**Title:**

Spatial Variability in Serial Response Learning and Performance by Pigeons (*Columba livia*)

**Journal Issue:**

**Author:**

Herbranson, Walter, Whitman College
Xi, Patricia M., Claremont Graduate University
Trinh, Yvan T., Whitman College alum

**Publication Date:**

**Publication Info:**

**Permalink:**

**Keywords:**

**Local Identifier:**

**Abstract:**

Pigeons can learn structured sequences of cued responses and perform them quickly, even when random variability is later introduced into the originally learned sequence, making some cue locations unpredictable. In order to determine if initial learning shows the same tolerance of spatial variability as steady-state performance, naïve pigeons were trained on random distortions around a structured sequence without having seen the original sequence itself. Learning was possible, but accommodated less variability than did performance of the same sequence previously learned in an undistorted context. Analysis of results indicated that performance of a randomly distorted sequence is best when birds are initially trained with little or no variability, and randomness is later introduced in a gradual fashion.

**Copyright Information:**

eScholarship
University of California

eScholarship provides open access, scholarly publishing services to the University of California and delivers a dynamic research platform to scholars worldwide.

**INTERNATIONAL**
JOURNAL OF **COMPARATIVE PSYCHOLOGY**

# Spatial Variability in Serial Response Learning and Performance by Pigeons (*Columba livia*)

**Walter T. Herbranson, Patricia M. Xi, and Yvan T. Trinh**
*Whitman College, U.S.A.*

Pigeons can learn structured sequences of cued responses and perform them quickly, even when random variability is later introduced into the originally learned sequence, making some cue locations unpredictable.  In order to determine if initial learning shows the same tolerance of spatial variability as steady-state performance, naïve pigeons were trained on random distortions around a structured sequence without having seen the original sequence itself.  Learning was possible, but accommodated less variability than did performance of the same sequence previously learned in an undistorted context.  Analysis of results indicated that performance of a randomly distorted sequence is best when birds are initially trained with little or no variability, and randomness is later introduced in a gradual fashion.

Modern comparative psychology has frequently benefitted from concepts and methods imported from human cognitive psychology.  While some cognitive concepts, such as awareness, are virtually impossible to study in nonhuman animals, others lend themselves quite readily to cross-species investigations.  Procedural memory, for example, refers to memory for how to perform a behavior (Squire, 2004).  As such, investigations of procedural memory are often easier to implement within comparative psychology, as they commonly utilize objective, performance-based measures such as response times, without needing to rely on self-report.

One classic example of a human procedural memory task is the Serial Response Time (SRT) procedure (Nissen & Bullemer, 1987).  The procedure involves a participant responding to a series of cues appearing sequentially in four different locations by pressing four corresponding buttons.  If the cued locations (and consequently, the required responses) appear in a perpetually repeating sequence, response times become faster as a participant learns to anticipate upcoming cues.  The change in response time is due to a participant having learned about the sequence, rather than mere familiarity with the general procedure and apparatus because response times slow immediately when the sequence is later changed to a random one.

A strength of the SRT task is that the dependent variable (median response time) is performance-based, and thus can be easily adapted for use with nonhuman animal populations.  Froehlich, Herbranson, Loper, Wood, and Shimp (2004), for example, had pigeons peck lighted cues that appeared on 3 different response keys in perpetually repeating nine-element sequences such as L-C-R-L-R-C-R-L-C- (where L, C, and R indicate the left, center, and right keys respectively).  Their results mirrored those of a typical human SRT experiment in that response times gradually became faster when cues appeared according to a repeating sequence, and then slowed immediately when cues appeared in a random order.  The difference in response times to predictable cues (in structured conditions) and to unpredictable cues (in random conditions) constitutes *global facilitation*, and is the standard measure of learning in SRT tasks.

While global facilitation is a convenient and concise measure of procedural learning, it may gloss over some details that a finer-grained analysis can reveal. Froehlich et al.'s analysis, for example, indicated that faster response times were a consequence of pigeons having learned probabilities associated with the specific key-to-key transitions that constituted a sequence. As such, some elements in a structured sequence produced faster response times than others. For example, in the sequence specified above, L appears three times, with two instances followed immediately by C, one by R. Correspondingly, the conditional probability of a target appearing on key C given a preceding response on key L is equal to two-thirds. For the same reason, the conditional probability of R given L is one-third. Response times are generally faster on keys having higher conditional probabilities. In this case, if a pigeon had just pecked L, it would respond more quickly to C than to R because the conditional probability $p(C|L)$ is greater than $p(R|L)$. One implication of this finding is that response times are not uniform across a learned sequence. Some transitions within the sequence appear more frequently, and thus contribute disproportionately to the global effect. This key-to-key effect is known as *local facilitation*.

An important consequence of parallel results from humans and pigeons is that the well-established frameworks from human cognition can provide a rich context for investigations of animal learning. For example, SRT learning can progress in humans independently of a functioning hippocampus (Nissen & Bullemer, 1987), as well as under the effects of the amnesia-inducing, anti-cholinergic drug scopolamine (Nissen, Knopman, & Schacter, 1987). These physiological characteristics provide several ready-made hypotheses for parallel manipulations of pigeon SRT learning. In addition to physiological frameworks, several cognitive frameworks simultaneously provide useful context. For example, procedural memory is one of several varieties of implicit memory, along with numerous other learning phenomena such as priming (Tulving & Schacter, 1990), habit formation (Knowlton, Mangels, & Squire, 1996), and artificial grammar learning (Reber, 1967), all of which have also been studied in nonhuman animals (Brodbeck, 1997; Bullock & Bitterman, 1962; Herbranson & Shimp, 2008). The primary characteristic that these different varieties of implicit learning all have in common is their lack of reliance on an intact hippocampus (Squire, 1992).

An additional characteristic that many implicit memory tasks share is that they support considerable variability. For example, weather prediction (Knowlton, Squire, & Gluck, 1994) and artificial grammar learning (Reber, 1967) tasks are two widely used implicit memory tasks, and both involve presentation of stimulus sets that are so complex as to have potentially infinite numbers of exemplars. While the standard SRT stands out among implicit memory tasks in that it does not involve such extreme variability (in fact, the response sequence is highly stereotyped), variability can be incorporated into the procedure and still produce similar results. Howard and Howard (1997) for example, introduced spatial variability into an SRT procedure by alternating predictable cue locations with random ones. Doing so naturally increased the number of different response sequences produced in a session. Nevertheless, even with the increased spatial variability, response time facilitation was still produced. The authors concluded that facilitation was based on learning of specific response pairs and triplets that appeared more frequently than others. This interpretation is consistent with theories about what is learned in other implicit memory tasks. Strings of characters generated by an artificial grammar, for example, may be identified because they are more likely to contain certain letter pairs (bigrams) and triplets (trigrams) than non-grammatical strings (Meulemans & Van der Linden, 1997).

Herbranson and Stanton (2011) took a similar approach to introducing spatial variability in a pigeon version of the SRT. They did so by manipulating cue validity, the likelihood that a particular cue would appear in the location dictated by the sequence. For example, with a cue validity of 0.9, each cue had a 90% chance of appearing on a specific key in accordance with the sequence, and a 10% chance of appearing on a different key, randomly selected from the remaining candidates. After learning the undistorted sequence, cue validity was reduced by 0.1 every 10 days (from 1.0 to 0.9 to 0.8 and so on). As cue validity fell, the magnitude of the global facilitation effect grew correspondingly smaller, down to a cue validity of 0.6. At

values less than 0.6, there was no global facilitation effect and responses to valid cues were as slow as responses to invalid ones. Importantly, in conditions that produced significant global facilitation, response times were still based on local predictability. In this case though, local predictability reflected not just the original sequence, but also the declining cue validity parameter. The likelihood that a cue would appear in a particular location (*true* local predictability) is equal to local predictability (the conditional probability as defined above), times cue validity (the probability that the local predictability figure actually applies), and consistently predicted response times across cue validity conditions.

Herbranson and Stanton's pigeons were able to respond quickly, even when cue validity was not 100% as a consequence of having previously learned the undistorted sequence. The intact fragments of the sequence could be executed quickly because they had already been learned in a consistent, unambiguous context. Invalid (randomly distorted) fragments were executed more slowly because they violated the content of that previous learning. When the proportion of intact fragments was high, the global facilitation effect was slightly diminished, but still strong because the majority of targets were valid. In contrast, when cue validity was low and a larger proportion of transitions had been distorted, there were not enough intact transitions to maintain a strong global facilitation effect. Those local transitional probabilities that underlie SRT learning however, might or might not have been learned in the first place had they been obscured by random variability all along.

There are some reasons to expect that it might be possible to learn a response sequence by performing random distortions of that sequence. For example, the original, undistorted sequence might play a role similar to that played by a prototype in categorization. In a classic study, Posner and Keele (1968) had participants learn categories consisting of randomly distorted dot patterns. Later, when shown the original prototype dot pattern from which training exemplars had been generated, its category membership was more easily classified than the actual training stimuli. Huber and Lenz (1996) showed a similar prototype effect with pigeons, training them on exemplar stimuli that varied across multiple dimensions. Later, pigeons produced enhanced responding to previously unseen stimuli that were near the central tendency of the learned category. In the context of categorization, it seems that both humans and pigeons are able to extract regularly occurring features and use those extracted features to classify prototypical stimuli. If a similar process is at work in serial response learning, then one might see a response time equivalent of a prototype effect: a drop in response times to a previously unseen prototype sequence following training on random distortions of that sequence.

The following experiment introduces spatial variability into a serial response time task in the manner of a category learning experiment: by incorporating variability during the acquisition phase. If learning is able to accommodate variability and pigeons can extract the prototype sequence, then we should see global and local facilitation effects in proportion to cue validity, down to a cue validity value of about 0.6. On the other hand, if acquisition is not as tolerant of variability as performance, then birds might only show learning at higher cue validity values, might show a weakened form of response time facilitation, or might not show response time facilitation at all.

## Method

### Animals

Twelve white Carneau pigeons were purchased from Double T Farm (Glenwood, IA). Each was maintained at 80-85% of free-feeding weight to approximate the condition of healthy wild pigeons (Poling, Nickel, & Alling, 1990). Birds were housed in individual cages in a colony room with a 14:10 hour light: dark cycle, and had continuous access to water and grit. The 12 birds were arbitrarily assigned to four groups of three, designated as: 90% cue validity, 80% cue validity, 70% cue validity, and 60% cue validity.

**Apparatus**

Four identical BRS/LVE operant chambers were used. Each had three circular response keys (2.5 cm in diameter) located in a horizontal row in the front wall, and a food hopper located directly below the center key. A houselight located on the front wall, directly above the center key was illuminated for the duration of each experimental session.

**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 1,000 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, illumination of one of the three response keys with white light (the *target key*), a peck to the illuminated target key, and (on a random 5% of trials) a reinforcement interval. Illuminated keys remained lit until pecked, and there were no programmed consequences for pecks to unlit keys. A single peck to any lit key either initiated a reinforcement interval of approximately 3 s (durations 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.

Target locations for each trial in a session were based on the same fixed nine-element sequence used by Herbranson and Stanton (2011): L-C-R-L-R-C-R-L-C-. Local predictability for individual items in the sequence were either 0.33 (Positions 1, 5, and 6) or 0.67 (Positions 2, 3, 4, 7, 8, and 9). The starting point in the sequence was randomly determined at the beginning of each session. Upon reaching the final element in the sequence, it restarted at the beginning and continued to repeat until the entire 1,000 pecks for a session had been made. The end of one sequence and the beginning of the next were not indicated in any way.

**Blocks of training**. The experiment consisted of three blocks of sessions. The first block lasted for 35 days, and each daily session during that block consisted of a mix of valid and invalid trials, according to the designated cue validity parameter for each group. Birds in the 90% cue validity group saw 90% valid cues and 10% invalid cues. Similarly, birds in the 80%, 70%, and 60% cue validity groups saw 80%, 70%, and 60% valid cues respectively, with the remainder of cues being invalid. The second block of sessions lasted for 15 days, and consisted entirely of valid cues, regardless of group. The third block of trials lasted for 10 days, and was otherwise identical to the first block, with mixed validity trials presented to each group according to its cue validity parameter (90%, 80%, 70%, or 60%).

During the first and third blocks, sequences deviated from the one specified above based on cue validity, the probability that the target key would be the one specified by the fixed nine-item sequence. Thus, a session with cue validity of 100% (as in Block 2) would consist exclusively of validly cued targets, and precisely match the sequence specified above. A session with imperfect cue validity (any value other than 100%) would consist of both validly cued and invalidly cued targets, and thus differ from the original. For example, in a 90% cue validity condition, each target appeared in the location specified by the sequence with $p = 0.9$, and in one of the other two possible locations (randomly chosen) with $p = 0.1$. 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 cue presented 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 during mixed validity blocks 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-*C*-R-C-R-L-C, if the second and fourth (italicized) elements in the sequence were invalid. Note that the remaining valid targets have retained their original relative positions. Each peck to a lit response key had a 5% chance of resulting in reinforcement, regardless of its location and whether or not it was a valid target.

## Results

## Average Response Times

Figure 1 displays the average daily median response time for each of the four groups across the entire experiment. Notice that response times for all groups were initially slow (600 ms or more), and dropped to a consistent level of around 400 ms over the first 10 days of training. By the end of Block 1 (35 days), response

times for each group were relatively consistent from day to day. During the 15 days that constituted Block 2, there appears some notable separation between groups, and that separation remains through the 10 days of Block 3. The source of the between-groups separation appears to be steadily decreasing response times for the 90% ($b = -4.55$, $p < 0.01$) and 80% ($b = -5.48$, $p < 0.01$) cue validity groups during Block 2. The 70% ($b = 0.31$, $p = 0.68$) and 60% ($b = -0.17$, $p = 0.88$) cue validity groups had slopes that were not significantly different from zero during the same block.
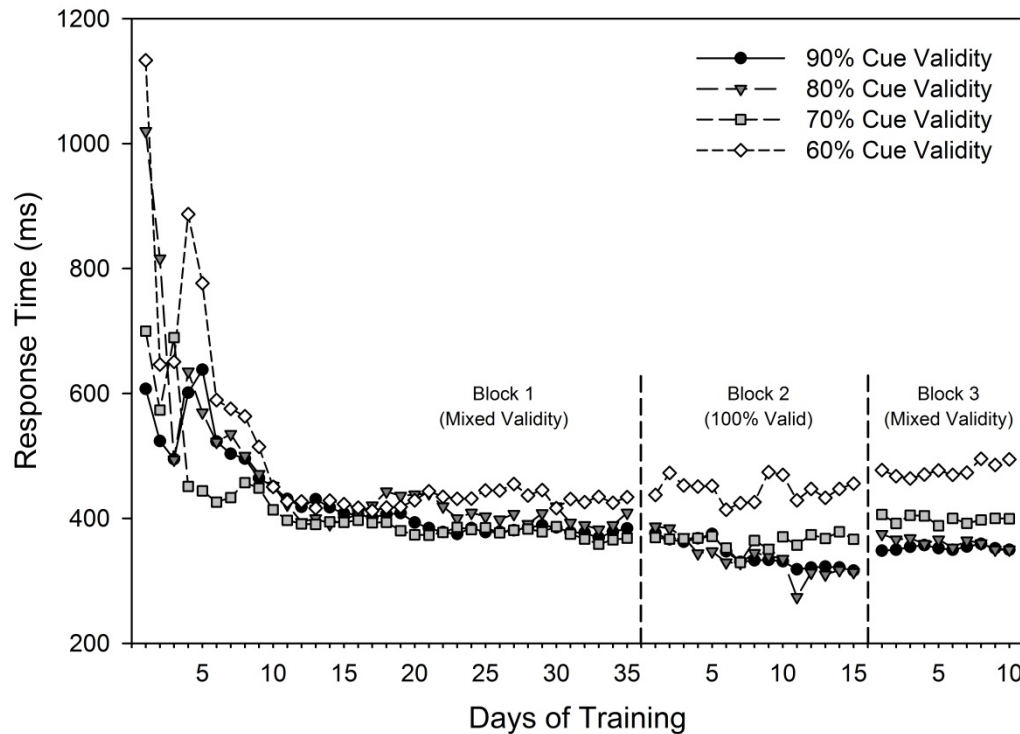


*Figure 1*. Average daily median response time for each cue validity group across the three blocks of the experiment.

In order to statistically assess the effects of the two independent variables (blocking and cue validity), we ran a 3 (block: mixed validity, 100% valid, mixed validity) x 4 (cue validity condition: 90%, 80%, 70%, 60%) mixed ANOVA on daily median response times from the final 10 days of each block. Block was a repeated-measures factor and condition was a between-groups factor. Averages for each combination of block and cue validity are shown in Figure 2. There was a main effect of blocks indicating that birds performed differently across the three blocks that constituted the experiment ($F(2, 72) = 112.60$, $p < 0.01$, partial $\eta^2 = 0.76$). There was a main effect of cue validity condition, indicating that birds performed differently depending on which of the four conditions they had been arbitrarily assigned to ($F(2, 36) = 405.35$, $p < 0.01$, partial $\eta^2 = 0.97$). Finally, there was an interaction between block and cue validity ($F(6, 72) = 36.16$, $p < 0.01$, partial $\eta^2 = 0.75$). These ANOVA results are consistent with the interpretation that cue validity did influence response times, and confirms some instances of prototype effects. Recall that a prototype effect would be indicated by enhanced responding when shown an undistorted pattern (Block 2) after initial training on distortions of that pattern (Block 1). A post-hoc Tukey test confirmed two prototype effects, as there were significant differences between Blocks 1 and 2 for the 90% and 80% cue validity groups ($p < 0.01$ with Bonferroni correction; marked by asterisks), but not for either of the other two groups.
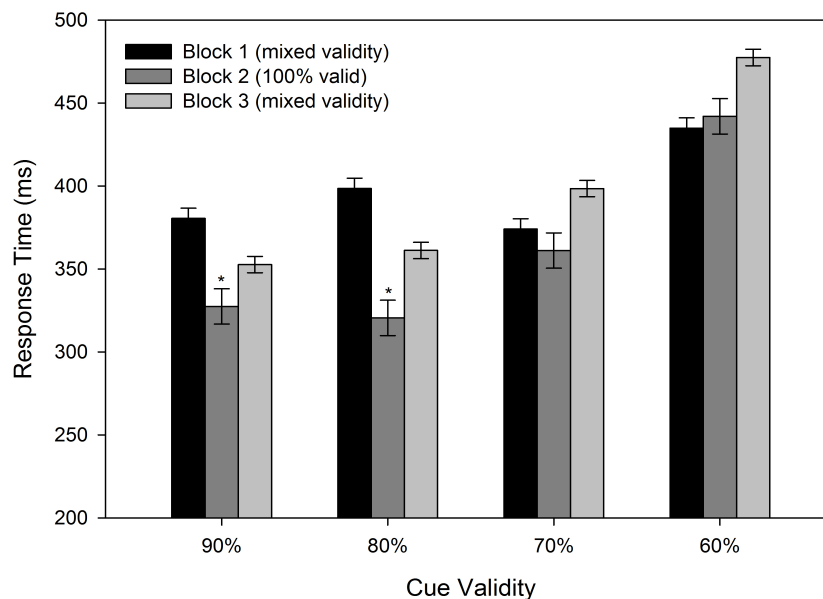
*Figure 2*. Response times for each cue validity group during Blocks 1, 2, and 3. Significantly reduced response times from Block 1 to Block 2 constitute *prototype effects*, and are designated by asterisks.

Because cue validity condition was a between-groups factor, we cannot confidently attribute differences between cue validity groups to cue validity itself. That is, birds in some groups may have been faster than others to begin with. If so, then we risk misattribution of response time differences between groups to a cue validity effect. Therefore, in order to isolate a possible cue validity effect, we must instead look to additional statistical analyses, presented below.

**Cue Validity Effect (Global Facilitation)**

Figure 3 shows the effect of cue validity during mixed-validity Blocks 1 (black bars) and 3 (grey bars). This basic measure of global response time facilitation is computed by subtracting the median response time for validly cued targets from the median response time for invalidly cued targets. In this way, general speed differences between birds are factored out under the assumption that they should affect both types of trials in the same manner. This difference could not be computed for Block 2 because it consisted entirely of validly cued targets. During Block 1, the only group to show a significant cue validity effect was the 90% cue validity group, as its 95% confidence interval does not encompass zero (no difference between valid and invalid trials). The remaining groups were approximately as fast to respond to invalidly cued targets as they were to validly cued targets. On Block 3, the cue validity effect for the 90% cue validity group remained significant, and in fact was considerably stronger than it was on Block 1. Furthermore, the 80% cue validity group showed a small but significant cue validity effect during Block 3, while the remaining groups continued to show no effect. Both mixed-validity blocks from the current experiment show a marked contrast with the data from Herbranson & Stanton (2011; white bars) where there is a significant effect at all four cue validity values.

In order to statistically confirm these observations, we ran a 2 (block: 1, 3) x 4 (cue validity condition: 90%, 80%, 70%, 60%) mixed ANOVA on response time facilitation. Block was a repeated-measures factor and condition was a between-groups factor. There was a main effect of block indicating a stronger facilitation effect during Block 3 ($F(1, 36) = 45.65$, $p < 0.01$, partial $\eta^2 = 0.56$). There was a main effect of cue validity

indicating that higher cue validity values produced stronger response time facilitation ($F(3, 36) = 94.21$, $p < 0.01$, partial $\eta^2 = 0.89$).  There was also a significant interaction between block and cue validity, ($F(3, 36) = 8.44, p < 0.01$, partial $\eta^2 = 0.41$).
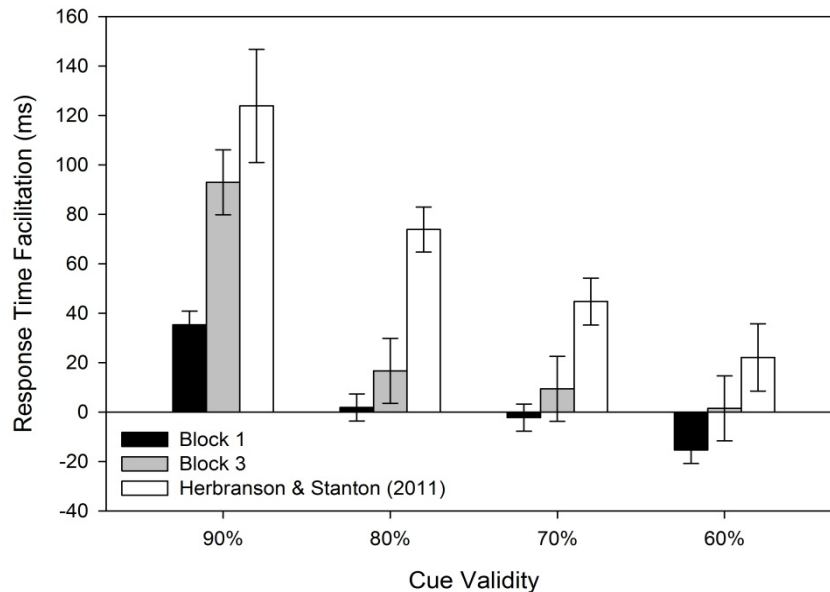


*Figure 3*.  Global response time facilitation (response time to invalid cues minus response time to valid cues) on Blocks 1and 3.  For comparison, white bars show the same measure from birds in Herbranson & Stanton (2011).  Error bars depict 95% confidence intervals.

## Local Predictability

Figure 4 shows data pertaining to local predictability, a possible basis for global response time facilitation.  Recall that local predictability is the conditional probability that a target will appear in a particular location, given the most recent response, and that targets in the sequence used could take on two possible values: 0.33 (hereafter referred to as *low predictability* targets) and 0.67 (hereafter referred to as *high predictability* targets).  Figure 4 shows that during Block 1, the only group to show a local facilitation effect, whereby birds responded to high predictability targets more quickly than low predictability targets, was the 90% cue validity group, and this finding is consistent with the previous analysis of cue validity (global facilitation).  For the remaining groups, birds were just as fast to respond to low predictability targets as they were to respond to high predictability targets.  During Block 2, when all four groups saw the undistorted sequence (100% cue validity), all four groups showed a local predictability effect, indicating that they were capable of learning something about the sequence.  Finally during Block 3, all four groups continued to show a local predictability effect, even when cue validity was no longer 100%.

To statistically confirm these observations, we ran a 3 (block: mixed validity, 100% valid, mixed validity) x 2 (local predictability: low, high) x 4 (cue validity condition: 90%, 80%, 70%, 60%) mixed ANOVA on median response times across the final 10 days of each condition.  Block and local predictability were repeated measures factors and condition was a between-groups factor.  Even though they did not all pertain to hypothesized effects, all of the main effects and interactions were statistically reliable.  There were significant main effects of block ($F(2, 72) = 80.57$, $p < 0.01$, partial $\eta^2 = 0.69$), local predictability ($F(1, 36) = 985.37, p < 0.01$, partial $\eta^2 = 0.97$) and cue validity ($F (3, 36) = 867.95, p < 0.01$, partial $\eta^2 = 0.99$).

There were significant two-way interactions between block and local predictability ($F(2, 72) = 187.36$, $p < 0.01$, partial $\eta^2 = 0.84$), between block and cue validity ($F(6, 72) = 23.67$, $p < 0.01$, partial $\eta^2 = 0.66$) and between local predictability and cue validity ($F(3, 36) = 59.55$, $p < 0.01$, partial $\eta^2 = 0.83$). Finally, there was a significant three-way interaction between block, local predictability, and cue validity ($F(6, 72) = 4.54$, $p < 0.01$, partial $\eta^2 = 0.28$).
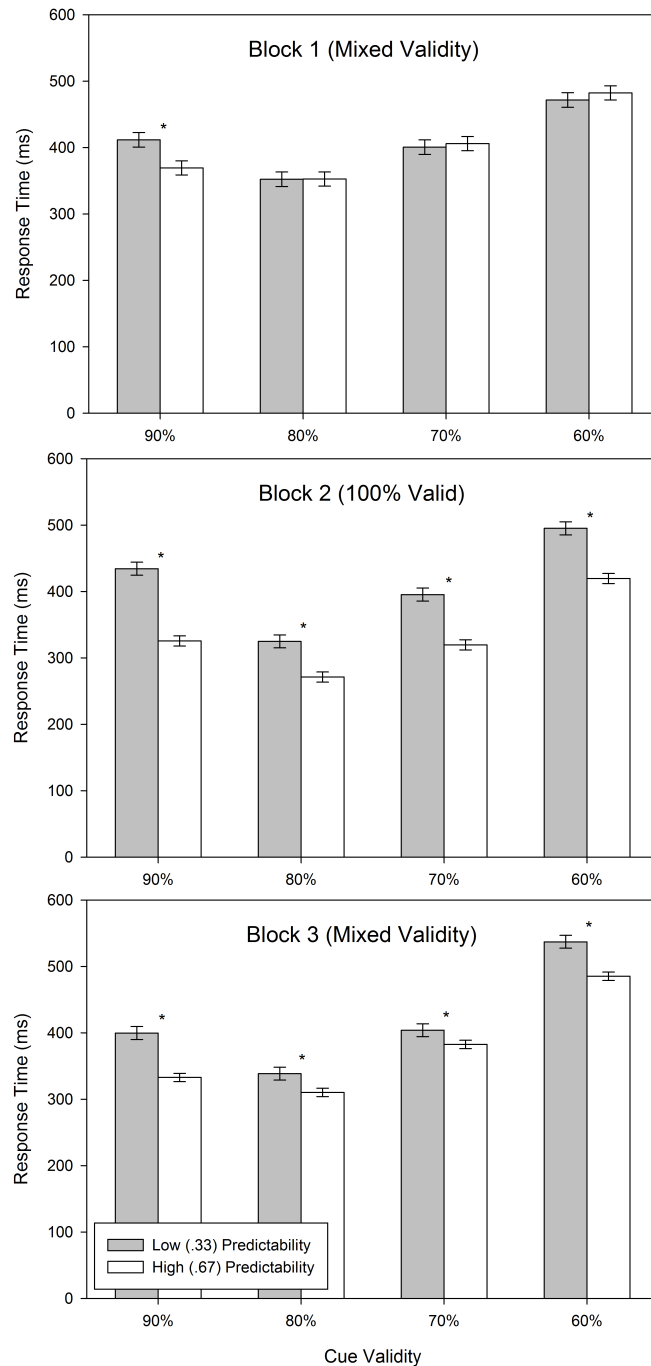


*Figure 4*. Response times as a function of local predictability during the final 10 days of each condition. Error bars depict 95% confidence intervals. Asterisks indicate significant differences between pairs of bars having high and low local predictability values for a given condition and block.

The significant 3-way interaction is essential, as the critical pairwise comparisons pertaining to local facilitation are between low and high local predictability means for a given block and cue validity condition (paired bars in Figure 4). We tested for these 12 possible differences using a post-hoc Tukey test. There were significant differences for all four cue validity groups during Blocks 2 and 3, and for only the 90% cue validity group during Block 1 (all $p$s < 0.01 with Bonferroni correction). These significant differences are indicated in Figure 4 by asterisks.

The results of this ANOVA confirm previously presented analyses, and are what one would expect to see if the global facilitation effect shown in Figure 3 were due to local predictability. The one group that showed global facilitation during Block 1 was also the only group to show a local predictability effect during that same block. The difference between groups that did and did not show global facilitation during Block 1 however, was not due to an inability of the lower cue validity groups to learn transitional probabilities because each showed local facilitation during Block 2. Neither was it due to the inability of the lower cue validity groups to apply knowledge of response transitions (once learned) to a distorted sequence, as each group showed local facilitation during Block 3 after having learned the undistorted sequence during Block 2.

## True Local Predictability

The previous two analyses indicated that response times during Blocks 1 and 3 were influenced by two factors: cue validity and local predictability. Presumably, these two factors both had their effects by influencing how consistently upcoming targets could be predicted. The two factors logically combine in that local predictability (whether high or low) is proportionally diluted by imperfect cue validity (the probability that cue validity applies and a target can actually be predicted by previous responses). Therefore, one can define *true* local predictability as the combined effect of these two factors: local predictability times cue validity. Figure 5 displays response times as a function of true local predictability. Notice that plotted true local predictability values along the x-axis are different for each group because each group's local predictability is modified by a different cue validity parameter. If birds' response times are a function of their ability to successfully predict the location of an upcoming target, then higher true local predictability should yield greater response time facilitation (again defined as the difference between response times to invalidly cued targets and validly cued targets).

During Block 1, the only group that showed either global facilitation or local facilitation was the 90% cue validity group. Figure 5 shows that this same group was the only one for which there was greater facilitation for the higher true local predictability value (indicated by a line with a positive slope). The other groups produced slopes that were essentially flat (in fact, slightly negative), indicating that response time facilitation did not increase with true local predictability. During Block 3 all four groups showed a positive slope, indicating that facilitation increased with true local predictability. The bottom panel in Figure 5 shows the same analysis for the data from Herbranson & Stanton (2011), and the same pattern emerges. While the data from Herbranson and Stanton are more tightly clustered, note that all of their data came from the same birds (cue validity was a within-groups factor) whereas our birds were divided into different cue validity groups (a between-groups factor). The most important pattern is that in both cases, response time facilitation was greater in a given condition when true local predictability was higher.

**Cue Repetition**

Notice in Figure 5 that there are negative response time facilitation values (see the data points representing birds in the 60% and 70% cue validity conditions of the present experiment). This is initially curious, because negative values indicate that birds were actually faster to respond to invalid targets than to valid ones. Consider though, that given the sequence used, consecutive valid targets could not appear in the same location, whereas an invalid target creates such a possibility. Consequently, a bird that is unable to effectively anticipate upcoming target locations, and instead simply monitors the display for a target in any location, might be especially fast to respond to certain invalid targets – specifically those that result in consecutive targets on the same key (and after only a 500 ms ISI), if only because the necessary movement is minimized.

Given the above pattern, it would seem that one specific type of local predictability that could be especially important is repetition of consecutive targets. Because the original, undistorted sequence did not contain any two consecutive targets in the same location, a bird that has learned the sequence would, in part, be anticipating a lack of immediate repetitions (consecutive cues in the same location). Imperfect cue validity (i.e., any value less than 100%) on the other hand, creates the possibility of repetition, and as cue validity declines, repetitions would become proportionally more probable. To assess the importance of cue repetition, we separately calculated median response times to repeated targets (those appearing in the same location as the previous target) and non-repeated targets (those appearing in a different location) during Block 3. Birds in the 90% cue validity group were considerably faster to respond to non-repeated targets ($M = 352.23$, $SD = 4.84$) than to repeated targets ($M = 475.17$, $SD = 23.30$) ($t$ (9) = 14.99, $p < 0.01$). Note that response times to repeated targets were slower despite the fact that by definition they require a bird to move a shorter distance. The other three groups all showed differences of less than 7 ms, indicating that they responded approximately as quickly to repeated and non-repeated targets (all $p$s > 0.25).

**Discussion**

The experiment duplicated the methods of Herbranson and Stanton (2011), using the same sequence and cue validity values, but with one critical difference: Herbranson and Stanton's pigeons were initially trained on the undistorted sequence and only later introduced to variability, whereas our pigeons saw variable sequences from the beginning. These differences in training had a profound influence over performance, as Herbranson and Stanton's pigeons showed significant response time facilitation with cue validities as low as 60%, whereas our pigeons only showed global facilitation if cue validity was 90% (and even then the effect was considerably weakened). After exposure to the undistorted sequence, the performance of our pigeons improved, but still fell short of the results from Herbranson and Stanton, as significant global facilitation was produced only down to 80% cue validity, and when produced was weaker in magnitude. We believe that the differences between the present results and those of Herbranson and Stanton are primarily attributable to the effects of variability during early training. Birds in both experiments were the same strain, experimentally naïve, tested in the same operant chambers, and received the same amount of initial training (35 days).

While all of our birds appeared to have reached steady levels of performance by the end of Block 1, it is also true that birds in different cue validity groups would have seen different numbers of valid trials during those 35 days. Thus, in some sense, they can be said to have received different amounts of training. However, consider that analysis of local predictability indicated that birds in all four groups learned the undistorted sequence within the 15 days that constituted Block 2 (a total of 15,000 valid trials). By comparison, birds in the 90%, 80%, 70% and 60% cue validity groups would have accumulated 15,000 valid trials in approximately 16.67, 18.75, 21.43, and 25.00 days of mixed validity training, respectively. In each case, the 35 day duration

of Block 1exceeded that number of trials, and birds in the 80% and below groups still displayed no traces of either global or local facilitation.
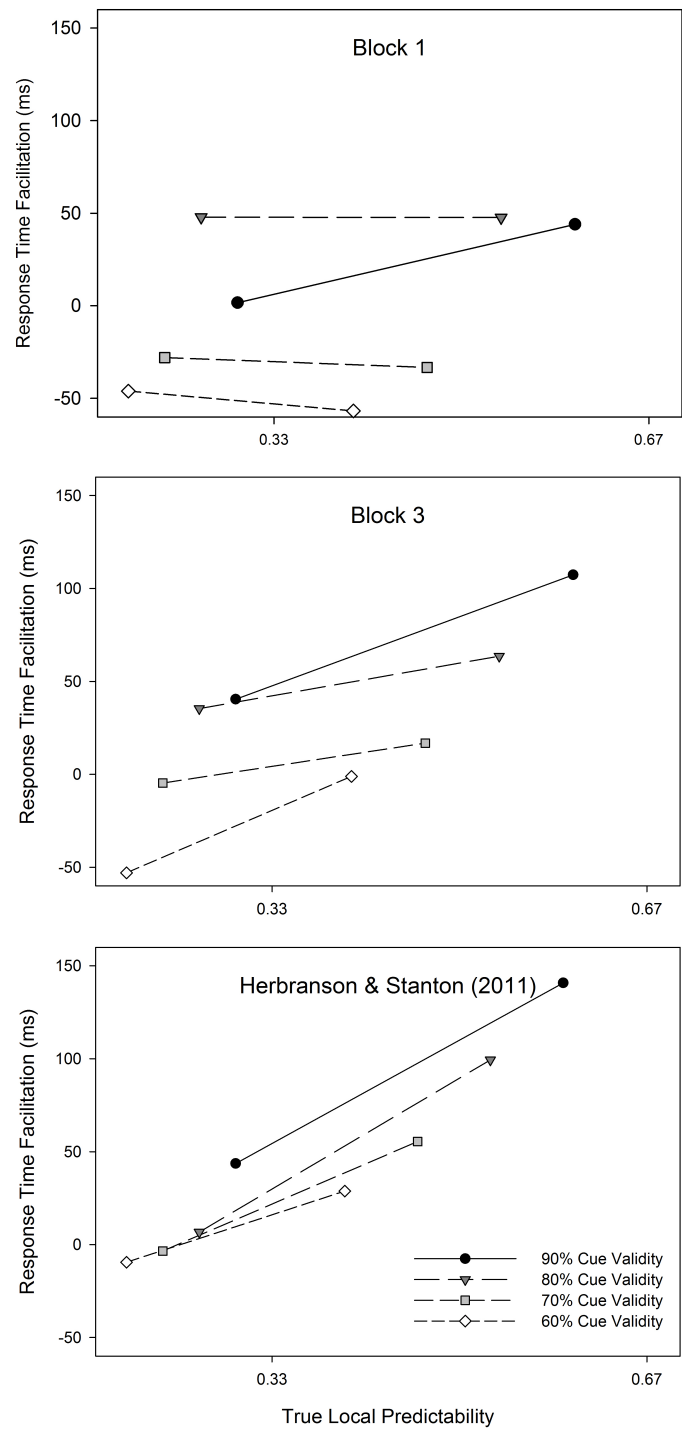


*Figure 5.* Response time facilitation as a function of true local predictability from the mixed validity Blocks 1 (top panel) and 3 (middle panel) and from Herbranson & Stanton (2011; bottom panel).

When interpreting the present results, it is important to remember that the design is somewhat different from a prototypical SRT, in which the critical comparison is between consecutive blocks of trials that are structured (i.e., predictable) and random (i.e., unpredictable). In the present experiment, the critical comparison is between predictable and unpredictable targets within the same block. Similar approaches have been used previously by Herbranson and Stanton (2011) and by Howard and Howard (1997), and bear the advantage of avoiding comparisons across blocks (introducing possible sequencing effects or history confounds) or across groups (making individual differences a concern). Nevertheless, the present approach might be strengthened by including a random control group that is presented with nothing but random targets throughout each block. This control condition would provide for potentially informative comparisons as it would involve the simplest of all strategies (search for any lit key and peck it), stripped of any possible local predictability effects.

Results from the current experiment and from Herbranson and Stanton (2011) both indicated that pigeons can skillfully *perform* variable response sequences with cue validities down to about 60%. The present results however, indicate that learning is less accommodating of spatial variability and that a response sequence can be reasonably *learned* only if cue validity during acquisition is no less than about 90%. Cognitive psychologists have long made a distinction between encoding and retrieval as distinct memory processes (e.g., Tulving & Thompson, 1973), and the present results would indicate that a similar distinction may also be relevant to pigeons' procedural memory, as acquisition and performance are not affected in the same way by variability. While encoding and retrieval are more commonly used in reference to explicit (or declarative) memory, recent neuropsychological evidence has revealed some striking parallels, both physiological and cognitive between implicit and explicit processes for both encoding (Turk-Browne, Yi, & Chun, 2006) and retrieval (Schott et al., 2005).

These results seem to show a loose parallel with those from classic research on categorization (Posner, 1968). That is, our pigeons showed the equivalent of a prototype effect, in which responding to an undistorted prototype is enhanced following training on random distortions around that prototype. In this case though, the prototype was not for a perceptual category, but for a motor response pattern. Despite the name of the effect, note that there is debate over whether individuals actually extract a complete representation of the prototype (see Medin & Smith, 1984). In the present case for example, it appears that pigeons learned about the probabilities of the various response transitions that made up the sequence (local transitional probabilities) rather than extracting a complete intact prototype sequence. While these results mirror the basic pattern of a prototype effect as seen in categorization research, there are some potentially important differences. First, prototype effects in categorization experiments can usually be seen upon first exposure to an undistorted stimulus, whereas the effect shown here did not appear immediately upon exposure to the undistorted sequence (compare the final day of Block 1 with the first day of Block 2 in Figure 1). The effect here refers only to the difference between the stable performance levels achievable under distorted and undistorted conditions. Second, in categorization experiments, one can ensure that the undistorted prototype is not presented until after training is complete. That was not possible in the present experiment. Because we introduced variability via cue validity, it is quite possible that pigeons (especially those in the high cue validity groups) at some point during training saw a sequence of nine consecutive valid targets. Such a chain of responses might or might not be comparable to exposure to an undistorted prototype. That is, the prototype sequence in the present experiment is ongoing and perpetually repeating with no defined beginning or end. Posner's (1968) prototypes on the other hand, were much more precisely defined. Nevertheless, these constitute potentially important differences and should not be overlooked when referring to results from the present experiment as prototype effects.

On a methodological note, these results demonstrate that different measures of performance on a single task can show markedly different sensitivities to learning. For example, our measure of global response time facilitation (Figure 3) by itself would indicate that the 70% and 60% cue validity groups showed no evidence of learning at any point in the experiment, and that even the 80% cue validity group only showed weak facilitation during Block 3. In contrast, our measure of local predictability (shown in Figures 4 and 5) indicates that all four cue validity groups learned something about the sequence by the end of the experiment. It would thus appear that local predictability is a more sensitive measure of response sequence learning than is global facilitation. Local facilitation, however, may not tell the entire story either. While true local predictability does provide an orderly account of the results, we must also consider the possibility that birds in different conditions learned different things. For example, our analyses indicated that birds in the 90% cue validity condition learned transitional probabilities, including the special case that repetitions were rare (one specific kind of local transition). Birds in other conditions did not appear to be as sensitive to target repetitions, but did still show local predictability effects, indicating that other transitional probabilities were learned.

We believe that these results have important consequences for the training of behaviors in a variety of contexts. In the case of SRT learning, it seems that there is a limit to the amount of variability that can be accommodated during training, and consistency is particularly important during early training of a response sequence. Once the sequence is established, variability can be incorporated, but is best incorporated gradually (as in Herbranson & Stanton, 2011). Despite eventual exposure to the undistorted sequence, our pigeons did not show the same magnitude of facilitation as Herbranson and Stanton's pigeons that were not exposed to variability until after the sequence had been learned. Thus, it would seem that the introduction of variability too early might retard learning, and possibly even prevent eventual performance at a level that would have been attainable had an undistorted (or less distorted) sequence been learned initially. Schwartz (1981) showed a similar effect in an experiment where pigeons had to produce a sequence of responses across keys to earn reinforcement. If progress through the sequence was indicated by systematic changes in a matrix of stimulus lights, pigeons developed stereotyped and efficient response sequences. If, however, the matrix light changed randomly, birds failed to develop those efficient response sequences. More importantly, when later exposed to the systematically changing display those birds that had initially seen the randomly changing light matrix failed to develop efficient response sequences. In other words, randomness during initial training prevented the acquisition of an efficient response sequence. This effect of randomness during initial training, however, is not necessarily universal. In humans, for example, it has been well established that variable practice conditions can actually enhance learning of motor skill relative to constant training (Coelho, Nusbaum, Rosenbaum, & Fenn, 2012; Wulf, 1997).

Finally, we believe that these results may reflect general learning processes that pertain to a variety of natural behaviors grounded in procedural memory (locomotion, foraging, social interaction, etc.). Exactly how our cue validity parameter translates into real-world variability is at this point unclear, but it is important to note that variability influences both learning and performance, and that these two processes apparently do not operate in perfect parallel. Specifically, learning appears to require considerably more consistency than does performance. Nevertheless, the natural world is highly variable, and so any useful learning process would almost certainly need to include some factor that can accommodate an appreciable (though not unlimited) degree of variability.

# References

Brown, P. L., & Jenkins, H. J. (1968). Autoshaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior, 11*, 1–8.

Brodbeck, D. R. (1997). Picture fragment completion: Priming in the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 461-468.

Bullock, D. H., & Bitterman, M. E. (1962). Habit reversal in the pigeon. *Journal of Comparative and Physiological Psychology, 55*, 958-962.

Coelho, C. J., Nusbaum, H. C., Rosenbaum, D. A., & Fenn, K. M. (2012). Imagined actions aren't just weak actions: Task variability promotes skill learning in physical practice but not in mental practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1759-1764.

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.

Herbranson, W. T., & Shimp, C. P. (2008). Artificial grammar learning in pigeons. *Learning and Behavior, 36*, 116-137.

Herbranson, W. T., & Stanton, G. L. (2011). Flexible serial response learning in pigeons and humans. *Journal of Comparative Psychology, 125*, 328-340.

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.

Huber, L., & Lenz, R. (1996). Categorization of prototypical stimulus classes by pigeons. *The Quarterly Journal of Experimental Psychology, 49B*, 111-133.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399 – 1402.

Knowlton B., Squire L., & Gluck, M. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1*, 106–120.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. In M. R. Rosenzweig (Ed.), *Annual Review of Psychology. Vol. 35*. (pp. 113-138). Palo Alto, CA: Annual Reviews.

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 1007-1028.

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., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353–363.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855–863.

Schott, B. H., Henson, R. N., Richardson-Klavehn, A., Becker, C., Thoma, V., Heinze, H. J., & Düzel, E. (2005). Redefining implicit and explicit memory: The functional neuroanatomy of priming, remembering, and control of retrieval. *Proceedings of the National Academy of Sciences, 102,* 1257–1262.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99*, 195-231.

Squire, L. R. (2004). Memory systems in the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*, 171–177.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301–306.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.

Turk-Browne, N.B., Yee, D, & Chun, M.M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron, 49*, 917-927.

Wulf, G. (1997). Variability of practice and implicit motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 987-1006.