4–2–1	—	—	288244	448228	—	—
3–5–1	—	—	284482	444444	—	—
4–5–1	—	—	442882	822844	—	—
3–2–6	—	—	822844	442882	—	—
4–2–6	—	—	444444	284482	—	—
3–5–6	—	—	448228	288244	—	—
4–5–6	—	—	824428	444444	—	—
1–3–2	444444	—	—	—	—	442882
6–3–2	824428	—	—	—	—	822844
1–4–2	448228	—	—	—	—	444444
6–4–2	288244	—	—	—	—	284482
1–3–5	284482	—	—	—	—	288244
6–3–5	444444	—	—	—	—	448228
1–4–5	822844	—	—	—	—	824428
6–4–5	442882	—	—	—	—	444444
2–1–3	—	444444	—	—	288244	—
5–1–3	—	442882	—	—	284482	—
2–6–3	—	822844	—	—	444444	—
5–6–3	—	824428	—	—	448228	—
2–1–4	—	448228	—	—	824428	—
5–1–4	—	444444	—	—	822844	—
2–6–4	—	284482	—	—	442882	—
5–6–4	—	288244	—	—	444444	—

Note. Each digit in the six-digit strings represents the frequencies for the corresponding version (e.g., the first digit represents the frequencies for Version 1). For example, in Version 5, the sequence 3–2–6 was followed four times by Location 3 and eight times by Location 4. Dashes indicate that transitions did not occur.

Appendix B

Frequencies for the Six Versions of the Sequential Structure of Target Locations in
Experiment 4

Previous target locations	Next target location					
	1	2	3	4	5	6
3–2–1	—	—	366344	636328	—	—
4–2–1	—	—	363682	633644	—	—
3–5–1	—	—	633644	363682	—	—
4–5–1	—	—	636328	366344	—	—
3–2–6	—	—	633644	363682	—	—
4–2–6	—	—	636328	366344	—	—
3–5–6	—	—	366344	636328	—	—
4–5–6	—	—	363682	633644	—	—
1–3–2	634436	—	—	—	—	632863
6–3–2	368236	—	—	—	—	364463
1–4–2	364463	—	—	—	—	368236
6–4–2	632863	—	—	—	—	634436
1–3–5	364463	—	—	—	—	368236
6–3–5	632863	—	—	—	—	634436
1–4–5	634436	—	—	—	—	632863
6–4–5	368236	—	—	—	—	364463
2–1–3	—	443663	—	—	286363	—
5–1–3	—	823636	—	—	446336	—
2–6–3	—	446336	—	—	823636	—
5–6–3	—	286363	—	—	443663	—
2–1–4	—	446336	—	—	823636	—
5–1–4	—	286363	—	—	443663	—
2–6–4	—	443663	—	—	286363	—
5–6–4	—	823636	—	—	446336	—

Note. Each digit in the six-digit strings represents the frequencies for the corresponding versions (e.g., the first digit represents the frequencies for Version 1). For example, in Version 4, the sequence 3–2–6 was followed six times by Location 3 and six times by Location 4. Dashes indicate that transitions did not occur.