

Flexible Serial Response Learning by Pigeons (*Columba livia*) and Humans (*Homo sapiens*)

Walter T. Herbranson and George L. Stanton
Whitman College

Experimental tasks designed to involve procedural memory are often rigid and unchanging, despite many reasons to expect that implicit learning processes can be flexible and support considerable variability. A version of the serial response time (SRT) task was developed, in which the locations of targets were probabilistically determined. Targets appeared in locations according to both a structured sequence and a cue validity parameter, and the time to respond to each target was measured. Pigeons (*Columba livia*) and humans (*Homo sapiens*) both showed response time facilitation at the highest tested value for cue validity, and the magnitude of that facilitation gradually weakened as cue validity was decreased. Both species showed evidence that response times were largely determined by the local predictabilities of individual cue locations. In addition, humans showed some evidence that explicit knowledge of the sequence affected response times, specifically when cue validity was 100%.

Keywords: serial response time, procedural memory, pigeon, probability learning

A compelling example of behavioral flexibility that has been the subject of extensive research across multiple species is perception. Many different species can classify objects into useful categories on the basis of their perceptual features, including, but not limited to, humans (Ashby & Maddox, 2005), pigeons (Herrnstein & Loveland, 1964), monkeys (D'Amato & van Sant, 1988), gorillas (Vonk & MacDonald, 2002), orangutans (Vonk & MacDonald, 2004), dogs (Range, Aust, Steurer, & Huber, 2008), and horses (Hanggi, 1999). The importance of flexibility in perception is underscored in category learning by the finding not only that individuals can learn new category structures, but that they can subsequently apply their category knowledge to novel, never-before-seen exemplars. In fact, it is the ability to transfer category knowledge to novel exemplars that often stands as the standard for true category learning (in comparison with memorization of specific, previously encountered exemplars).

A related and similarly flexible cognitive process is memory. Memory ensures that all appropriate responses need not be rigidly preprogrammed into a nervous system. A foraging raccoon can remember the location of a rich food supply that accompanies a new housing development. Similarly, a deer can learn to avoid hunting trails, and even an absent-minded college professor can (occasionally) remember the names of new acquaintances.

The simple characterization of memory just provided, however, leaves out the myriad ways that memory interacts with behavior and allows organisms to actively manipulate their environments.

For example, a bird learns to skillfully alight on chaotically wavering tree branches. A chimpanzee learns to open hard-shelled nuts with a series of well-aimed hammer strikes. And an amateur athlete learns to pocket billiard balls while precisely controlling the speed, trajectory, and final resting position of the cue ball. These are all examples of procedural memory, which is a variety of implicit or nondeclarative long-term memory that mediates the acquisition of skills and habits, and that presumably relies on subcortical neural structures (Squire, 2004).

One laboratory task that has become a standard tool for investigating procedural memory is the serial response time (SRT) task developed by Nissen and Bullemer (1987). In the simplest version of the task, participants respond to four locations on a display by pressing one of four keys having the same relative spatial arrangement. During a structured condition, targets appear in four locations according to a repeating pattern, and response times become faster as participants gain experience. Later, when the target locations are determined randomly, there is a marked slowing of response times. This “global facilitation” effect is simply the difference between the overall median response times during structured and random conditions, and is usually considered to be reflective of procedural memory; when probed, participants often claim to have no awareness that targets appeared according to any consistent pattern even though their response times indicate that they were able to implicitly anticipate and respond more quickly to those target locations. The interpretation that SRT learning involves procedural memory is further supported by physiological experiments showing normal SRT learning in amnesic patients (Nissen & Bullemer, 1987) and in normal participants given the amnesia-inducing anticholinergic drug scopolamine (Nissen, Knopman, & Schacter, 1987). In both cases, participants had no explicit memory of the learning process, but their response times did show a global facilitation effect. Thus, explicit memory does not appear to be necessary for normal SRT learning.

This article was published Online First May 16, 2011.

Walter T. Herbranson and George L. Stanton, Department of Psychology, Whitman College.

This research was supported by a grant from the National Institutes of Health.

Correspondence concerning this article should be addressed to Walter T. Herbranson, Department of Psychology, Whitman College, 345 Boyer Ave., Walla Walla, WA 99362. E-mail: herbrawt@whitman.edu

What is actually learned during an SRT experiment is not yet entirely clear, and must be based on a number of situational factors. Jimenez, Vaquero, and Lupianez (2006) suggested that learned sequences may be implicit, explicit, or both, depending on the learning circumstances. Russeler and Rosler (2000) further proposed that individuals may simultaneously acquire separate motor and perceptual representations of a sequence. According to Remillard and Clark (2001), the implicit representation of a sequence generally consists of knowledge about first-, second-, and third-order transition probabilities. For example, in the sequence originally used by Nissen and Bullemer (1987)—D B C A C B D C B A—Response C is followed by Response A once, and by Response B twice. This yields a first-order conditional probability $p(\text{BIC}) = 2/3$ or .67. Similarly, the first-order probability $p(\text{AIC}) = 1/3$ or .33. Consequently, if a participant has just responded on C, response times are generally faster to B than to A because the first-order conditional probability of BIC is higher than that of AIC. Similarly, second-order probabilities may be determined in the same way, using the two consecutive preceding responses. For the same sequence, the second-order probabilities $p(\text{DICB})$ and $p(\text{AICB})$ are both equal to $1/2$ or .5, because each occurs exactly once in the sequence. An important implication of response times being dependent on local transitional probabilities is that the global response time effect in SRT experiments is not constant. Some parts of the sequence yield consistently faster response times than do others, and contribute disproportionately to the response facilitation seen in comparisons between structured and random conditions.

The SRT methodology has recently been translated into comparative psychology research, with parallel data having been obtained from pigeons (Froehlich, Herbranson, Loper, Wood, & Shimp, 2004) and rats (Christie & Dalrymple-Alford, 2004). These comparative data suggest that procedural memory is a fundamental memory process that is useful in a wide variety of ecological niches, and may be evolutionarily quite old. Because SRT is a performance-based measure of memory, rather than one based on recall or recognition, it is especially well suited to investigations of animal learning. Thus, as with many other performance-based implicit memory tasks (such as rotor pursuit or mirror tracing), the SRT carries the advantage of not being dependent on measures that can be easily misinterpreted.

One important feature that the tasks mentioned above have in common is that they involve repetition and practice—they involve a behavior that is performed in more or less the same way every time. However, such rigidity may not be a necessary characteristic of implicit memory, because many other implicit memory tasks involve considerable flexibility. Weather prediction tasks (Knowlton, Squire, & Gluck, 1994) and artificial grammar learning (Reber, 1967), for example, involve considerable variability, yet also display the defining features of implicit memory. Thus, it may be that the SRT task, which usually involves an unchanging, constantly repeating sequence, could still involve a learning mechanism that accommodates variability when it is present.

Experiment 1

A simple way to add variability to an SRT sequence is to manipulate cue validity. Cue validity is the probability that an individual target will appear in the location dictated by the structured sequence. Using the standard methodology, “structured con-

ditions” are conditions that contain a large proportion of validly cued targets (100% valid targets, to be precise). Meanwhile, “unstructured” or “random” conditions are ones that consist almost entirely of invalidly cued targets. However, note that there are many possible values for cue validity that fall somewhere in between 1.0 and $1/n$, where n is the number of possible responses. In this paper, cue validity may be interpreted in much the same way that the term *validity* has been used in spatial attention research. For example, Posner (1980) cued one of two spatial locations (left or right) with a flashing stimulus. Subsequently, his human participants were faster to respond to targets appearing in the cued location (“validly cued trials”) than were those appearing in the uncued location (“invalidly cued trials”). Shimp and Friedrich (1993) found a similar cue validity effect in pigeons, showing that the importance of validity is not exclusive to humans. The present experiment takes a similar approach to the serial response time task. However, the targets, instead of being cued by a flashing stimulus, are cued by the participants’ own previous responses. A validly cued trial is one in which the target appears in a location that is consistent with the preceding responses, given the structured sequence. An invalidly cued trial is one in which the target appears in a location that is inconsistent with the preceding responses. An appealing feature of this design is that validity can easily be manipulated in a precise and continuous manner, with possible values ranging from 0.0 to 1.0. In contrast, the modal SRT procedure manipulates cue validity in an all-or-none fashion by comparing performances in structured and random conditions. A second advantage is that response times to valid and invalid trials can be compared within the same session, reducing possible ordering effects or historical confounds.

Methods

Animals. Three White Carneaux pigeons (*Columba livia*) were obtained from Double-T Farm (Glenwood, IA). Each was maintained at approximately 80% of its free-feeding weight (Poling, Nickel, & Alling, 1990)¹ in a temperature-controlled colony room with a 14:10 light:dark cycle. Each pigeon was housed in an individual cage with free access to water and grit. All experimental sessions took place at approximately the same time of day, 5 days per week.

Apparatus. Three BRS/LVE operant chambers were used. The front wall of each chamber contained three pecking keys and a feeder through which birds could gain access to mixed grain. Each chamber was interfaced to a computer that controlled all experimental events, recorded data, and computed daily statistics.

Procedure. All birds were pretrained in sessions consisting successively of habituation, magazine training, and autoshaping (Brown & Jenkins, 1968) until consistent responding was achieved on each of the three response keys. Stimuli during autoshaping were white key lights.

A daily experimental session consisted of 1000 trials (gradually increased from 10 over the first 10 days of the experiment). Each trial consisted of a response-to-stimulus (RS) interval of 0.5 s, the

¹ Birds maintained at 80% of free-feeding weight would thus appear to be best described as “naturally motivated,” rather than “deprived.” It is our hope that the results presented here draw on cognitive processes that govern the behavior of birds under typical conditions experienced in the wild.

illumination of one of the three keys with white light (the “target key”), a corresponding peck to the illuminated target key, and (on a random 5% of trials) a reinforcement interval. Illuminated keys remained lit until pecked, with no consequences for pecks to any dark keys. At that point, a single peck either initiated a reinforcement interval of approximately 3 s (reinforcement intervals varied slightly between birds in order to maintain individual running weights), or else passed directly to the RS interval for the next trial. The first trial of a daily session began with the illumination of the target key (without an RS interval), and the rest of the trial progressed as normal.

Three baseline conditions were run at the beginning of the experiment: a structured condition (35 days), followed by a random condition (10 days), and finally a replication of the structured condition (10 days). During both of the structured conditions, illuminated keys appeared on the left (L), center (C), and right (R) keys according to the following fixed nine-item sequence: L C R L R C R L C. First-order local predictability for individual items in the sequence are either .33 (Positions 1, 5, and 6) or .66 (Positions 2, 3, 4, 7, 8, and 9). For example, C follows L two of three times in the sequence ($p = .66$) and R follows L one of three times ($p = .33$). The second- and third-order local predictabilities are all 1.0. The starting point in the sequence was randomly determined by the computer at the beginning of each session. Upon reaching the end of the sequence, it restarted at the beginning and continued to repeat until the required 1000 pecks had been made. The end of one sequence and the beginning of the next were not indicated in any way. The intervening random condition was identical with the exception that the illuminated key on each trial was randomly determined by the computer (with $p = .33$ for each of the three keys).

Cue validity is the likelihood that the target key illuminated on a given trial is the one specified by the fixed nine-item sequence. Thus, the two baseline structured conditions had cue validity of 100% and consisted exclusively of validly cued targets. Following the completion of the second structured condition, we reduced cue validity by 10% every 10 days (i.e., to 90%, then 80%, and so on until reaching the final value of 10%). In the 90% cue validity condition, each target appeared according to the sequence with $p = .9$, and in one of the other two possible locations (randomly chosen) with $p = .1$. Thus, 90% of trials presented validly cued targets, whose locations could be predicted on the basis of the preceding trials. The remaining 10% of trials presented invalidly cued targets, whose locations were inconsistent with the preceding trials. The probabilities applied individually to each and every response during a session. Thus, depending on random variation, each iteration of the sequence might have a larger or smaller percentage of valid trials.

Position within the sequence was always preserved, even when invalid targets were presented. For example, the undistorted sequence L C R L R C R L C might appear as L R R L R C R L C, if the second item in the sequence was invalid. Note that the remaining valid targets have retained their original relative positions.

Results

Pretraining and baseline conditions. Figure 1 displays average² response times across all 3 birds and the final 10 days of each condition for each of the three pretraining conditions. A repeated-measures analysis of variance (ANOVA) with 3 levels

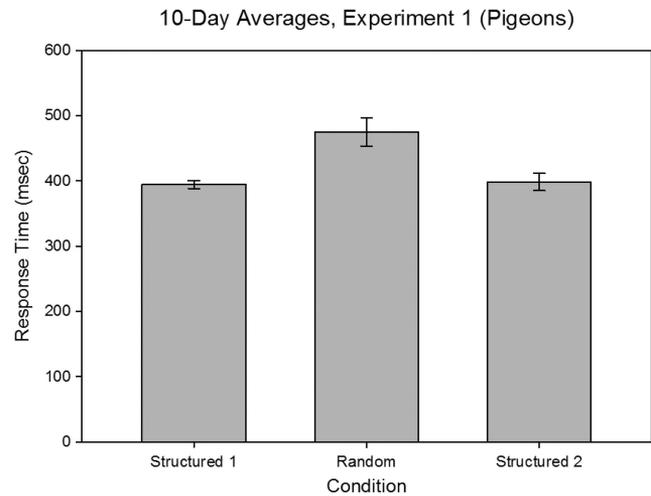


Figure 1. Average response times across all 3 birds and the final 10 days of the three pretraining conditions. Error bars depict 95% confidence intervals.

(Structured 1, Random, Structured 2) shows significant differences among conditions, $F(2, 18) = 55.87, p < .05$. The 95% confidence intervals (CIs) reveal that response times were markedly slower for the middle, unstructured condition ($M = 474.87$ ms) than for either of the structured conditions ($M = 394.22$ ms for Structured 1, and $M = 398.08$ ms for Structured 2). The confidence interval for the difference between the Structured 1 and Random conditions is (CI = 58.03, 103.27, $d = 2.55$). Similarly, the confidence interval for the difference between the Structured 2 and Random conditions is (CI = 54.63, 98.94, $d = 2.48$). Note that neither of these confidence intervals contains a value of 0 (no difference), indicating that the unstructured condition was reliably slower than were both structured conditions. In contrast, the confidence interval for the difference between the two structured conditions was (CI = $-15.48, 7.74, d = 0.24$), indicating no meaningful difference between these two conditions. This replicates the standard global facilitation effect seen in most SRT experiments.

Manipulation of cue validity. Figure 2 shows the overall median response times for all 3 birds in each of the nine conditions having imperfect cue validity, ranging from 90% to 10%. Note that response times became slower as cue validity dropped. Average response times for the 90%, 80%, and 70% conditions were not significantly different from the average response time for the baseline random condition, and were slower than the average response time for both baseline structured conditions (see 95% confidence intervals represented as error bars in Figure 2). The remaining cue validity conditions (60% and lower) were all slower than each of the baseline conditions, whether structured or random.

Validity effect. Figure 3 shows a comparison of response times to valid and invalid targets during each of the imperfect cue validity conditions. A 2 (target type: valid, invalid) \times 9 (cue validity: 90%–10%) repeated-measures ANOVA shows significant main effects for target type, $F(1, 9) = 686.19, p < .05$, and

² Statistics are computed from median response times for each individual session and bird, in order to minimize the effects of response time outliers.

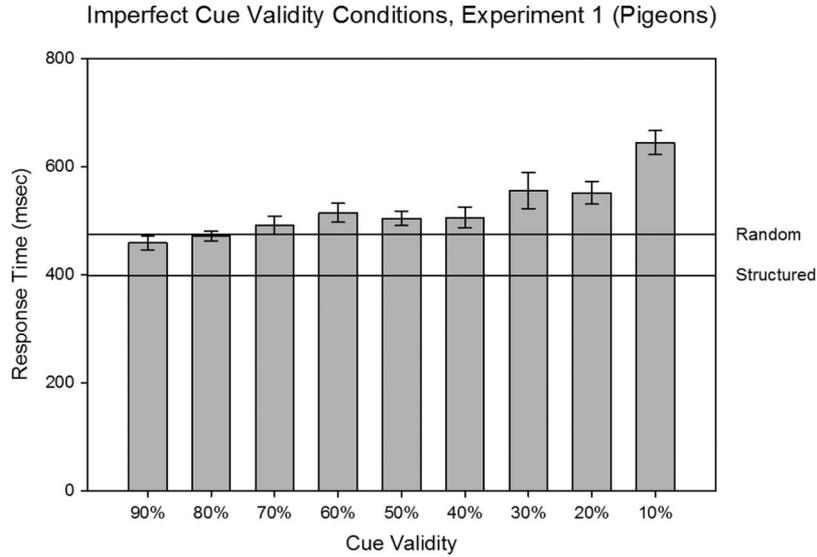


Figure 2. Average response times across all 3 birds and all 10 days of the nine conditions having imperfect cue validity. Error bars depict 95% confidence intervals. Reference lines show the average response times for the final two pretraining conditions (random and structured).

cue validity, $F(8, 72) = 33.80, p < .05$, as well as a significant interaction, $F(8, 72) = 29.40, p < .05$. To further explore the interaction, we compared pairs of valid and invalid means from each cue validity condition (see Table 1). The numbers in Table 1 indicate that the valid and invalid means at 90%, 80%, 70%, and 60% cue validities were reliably different from each other. In contrast, the remaining pairs, at 50% cue validity and below, showed no reliable differences.

Possible sources of the validity effect. There are multiple possible sources of the validity effect shown in Figure 3. One possibility is that pigeons learned the sequence, and were capable of anticipating the location of the next target, but only if they had a reference point indicating the current position in the sequence. Obviously, invalid targets produce a slower mean response time because they appear in an unanticipated location. However, there is a possible secondary effect of those invalid targets: interruption

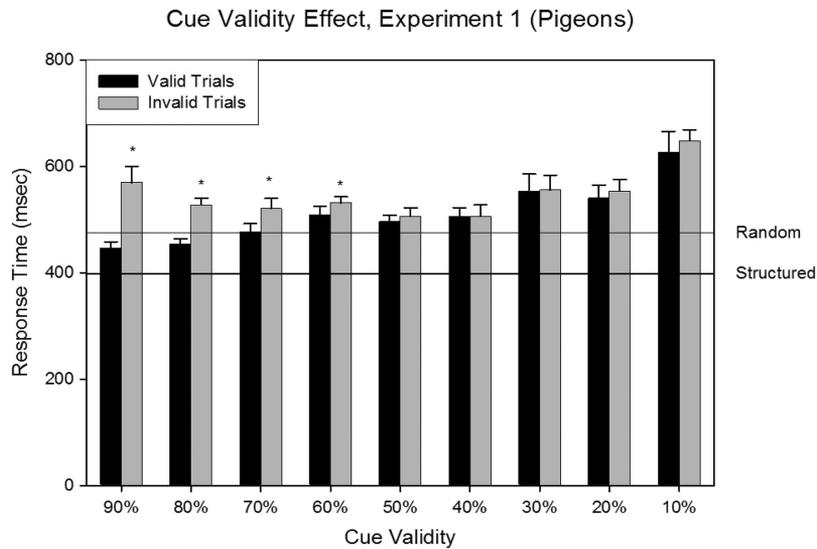


Figure 3. Average response times to valid and invalid trials during each of the nine conditions having imperfect cue validity. Error bars depict 95% confidence intervals. Reference lines show the average response times for the final two pretraining conditions (random and structured). Asterisks indicate statistically reliable differences between valid and invalid trials for a condition.

Table 1
Validity Effect on Response Times Across Conditions, Experiment 1 (Pigeons)

Cue validity	Valid	Invalid	Difference	95% CI	Cohen's <i>d</i>
90%	446.12	569.95	-123.83	-146.70, -100.97	3.87
80%	454.10	527.94	-73.84	-82.95, -64.74	5.81
70%	475.83	520.53	-44.70	-54.22, -35.18	3.36
60%	508.48	530.55	-22.07	-35.67, -8.46	1.16
50%	495.98	506.12	-10.13	-21.04, 0.78	0.66
40%	504.57	506.33	-1.77	-11.15, 7.62	0.13
30%	553.07	555.42	-2.35	-12.95, 8.26	0.16
20%	540.38	553.02	-12.63	-27.72, 2.45	0.60
10%	626.08	648.50	-22.42	-52.98, 8.15	0.52

Note. Rows printed in boldface type indicate conditions in which there was a statistically reliable difference between response times to valid and invalid trials. CI = confidence interval.

of working memory for the current position in the sequence. If this is the case, then the slow response times characteristic of invalid trials may persist into one or more of the subsequent valid trials, at least until the current position in the sequence can be redetermined. Note that there is some preliminary evidence for this interpretation in Figure 3: The mean response times to valid trials in all of the imperfect cue validity conditions were slower than the mean response time during the structured pretraining condition (which consisted entirely of valid trials). By the current interpretation, this could be because the valid means are contaminated by valid targets that could nevertheless not be anticipated because of the disorientation produced by immediately preceding invalid targets. The closer the preceding invalid target, the more likely that this disorientation is to have an effect. In contrast, the more consecutive valid targets that have been seen, the greater the likelihood that a pigeon has identified the current location in the sequence, and will be able to anticipate subsequent targets as it did in the baseline structured condition.

Figure 4 addresses this possibility, and shows response times during the 90% cue validity condition³ as a function of the most recent invalid target. Note that valid targets immediately following invalid targets ($i + 1$) produce response times that closely match the average invalid target (i). In contrast, valid targets that appear more than 1 response after an invalid target ($i + 2$ or more) are very similar to the overall response time for valid targets. The 95% confidence intervals shown in Figure 4 support this interpretation: The confidence intervals for i and $i + 1$ overlap with the average invalid target, but not the average valid target. Meanwhile, the confidence intervals for $i + 2$ or more overlap with the average response time to valid, but not invalid, targets. Note that immediately following an invalid target ($i + 1$), there is an interruption of the context that presumably underlies fast response times. Following one valid target ($i + 2$), there is some context available, in the form of first-order local predictability, but not enough to completely specify the location of the next target. It is only at $i + 3$ that an animal with knowledge of the sequence could predict with perfect accuracy the next valid location. Because there is an abrupt shift between $i + 1$ and $i + 2$, this first explanation seems unlikely. Instead it may be that local predictability (in particular, first-order local predictability) served as the primary mechanism for the response time facilitation seen in this experiment.

The following analyses address a second possible source for the global facilitation effect from Figure 3, namely local predictability.

Figure 5 displays response times to items with first-order local predictabilities of .33 and .66 in each of the five conditions for which there was a cue validity effect (and the final baseline condition). In each condition, response times are faster to targets with the higher local predictability. In addition, response times for a given local predictability were faster if presented in a condition with higher cue validity. Given that local predictability and cue validity are both variables that affect the ability of a bird to predict the location of the next target, the relationship depicted in Figure 5 is sensible, and lends legitimacy to the possibility that birds may not actually be learning the sequence as a whole, but rather are learning the individual transitional probabilities associated with each key.

The slowing of response times and decreasing local facilitation as cue validity decreases seen in Figure 5 make sense when one considers that local predictability and cue validity both influence the likelihood of the next target appearing in a particular location. In other words, they are two ways of manipulating the same thing: predictability. Figure 6 combines these two into "true first-order local predictability," which is equal to the likelihood that the next target will appear in a particular location given the most recent key pecked. Mathematically, this is equal to first-order local predictability multiplied by cue validity (i.e., the probability that a particular key follows in the structured sequence times the probability that the next key will actually be determined by the sequence). When cue validity and first-order local predictability are combined in this way, the data points from Figure 5 align themselves in a strikingly orderly fashion when plotted against response time (with a robust r^2 of .94). Thus, local predictability appears to be a more powerful (and parsimonious) explanation than does the disorientation possibility discussed earlier.

Discussion

The results of Experiment 1 show that pigeons' performance of a learned response sequence is flexible enough to tolerate random noise, even with cue validity as low as 60%. Furthermore, the

³ This analysis and the parallel analysis in Experiment 2 use data from the 90% cue validity conditions because they provided useful quantities of data for each offset distance from invalid trials. Conditions with lower cue validities produced frequencies that were too small to be of use, particularly at longer offsets.

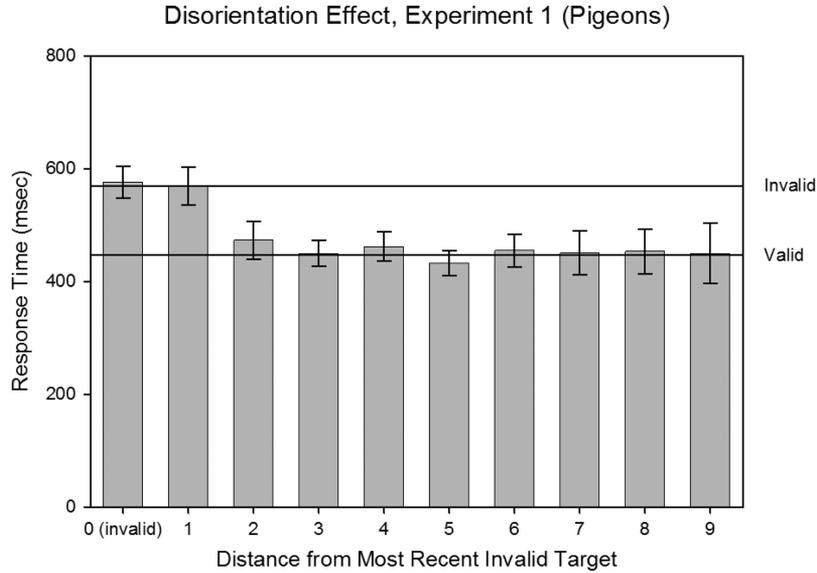


Figure 4. Average response times to valid trials following invalid trials. Data are from the 90% cue validity condition. Error bars depict 95% confidence intervals. Reference lines show the average response time to all valid and all invalid trials from the same condition.

global response time facilitation seems to be based on local predictability—a pigeon’s ability to anticipate the next target on the basis of the preceding response or responses. This kind of flexibility, though not normally a part of SRT research, would undoubtedly be part of most responses acquired in a natural setting (such as those involved in flight, foraging, social interaction, etc.).

One way of interpreting these pigeons’ performance draws on the parallel established earlier with perception. Category learning usually shows acquisition of a central concept, despite considerable variability between the exemplars of the category. For example, pigeons can easily learn to identify and respond appropriately to exemplars of various natural categories such as

bird, fish, or human (Herrnstein & Loveland, 1964). Performance remains high, even when novel exemplars are introduced, and when those exemplars show considerable variability in features. Similarly, in the present experiment, pigeons were initially trained on a central “prototype” sequence, but were

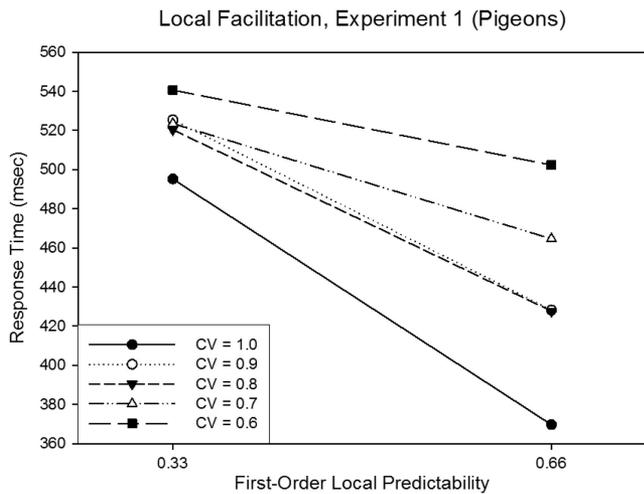


Figure 5. Average response times to valid trials having first-order local predictabilities of .33 (low) and .66 (high) during each of the conditions showing a cue validity effect.

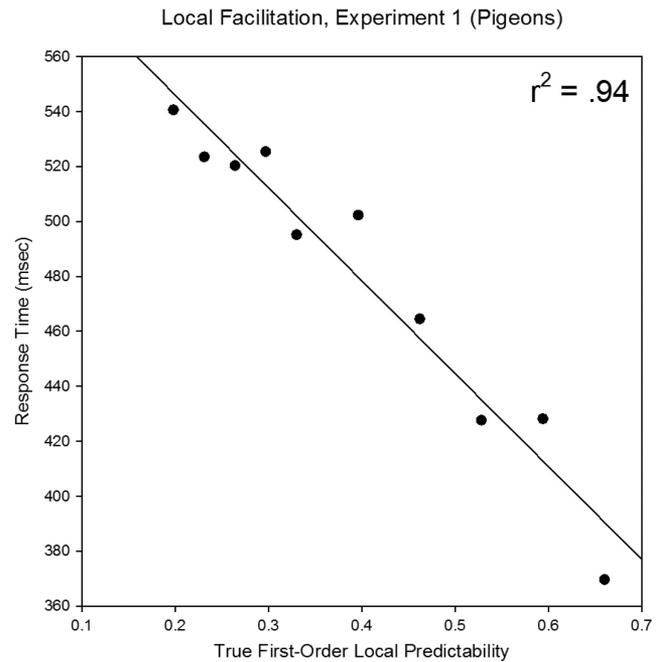


Figure 6. Response times as a function of true first-order local predictability. True first-order local predictability is equal to first-order local predictability times cue validity, and is the probability that the next target will appear in a particular location.

subsequently able to quickly and efficiently deal with variations on that prototype because they still shared a large proportion of its component response transitions.

This experiment then, can be considered as a motor analog of various category learning experiments. Posner and Keele (1968), for example, generated random dot patterns, and presented participants with distortions of those patterns. With experience, participants were able to recognize the distortions as members of the same category, and their ability to do so was dependent on the amount of random distortion applied to the original prototype. These kinds of prototype effects are not exclusive to humans, and also appear in studies of pigeon visual category learning (Aydin & Pearce, 1994). A prototype effect of sorts was obtained in the current experiment, but in this case, the distorted prototype was not a visual pattern, but a motor pattern.

Froehlich et al. (2004) found that pigeons' performance of a similar SRT task with 100% cue validity was based primarily on local probability. The birds in the present experiment, when presented with response sequences based on imperfect cue validity also relied on first-order local probabilities. Thus, it seems that even when flexibility is not required for a task, pigeons nevertheless make use of a cognitive system that can deal with random variability. This finding also sheds some light on what it means to say that a pigeon "knows the sequence." The analyses presented here would imply that pigeons' knowledge of the sequence consists of information about key-to-key transitional probabilities. In other words, a pigeon's experience with the task prepares it to respond quickly to the most likely location, given the key it has just pecked. That is quite different from saying that a pigeon is aware of the sequence in its entirety, or that it executes it as a unit.

On the basis of the above conclusions, we are inclined to think of the SRT as a task that relies primarily on probability learning. Although this is a subtle change from the standard probability learning methodology (e.g., Bullock & Bitterman, 1962), it underscores the flexible nature of how probability can be utilized by animals. For example, Blaisdell, Sewa, Leising, and Waldmann (2006) have also shown that rats use knowledge of probabilities to form the basis of causal reasoning. The present data build on the apparent usefulness of probability learning by showing that animals can not only use probability to predict the likelihood of the possible outcomes of behavior, but can also form the basis of complex behaviors themselves.

Although the SRT methodology has its origins in human cognitive psychology, and the general results are parallel, we do not know whether humans would respond to imperfect cue validity in the same way. Given the short length of the sequence, it may certainly be the case that humans can learn the sequence through an explicit memorization process, rather than by learning transition probabilities. Indeed, some humans report an accurate understanding of the response sequence in postexperimental interviews (Curran & Keele, 1993). Furthermore, humans can be comparatively impaired in relation to pigeons on some kinds of probability-based tasks (e.g., Herbranson & Schroeder, 2010), perhaps because they rely inordinately or inappropriately on explicit strategies, and ignore the actual frequencies of relevant events as they accumulate.

Experiment 2

Comparative psychology emphasizes the importance of cross-species comparisons in understanding behavior. Of Tinbergen's (1963) four causes of behavior (mechanism, ontogeny, adaptation, phylogeny), the majority of recent SRT research seems to emphasize the former two, proximate factors. However, just as important to a complete understanding of a behavioral phenomenon are the latter, ultimate causes. Those ultimate causes can be effectively approached by studying the similar behaviors in different species. This goal is the focus of Experiment 2.

Froehlich et al. (2004) showed that pigeons seem to learn response sequences in much the same way as do humans. Not only did they demonstrate the standard global facilitation effect, but their analysis showed that learning was based largely on local probabilities—the same conclusion drawn by Remillard and Clark (2001) about human SRT learning. On the basis of this comparison, one might expect that both humans and pigeons would perform comparably in an experiment like Experiment 1. Furthermore, similar parallels in the spatial attention and visual categorization experiments that motivated Experiment 1 support this expectation.

Despite these parallels, there are some important differences between humans and pigeons that might call into doubt the likelihood of parallel results. One possible difference lies in the structures of various cognitive systems in different species. For example, Willingham, Salidis, and Gabrieli (2002) outlined parallel neural systems for implicit and explicit skill learning in humans. Although both pigeons and humans quickly and effectively learn skills such as the SRT, the contributions from implicit and explicit memory need not be identical for the two species. For example, humans may rely more heavily on explicit memory, for which their characteristically large frontal lobes are well suited. Ashby, Alfonso-Reese, Turken, and Waldren (1998) have proposed a similar multisystem model of categorization, in which there is direct competition between verbal and implicit systems. Again, one might expect that a verbal system would be considerably more sophisticated in humans, and serve as a possible source of cross-species differences. However, note that although humans' explicit, verbal cognitive abilities are powerful, they may not always lead to an advantage. Herbranson and Schroeder (2010) found that pigeons outperformed humans on a probability puzzler, and concluded that their human participants' failure may have been due to misdirected verbal reasoning. Considering the statistical nature of the SRT, sophisticated verbal abilities might possibly hinder performance, rather than help it.

Methods

Participants. Twelve undergraduates were recruited from introductory level psychology classes. Some participants received course credit for participation.

Apparatus. Stimuli were displayed on a 15-inch flat-panel computer monitor. Three white squares (3 cm × 3 cm) were constantly visible on the monitor, evenly spaced in a horizontal row, to indicate possible target locations. Targets appeared as green squares that replaced the white squares. Participants responded on three adjacent keys (1, 2, and 3) on the number pad of a computer keyboard.

Procedure. Each participant completed 15 blocks with 100 trials per block. Each trial consisted of a response-to-stimulus (RS) interval of 0.5 s, presentation of a target in one of the three marked locations, and a corresponding keypress. Targets remained on the display until the corresponding key was pressed (leftmost key for the leftmost target, center key for the center target, rightmost key for the rightmost target), with no consequences for incorrect keypresses. At the end of each block, participants were allowed to rest for as long as they wished before beginning the next block. The first trial of each block began with the illumination of a target location (without an RS interval), and the rest of the block progressed as normal. Before beginning, participants were shown the display and response keys, and told to respond to each cue as quickly as possible. No mention was made of a possible pattern in the appearance of targets.

During the first five blocks, targets appeared according to the same fixed nine-item sequence that pigeons saw in Experiment 1: L C R L R C R L C. The starting point in the sequence was randomly determined by the computer at the beginning of the experiment. Upon reaching the end of a sequence, the sequence restarted at the beginning and continued to repeat until the required 100 responses for a block had been made. The end of one sequence and the beginning of the next were not indicated in any way.

Beginning with the sixth block, cue validity was reduced by 10% every block (i.e., to 90%, then 80%, and so on until reaching the final value of 0% on Block 15). In the 90% cue validity condition, for example, each target appeared according to the sequence with $p = .9$, and in one of the other two possible locations (randomly chosen) with $p = .1$.

Results

Structure of blocks. The first five bars in Figure 7 show the overall median response times to the first five blocks of training, during which all targets were valid. Note that response times become progressively faster. Blocks 4 and 5 were both significantly faster than was Block 1 (Block 4 CI = 36.25, 131.25, $d = 1.12$; Block 5 CI = 50.99, 137.51, $d = 1.38$), indicating that

participants learned something about the sequence over the first 400 or so trials of the experiment that allowed them to respond more quickly to targets. Although there was some marginal improvement from Block 4 to Block 5, the difference was small and not statistically reliable (CI = -14.90, 35.90, $d = 0.26$), indicating that learning had slowed to an insignificant rate by 500 trials.

Manipulation of cue validity. The final nine bars of Figure 7 show the overall median response times for each of the nine conditions with imperfect cue validity ranging from 90% to 0%, and corresponding to the pigeon data presented in Figure 2. Note that response times quickly become slower as cue validity dropped below 100%.

Validity effect. Figure 8 shows a comparison of response times to valid and invalid targets during each of the imperfect cue validity conditions, and corresponds directly to Figure 3. A 2 (target type: valid, invalid) \times 9 (cue validity: 90%–10%) repeated-measures ANOVA shows a significant main effect for target type, $F(1, 11) = 26.85$, $p < .05$, but not for cue validity, $F(8, 88) = 1.43$, $p > .05$. There was a significant interaction, $F(8, 88) = 13.98$, $p < .05$. To further explore the interaction, pairs of valid and invalid means from each cue validity condition were compared (See Table 2). These numbers indicate that the valid and invalid means at 90%, 80%, 70%, and 60% cue validity were reliably different from each other. In contrast, the remaining pairs, at 50% cue validity and below, were not reliably different. Note that this is precisely the same pattern of results shown by pigeons in Experiment 1.

Possible sources of the validity effect. As in Experiment 1, it is possible to investigate sources of the response time facilitation effect shown in Figure 8, namely, disorientation and local predictability. Figure 9 is the equivalent of Figure 4, and shows response times as a function of the most recent invalid target. The overall pattern is similar to that shown by the pigeons in Experiment 1, in that targets that immediately follow an invalid target are indistinguishable from the invalid targets that precede them (see reference lines), and targets that are removed by three or more responses from an invalid target are indistinguishable from the mean response time to valid targets (again, note that at $i + 3$, the target can be unambiguously predicted by an individual who knows the sequence). However, there is one notable and important difference. In particular, the response times to targets falling 2 responses after an invalid target ($i + 2$) fall at an intermediate value, different from their immediate neighbors ($i + 1$ and $i + 3$) and from the overall means for valid and invalid targets. Note that this implies that there is some factor at work besides first-order local probability, and is consistent with what one might expect if lost position were a factor.

Recall that a second possible interpretation of response time facilitation is based on first-order local probability. Figure 10 corresponds to Figure 5, and shows response times to items with first-order local probabilities of .33 and .66 in Blocks 5–9 (the final training block, and blocks with cue validities of 90%–60%, for which there was a cue validity effect). As was the case for pigeons, response times were faster to targets with higher local predictability. In addition, response times for a given local predictability were faster if presented in a block with higher cue validity. However, note that the lines corresponding to different blocks are not as evenly spaced as are those in Figure 5. In particular, the line corresponding to 100% valid targets is quite distant from the rest.

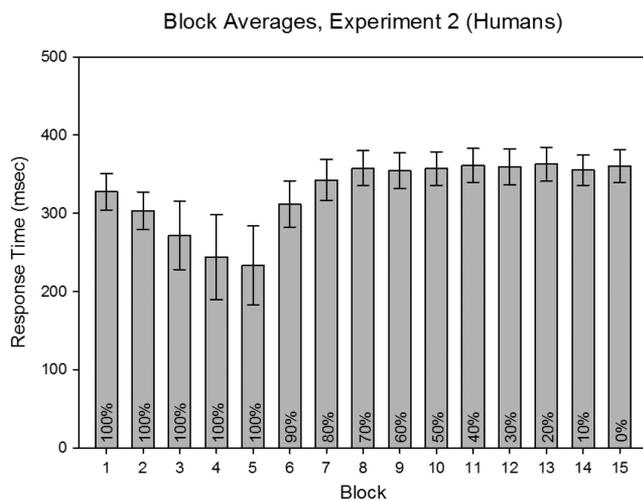


Figure 7. Average response times for all 12 participants in each of the 15 blocks of Experiment 2. Error bars depict 95% confidence intervals.

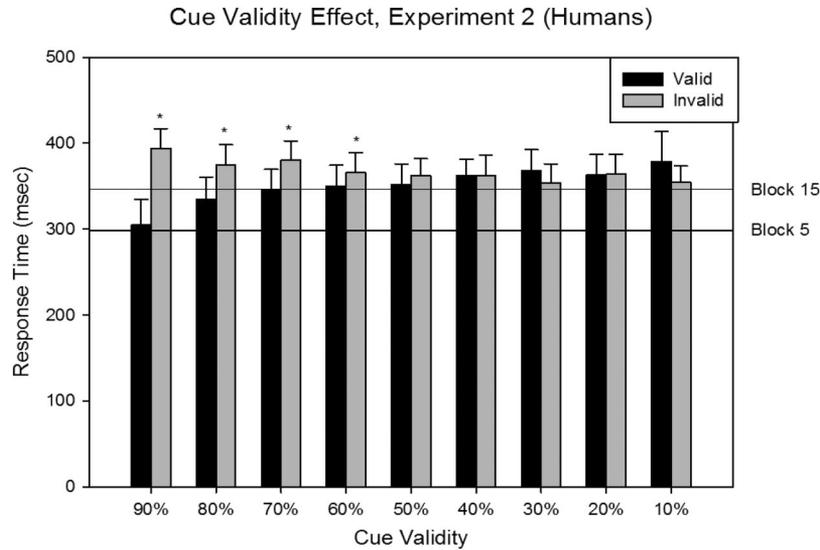


Figure 8. Average response times to valid and invalid trials during each of the nine blocks having imperfect cue validity. Error bars depict 95% confidence intervals. Reference lines show the average response times for blocks having 100% cue validity and 0% cue validity (Blocks 5 and 15, respectively). Asterisks indicate statistically reliable differences between valid and invalid trials for a block.

Figure 11 translates all of these data points into a common measure of true first-order local probability (as in Figure 5, true first-order local probability is equal to cue validity times first-order local probability). Note that there is considerably more variability among data points, and the r^2 is a more modest .46. A close examination of Figure 11 reveals that the two most extreme outliers at the bottom of the figure correspond to the two data points from 100% cue validity block. If, as the previous analysis suggests, human participants made use of explicit knowledge of the sequence, the 100% valid condition might produce a notably different pattern in Figures 10 and 11. Figure 12 removes these two data points, and produces a more robust r^2 of .72. Given that a condition with 100% cue validity is likely the one in which explicit knowledge of the sequence would be most valuable, the increased explanatory power when these two data points are removed makes intuitive sense. Rather than plotting data that rely on two different cognitive processes (implicit knowledge of local predictability and explicit memory of the sequence), the smaller data set draws more

exclusively on a single cognitive process by interfering with the usefulness of explicit memory.

Discussion

Experiment 2 found that humans produced the same kind of flexible SRT performance in conditions with imperfect cue validity that was shown by the pigeons in Experiment 1. In fact, a global facilitation effect was obtained under exactly the same conditions (i.e., from 90% cue validity down to 60%, but no lower). Nevertheless, there were some subtle differences between the two species. In particular, first-order local predictability was not as strong a predictor of response times in humans, especially when cue validity was 100%. Thus, it seems that humans' and pigeons' SRT performances are not in all ways identical. In this case, human participants' knowledge of the sequence transitions was supplemented by something else—possibly some form of explicit learning—but only in the 100% validity condition. In conditions with

Table 2
Validity Effect on Response Times Across Conditions, Experiment 2 (Humans)

Cue validity	Valid	Invalid	Difference	95% CI	Cohen's d
90%	304.63	393.92	-89.29	-126.12, -52.57	1.54
80%	334.54	374.08	-39.54	-61.82, -17.26	1.13
70%	345.79	379.88	-34.08	-46.23, -21.93	1.78
60%	350.13	365.71	-15.58	-30.56, -0.61	0.66
50%	351.50	361.58	-10.08	-22.94, 2.78	0.50
40%	362.46	361.83	0.63	-8.11, 9.36	0.05
30%	367.92	353.46	14.46	-3.57, 25.34	0.84
20%	362.71	364.00	-1.29	-18.34, 15.76	0.05
10%	378.58	354.58	24.00	-4.49, 52.49	0.54

Note. Rows printed in boldface type indicate conditions in which there was a statistically reliable difference between response times to valid and invalid trials. CI = confidence interval.

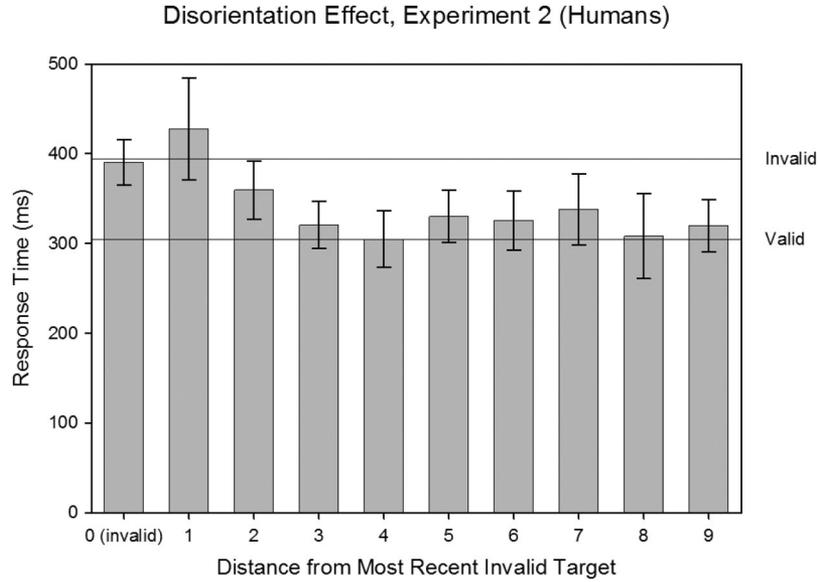


Figure 9. Average response times to valid trials following invalid trials. Data are from the 90% cue validity block. Error bars depict 95% confidence intervals. Reference lines show the average response time to all valid and all invalid trials from the same block.

imperfect cue validity, local predictability did a respectable job of describing both human and pigeon response times.

During postexperimental debriefing, 10 of the 12 participants reported that they noticed a pattern in the target locations. Of those 10, 8 were able to accurately state at least part of the sequence. As the results show, this kind of explicit awareness would be quite useful at 100% cue validity, when the resulting expectations are consistently correct. At imperfect cue validities, however, the same knowledge could be a hindrance to performance, and actually slow response times, because some targets will be invalid and appear in unexpected locations. Note though, that even when imperfect cue validity interfered with participants' ability to apply explicit mem-

ory, local facilitation was still obtained in all blocks having cue validity values greater than or equal to 60%. Thus, we can conclude that along with an explicit understanding of the sequence, participants were also acquiring implicit knowledge about local

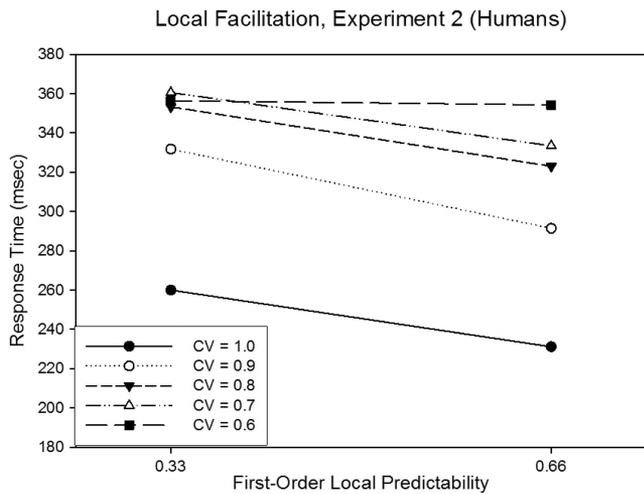


Figure 10. Average response times to valid trials having first-order local predictabilities of .33 (low) and .66 (high) during each of the blocks showing a cue validity effect.

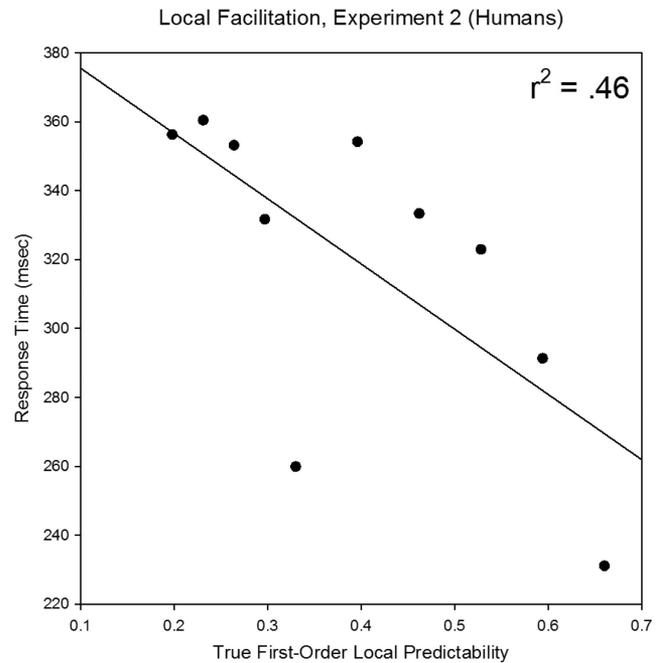


Figure 11. Response times as a function of true first-order local predictability. True first-order local predictability is equal to first-order local predictability times cue validity, and is the probability that the next target will appear in a particular location. Data are from Blocks 5–9 (cue validities of 1.0–0.6).

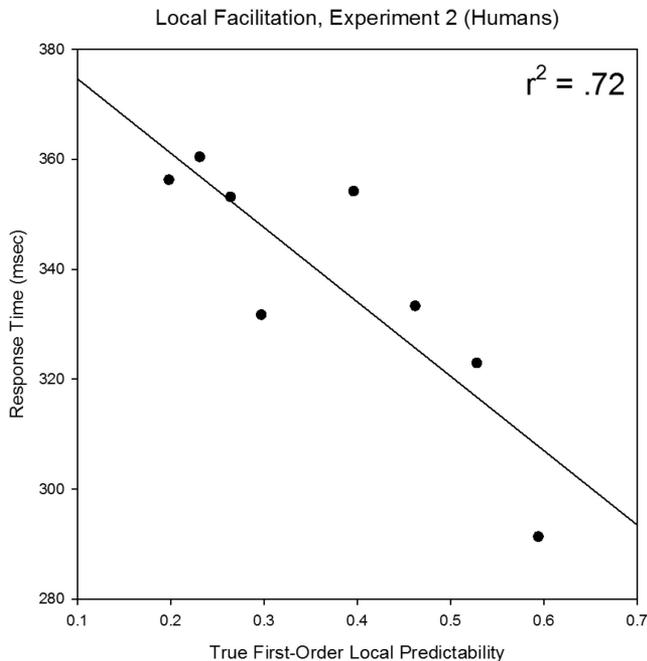


Figure 12. Response times as a function of true first-order local predictability. True first-order local predictability is equal to first-order local predictability times cue validity, and is the probability that the next target will appear in a particular location. Data are from Blocks 6–9 (cue validities of 0.9–0.6).

response transitions. This interpretation might well have been predicted by Willingham and Goedert-Eschmann (1999), who found parallel development of implicit and explicit knowledge using a slightly different SRT procedure.

The primary conclusion of the present experiment is similar to that of the previous experiment: Humans can use knowledge of a response sequence in a flexible manner, even when that sequence is subject to random distortion. This laboratory manipulation may parallel the kind of flexibility shown in skilled performance when, for example, a tennis player subtly adjusts his serve to accommodate a gust of wind, or when a jazz musician improvises on a known melody, performing the song in a novel, but nevertheless recognizable, form.

The knowledge acquired during the experiment, especially at imperfect cue validities, consists largely of item-to-item transition probabilities. We suspect that this knowledge is implicit in nature, whether or not some explicit knowledge develops in parallel. It seems intuitively plausible that imperfect cue validity would severely impair the use of explicit memory, but would not necessarily undermine the usefulness of implicit memory. If that is the case, one would expect results such as those seen in the present experiment. At this point, this interpretation is nothing more than speculation. Fortunately, the memory systems that are presumably involved are biologically dissociable. A biological manipulation that interfered with one memory system but not the other would be the ideal test of the proposition. For example, we would anticipate that amnesic participants' local facilitation data would look more similar to the cleaner function provided by pigeons in Experiment 1.

General Discussion

The parallel facilitation effects shown here in humans and pigeons may not be particularly surprising, because SRT research has been successful in both species before. However, the procedure utilized here introduced variability in a new way, by reducing cue validity. On the other hand, the results could be considered novel in that procedural learning is often investigated using rigid, unvarying methods. In other words, both pigeons and humans perform as if they "know" a specific sequence of cued responses, even in cases in which they have not performed that precise sequence before. Because the initial stages of training involved nothing but valid sequences, the introduction of invalid trials necessarily produced novel sequences, and members of both species responded to those novel sequences with characteristically fast response times.

Both pigeons and humans showed a consistent validity effect that was quite robust at 90% cue validity, and gradually weakened down to 60% cue validity. Below 60% cue validity, there was no response time difference between validly and invalidly cued responses. That the two species showed parallel results to such a degree (see Figures 3 and 8) is striking. The parallel could certainly be coincidence, and only future research will provide the additional necessary details. Given that motor sequences are so generally useful, it is entirely possible that SRT learning relies on a common, evolutionarily old neural system (homology). Alternatively, it could be that the SRT task reflects a fundamental kind of behavior for which multiple species have derived varying solutions (homoplasy). In either case, the task appears to have potential as a useful tool for comparative psychology research.

Even though the similarity between pigeon and human results in these experiments is quite striking, at this point we cannot be certain that the memory systems used by the two species are truly analogous. The memory systems involved in human SRT learning have been identified using various physiological research techniques (see Willingham et al., 2002). Future research should aim to identify the memory systems involved in pigeon SRT learning, using a similar approach. In the meantime, we propose that it should not be surprising if a variety of animals (including, but not limited to, pigeons and humans) possess well-developed procedural memory systems. Furthermore, it should not be surprising that implicit memory demonstrates considerable flexibility. What may be surprising is the impressive degree of complexity that implicit memory can apparently accommodate. Though evidence for the influence of explicit learning of sequences was found only in human participants, we make no claims about the presence or absence of equivalent learning systems in pigeons or other non-human animals. With regard to pigeons in particular, we consider the newer nomenclature for avian brain structures (see Jarvis et al., 2005) as a primary motivating factor to continue to explore possible parallels in avian and mammalian cognition.

The results presented here support the idea that implicit memory may not be quite as rigid as is sometimes implied by procedural tasks that involve exact repetition. Nevertheless, the data also imply that even in the presence of variability, learning still relies on consistent, rule-like factors. In this case, both pigeons and humans appeared to make use of first-order local probability to learn the sequence. In addition, humans may have supplemented first-order local probability with another parallel learning system

(explicit memory) when cue validity was 100%. In conditions with imperfect cue validity, first-order local probability was a better predictor, and sufficient to account for the observed response time facilitation. This finding may have some important implications for cognitive psychological investigations of implicit memory. In particular, imperfect cue validity may be a simple way of eliminating interference from explicit memory, and creating a purer and more direct measure of procedural memory. This is the same conclusion implied by Howard and Howard's (1997) alternating serial response time task, in which a structured sequence alternates regularly with randomly located targets. Their conclusion was that participants responded quickly to particular response triplets that occurred with high frequency because of the regularly alternating structure. Another option is to simply do research using pigeons, whose procedural learning appears to be consistent at all tested cue validities, including 100%.

One exciting theoretical implication of the results presented here is the multimodal applicability of flexible categories. The perceptual world is efficiently divided up into functional categories—classes of stimuli that are treated as equivalent. The present results suggest a similar mechanism for motor responses. It is likely that complex motor behaviors may produce a cognitive structure similar to a prototype in categorization. These “mototypes” display a striking flexibility and tolerance for ambiguity, and gracefully degrade as exemplars become more dissimilar to the prototype. At this point, it is unclear what the necessary conditions for the establishment of such a structure might be. Although the current methodology was effective as a training procedure, we suspect that the order of conditions may be especially important. In the present experiments, participants were first exposed to a perpetually repeating, undistorted sequence. The real world on the other hand, rarely provides such uniformity. If trained using only distortions, it seems likely that extraction of the response transitions that compose the original sequence would be more difficult.

Finally, it is worth remembering that the SRT was initially developed as a tool for investigating the neurobiology of memory. Thus, it seems appropriate to connect the current experiments back to that theme. The brain systems involved in SRT performance have been outlined elsewhere (Willingham et al., 2002). Given that various levels of cue validity may bias individuals toward the use of different memory systems, hypotheses about the effects of physiological manipulations can be derived from the results of these experiments. It is thus our hope that the SRT task may thus continue to provide insight into both the proximate and ultimate causes of sequence learning.

References

- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, 3, 442–481.
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, 56, 149–178.
- Aydin, A., & Pearce, J. M., (1994). Prototype effects in categorization by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 264–277.
- Blaisdell, A. P., Sewa, K., Leising, K. J., & Waldmann, M. R. (2006). Causal reasoning in rats. *Science*, 311, 1020–1023.
- Brown, P. L., & Jenkins, H. J. (1968). Autoshaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior*, 11, 1–8.
- Bullock, D. H., & Bitterman, M. E. (1962). Probability-matching in the pigeon. *American Journal of Psychology*, 75, 634–639.
- Christie, M. A., & Dalrymple-Alford, J. C. (2004). A new rat model of the human serial reaction time task: Contrasting effects of caudate and hippocampal lesions. *Journal of Neuroscience*, 24, 1034–1039.
- Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 189–202.
- D'Amato, M. R., & van Sant, P. (1988). The person concept in monkeys (*Cebus paella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 43–55.
- Froehlich, A. L., Herbranson, W. T., Loper, J. D., Wood, D. M., & Shimp, C. P. (2004). Anticipating by pigeons depends on local statistical information in a serial response time task. *Journal of Experimental Psychology: General*, 133, 31–45.
- Hangi, E. B. (1999). Categorization learning in horses (*Equus caballus*). *Journal of Comparative Psychology*, 113, 243–252.
- Herbranson, W. T., & Schroeder, J. (2010). Are birds smarter than mathematicians? Pigeons (*Columba livia*) perform optimally on a version of the Monty Hall Dilemma. *Journal of Comparative Psychology*, 124, 1–13.
- Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concepts in the pigeon. *Science*, 146, 549–551.
- Howard, J. H., Jr., & Howard, D. V. (1997). Age differences in implicit learning of higher-order dependencies in serial patterns. *Psychology and Aging*, 12, 634–656.
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., . . . Butler, A. B. (2005). Avian brains and a new understanding of vertebrate evolution. *Nature Reviews Neuroscience*, 6, 151–159.
- Jimenez, L., Vaquero, J. M., & Lupianez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning Memory, and Cognition*, 32, 475–490.
- Knowlton B., Squire L., & Gluck, M. (1994). Probabilistic classification learning in amnesia. *Learning and Memory*, 1, 106–120.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32.
- Nissen, M. J., Knopman, D. S., & Schacter, D. L. (1987). Neurochemical dissociation of memory systems. *Neurology*, 37, 789–794.
- Poling, A., Nickel, M., & Alling, K. (1990). Free birds aren't fat: Weight gain in captured wild pigeons maintained under laboratory conditions. *Journal of the Experimental Analysis of Behavior*, 53, 423–424.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353–363.
- Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs. *Animal Cognition*, 11, 339–347.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- 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.
- 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.
- Shimp, C. P., & Friedrich, F. J. (1993). Behavioral and computational models of spatial attention. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 26–37.
- Squire, L. R. (2004). Memory systems in the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, 82, 171–177.
- Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie*, 20, 410–433.
- Vonk, J., & MacDonald, S. E. (2002). Natural concepts in a juvenile gorilla

- (*Gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of Behavior*, 78, 315–332.
- Vonk, J., & MacDonald, S. E. (2004). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology*, 118, 3–13.
- Willingham, D. B., & Goedert-Eschmann, K. (1999). The relation between implicit and explicit learning: Evidence for parallel development. *Psychological Science*, 10, 531–534.
- Willingham, D. B., Salidis, J., & Gabrieli, J. D. (2002). Direct comparison of neural systems mediating conscious and unconscious skill learning. *Journal of Neurophysiology*, 88, 1451–1460.

Received September 27, 2010

Revision received February 8, 2011

Accepted February 9, 2011 ■

Online First Publication

APA-published journal articles are now available Online First in the PsycARTICLES database. Electronic versions of journal articles will be accessible prior to the print publication, expediting access to the latest peer-reviewed research.

All PsycARTICLES institutional customers, individual APA PsycNET® database package subscribers, and individual journal subscribers may now search these records as an added benefit. Online First Publication (OFP) records can be released within as little as 30 days of acceptance and transfer into production, and are marked to indicate the posting status, allowing researchers to quickly and easily discover the latest literature. OFP articles will be the version of record; the articles have gone through the full production cycle except for assignment to an issue and pagination. After a journal issue's print publication, OFP records will be replaced with the final published article to reflect the final status and bibliographic information.