THE EFFECT OF DISPLAY TIMING ON CHANGE BLINDNESS IN PIGEONS (COLUMBA LIVIA)

WALTER T. HERBRANSON AND EVA T. DAVIS

WHITMAN COLLEGE

Change blindness is a phenomenon in which even obvious changes in a visual scene may go unnoticed. Recent research has indicated that this phenomenon may not be exclusive to humans. Two experiments investigated change blindness in pigeons, using a variant of the widely-used flicker task to investigate the influence of display timing on change blindness. Results indicate that the duration of time during which a stimulus display is visible influences change detection accuracy, with the effect due to additional search time. The results are discussed in relation to the value of comparative cognition and cross-species investigations of behavior.

Key words: change detection, change blindness, attention, pigeon

Behavioral science has always been centrally concerned with learning, and some of the most fundamental and useful indicators of learning are correct responses. In an operant conditioning scenario, for example, the dependent variable is often the number of times a particular desired behavior is produced; with training those responses generally increase. The number and rate of correct responses can be easily quantified and seen, for example, in a cumulative record. Nevertheless, one need not rely solely on correct responses, as other performance-based measures can be just as useful, and may help to further characterize aspects of behavior not easily seen by focusing only on correct responses.

Response latencies, for example, have become increasingly useful tools. Within the tradition of behavior analysis, interresponse times have been widely used, and have provided some valuable insights into the nature of different schedules of reinforcement (Shimp, 1969). Response times can also constitute an alternative measure of learning by themselves that carry distinct advantages over more traditional measures such as correct responses: They can be sufficiently sensitive to detect differences

between conditions that might otherwise go undetected (e.g., Herbranson & Stanton, 2011).

In addition to the approaches already mentioned, another potentially important indicator of performance is incorrect responses; they can provide some valuable clues about the learning mechanisms at work. In a matching to sample task, it is often the incorrect responses that provide the greatest insight into an animal's performance. Introduction of a retention interval in delayed matching to sample generally produces more errors, and those errors increase with the duration of the delay (Blough, 1959), reflecting the influence of memory. Errors are also more likely to be made among similar stimuli, and again these kinds of errors can be used to make inferences about the behavioral and/or cognitive processes involved. Roitblat (1980), for example, cleverly used errors in a symbolic matching-to-sample task to investigate whether pigeons use prospective or retrospective coding to bridge the delay. The important implication is that errors, when they occur, are not random and arbitrarily distributed; some kinds of errors are more likely to occur than others, and the pattern of errors obtained can be highly informative.

The same observation applies to other areas of psychology. Many of the classic tools used to study cognition also play a significant role for sophisticated analyses of errors. Signal detection theory (Green & Swets, 1966), for example, considers correct responses (hits and correct rejections) as well as incorrect responses (misses and false alarms). Again, it is the incorrect responses that often provide the

greatest insight into the processes that influence behavior but cannot be directly observed (i.e., bias, sensitivity, etc.).

The field of comparative cognition stands to draw from its intellectual ancestors in similar ways. Analysis of specific kinds of errors have formed the basis of many research traditions: Rotational errors have provided insights into spatial navigation (Cheng, 1986), anticipation and perseveration errors have provided insights into midsession reversal learning (Stagner, Michler, Rayburn-Reeves, Laude, & Zentall, 2013), and within- and between-category errors have provided insights into the basis of categorization (Wasserman, Kiedinger & Bhatt, 1988). In each case, not all kinds of errors can or should be treated as equivalent; some are more likely to occur than others, and different types of errors happen for different reasons.

Change Blindness in Pigeons

An increasingly important phenomenon in cognitive psychology that is largely defined by a particular pattern of errors is change blindness: the failure to notice something different about a display from one moment to the next (Simons & Ambinder, 2005). Importantly, these failures to detect changes do not happen arbitrarily. The likelihood of an individual failing to detect a displayed change is systematically affected by a variety of factors. Grimes (1996), for example, demonstrated that sudden changes in a visual display were more likely to go unnoticed if their timing coincided with an eye saccade. The same changes were easily seen if presented between saccades, while the eyes were fixated.

Change blindness thus has some obvious real-world implications. Missing a prominent change in one's immediate environment has clear potential dangers (e.g., failure to notice the appearance of a danger or threat). An additional fascinating (as well as concerning) aspect of change blindness is that it persists even if the participant is actively searching for the change, and even if the participant is aware of the change blindness phenomenon ("change blindness blindness"; Levin, Momen, Drivdahl, & Simons, 2000). These surprising aspects of change blindness bring up the question of whether it is a unique quirk of human behavior or if it might also be seen in other animals.

Pigeons also would seem to possess the basic cognitive abilities required to test for change blindness. In particular, they can be trained to search for and notice differences between sequentially presented visual displays. Cook, Kelly, and Katz (2003) trained pigeons on a successive same/different task, in which they viewed alternating presentations of two photographic images. Pecks were reinforced only if the two images were the same. Birds learned the task and eventually showed above-chance discrimination by the presentation of the second image, the earliest point possible. Furthermore, pigeons can learn to search for and selectively peck a localized change within a larger visual display. Wright et al. (2010) created a change detection task in which pigeons had to peck one of two colored circles in a test array, and pecks to the circle that was not the same color as in the previously displayed sample (i.e., changes) were reinforced (also see Gibson, Wasserman, & Luck, 2011, and Lazareva & Wasserman, 2015, for other change detection tasks that have been successfully deployed for pigeons). Their results indicated that pigeons could learn to reliably detect localized color changes, and could do so even when successive displays were separated by a time delay.

One of the most commonly used procedures for studying human change blindness in a controlled laboratory setting is the "flicker paradigm" (Rensink, O'Regan & Clark, 1997). In this procedure, two stimulus displays (such as arrays of alphanumeric characters, or digitally manipulated photographic images) are displayed in successive alternation. The displays are identical, save for a single feature, and a participant is asked to find the difference. If the two displays are shown in immediate succession, with no temporal gap between them, then the difference is usually identified very quickly. Most participants report that the change "jumps out," and automatically captures their attention. On the other hand, if there is a brief inter-stimulus interval (ISI) between the displays, during which neither display is visible (producing the flickering display for which the procedure is named), then the difference is much more difficult to find, requiring more time and producing lower accuracy rates.

The flicker paradigm has been appealing to researchers interested in change blindness because it is easy to isolate and manipulate specific stimulus characteristics, such as the size, location, and salience of the change, as well as the timing of stimulus displays and ISIs. It also

provides a concise operational definition of change blindness: the difference in accuracy between trials featuring an ISI, and those without. Furthermore, awareness of the phenomenon and extensive training do not seem to eliminate the effect. These characteristics would hold the same appeal for investigations of attention in nonhuman animals. Other established procedures for studying change blindness and/or inattentional blindness can be considerably more complex, and would not be as easily adapted for use with nonhuman animals.

Based on the aforementioned appealing features, Herbranson et al. (2014) developed a version of the flicker task that could be implemented in an operant chamber. Although a standard operant chamber has a limited number of response keys, multiple features can be simultaneously presented on each. Their methodology utilized displays containing up to 24 distinct line features across the three response keys. As in the flicker paradigm previously used with human participants, each trial presented two alternating displays that differed by a single feature, either with or without an ISI. At the end of a trial, pigeons received access to mixed grain if and only if they pecked the key on which there was a difference between the displays.

Herbranson et al. (2014) produced two primary results relevant to change blindness; both paralleled results from human change blindness research. The primary result involved a comparison between accuracy on trials with an ISI and trials with no ISI. Although birds performed at above chance levels on both kinds of trials, they were significantly better on trials with no ISI. The second important result involved the question of whether change detection involves a serial search process as it does in humans. Consistent with a serial, location-by-location search, pigeons' accuracy increased with the total number of repetitions on a trial. In addition, a third finding was that pigeons' effective search area became larger, encompassing more locations with better-than-chance performance, as the number of repetitions increased. Specifically, on shorter trials with fewer repetitions, pigeons could detect changes in two of the three possible locations, but they were no better than chance when changes occurred in the third. On longer trials, they could reliably detect changes in all three possible locations.

Experiment 1

Although the primary results from Herbranson et al. (2014) are consistent with change blindness as it is seen in humans, we do not yet know the extent to which the mechanisms that produce parallel results in the two species are the same. Such processes can often be inferred by manipulating various stimulus characteristics and observing what effect (if any) they have on performance. As in the research traditions cited above, potentially informative errors are reflected by differences in overall accuracy.

Pashler (1988), for example, manipulated the timing of visual displays in the flicker task. He found that accuracy was strongly influenced by the duration of the ISI: as the ISI became longer, more errors were produced. In contrast, there was no such influence of the duration of each individual display: As the duration of individual alternating displays between ISIs changed, accuracy remained constant. Consequently, he could infer that errors increased with ISI duration (i.e., overall accuracy decreased), but not with stimulus duration. Again, the different accuracy rates revealed a specific and nonrandom pattern of change detection failure.

Herbranson et al. (2014) systematically manipulated ISI duration in the same manner as Pashler (1988) and found parallel results: Longer ISIs yielded more change detection failures. They did not, however, manipulate the durations of the stimulus displays in the same manner. Thus, we do not know if this additional aspect of stimulus timing considered by Pashler would produce a similar effect (or in this case, non-effect) in pigeons. Experiment 1 manipulated exposure duration for the stimulus displays while holding ISI duration constant. Results should reveal whether display duration has any consistent effect on accuracy (as ISI duration has been shown to have).

Method

Animals. Four white Carneau pigeons (*Columba livia*) were purchased from Double-T Farm (Glenwood, IA). Each bird was fed mixed grain and maintained at 80-85% of free-feeding weight to approximate the condition of healthy wild birds (Poling, Nickel, & Alling, 1990). Birds were housed in individual cages in a colony room with a 14:10-hr light:dark cycle and had unrestricted access to water and grit.

All four had previous experience with a serial response time task (Herbranson & Stanton, 2011) and a change detection task (Herbranson et al., 2014; Herbranson, 2015).

Apparatus. Four identical operant chambers were used (BRS/LVE, Laurel MD: Cubicle SEC-002 with response panel PIP-016). Each had three circular response keys (2.5 cm in diameter) located in a horizontal row on the front wall, 5.8 cm apart, edge-to-edge. A food hopper with a 5.0×5.8 cm (width $\times$ height) opening was located 9.0 cm directly below the center key. A houselight located on the front wall, 4.5 cm directly above the center key, was illuminated for the duration of each experimental session.

Stimuli. Stimuli consisted of straight white lines back-projected onto each response key using stimulus projectors (Industrial Electronic Engineers, Van Nuys, CA) that had been retrofitted with LED light sources (Martek Industries, Cherry Hill, NJ). The LED light modifications were necessary because their onset and offset times (approximately 30 μs) are much faster than incandescent bulbs, allowing for precise control of even very fast stimulus presentations and ISIs. The three keys each could display up to eight radial lines, with each line spanning the full diameter of the key. The lines were positioned at evenly spaced orientations corresponding to 0.0°, 22.5°, 45.0°, 67.5°, 90.0°, 112.5°, 135.0°, and 157.5° from vertical.

On each trial, a base stimulus was generated according to the following parameters: each of the eight lines on each of the three keys independently had a .5 chance of being present and a .5 chance of being absent. Consequently, each stimulus could consist of anywhere from 0 to 24 lines across the keys (0-8 lines per key). A modification of that base stimulus was then generated by reversing the display status of two of the lines on a single key. If a line to be reversed was present in the base display, then it was not present in the modified display. Conversely, if it was not present in the base display, then it was present in the modified display. The critical, changing key was equally likely to be any of the three keys on a given trial, and each change was equally likely to occur in any of the eight orientations on the critical key.

Each trial consisted of alternating presentations of the original and modified displays (the duration of each display varied by condition

between 15 ms and 125 ms[1]; see below). The alternating displays were presented 1, 2, 4, 8, or 16 times (randomly determined on each trial with $p = .2$ for each). Each presentation of the original display was followed by the modified display and each presentation of the modified display was followed by either by a repetition of the original display or a trial-terminating display consisting of all three keys uniformly illuminated with white light (if and only if it was the final repetition of the trial).

Half of the trials presented the two alternating displays with no time delay in between. The modified display was presented immediately after the base display, so that there was no time when one of the two displays was not present on the response keys until the trial-terminating display of white keys. The other half of the trials contained a 30-ms ISI between the displays, during which the keys were completely dark and no lines were visible. The ISI was then followed immediately by the modified display. Thus, on trials with an ISI, the same number of repetitions took longer because each stimulus display was followed by an ISI delay. Figure 1 shows the structure of two stimulus displays, one with an ISI and one without.

Procedure. The experiment was conducted over four blocks of 10 days each (40 days total). Daily sessions consisted of 120 trials, with each trial separated by a 5-s intertrial interval (ITI), during which the houselight remained illuminated. During this ITI, the computer program generated original and modified displays, as well as determined the number of repetitions and whether or not to include an ISI. At the conclusion of the ISI, the stimulus was automatically presented, with no preceding cue or required response. Pecks during stimulus presentation were not recorded and had no programmed consequences. Following completion of the entire stimulus display, all three keys were uniformly illuminated with white light and the first peck on any key was automatically recorded. If the peck corresponded to the location of the stimulus change, then the bird was presented

---

[1] Several aspects of the stimulus displays were chosen based on previous research in hopes of yielding a moderate level of accuracy (to avoid potential floor and/or ceiling effects), and maximizing the contrast between ISI and no-ISI trials. Number of changes was based on results from Herbranson (2015). Values for stimulus and ISI durations were informed by the results from Herbranson et al. (2014).
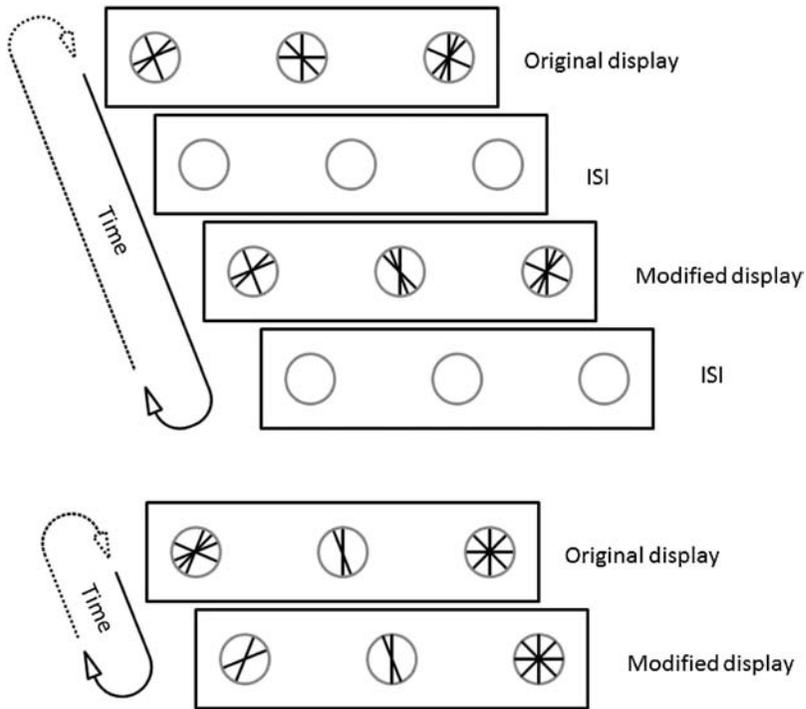
Fig. 1. Example stimulus displays with (top) and without (bottom) an ISI. Adapted from Figure 1 of Herbranson, 2015.

with approximately 3-s access to mixed grain (access times varied from bird to bird in order to maintain individual running weights). If the peck corresponded to either of the other two, unchanging locations, then a 10-s error signal was presented, during which the houselight flashed on and off every 0.5 s. After either a reinforcer or the error signal, the session continued with an ITI, followed by the next trial.

Conditions. Four blocks of 10 days were identical with the exception of the duration of the stimulus displays. During the first block, each presentation of a line feature display (both original and modified) was visible for 125 ms, regardless of any other stimulus characteristics (ISI presence, number of repetitions, etc.) This was the baseline condition and paralleled the procedure from Herbranson et al. (2014), with the exception of the display timing (ISI and stimulus durations) and number of changing features on the critical key (two rather than one). The subsequent three blocks were identical except for stimulus durations of 60, 30, and finally, 15 ms.

Because all four birds had previous experience on the flicker task, no pretraining was

necessary and data collection could begin immediately.

Results

Factors influencing accuracy. A 2 (ISI: present, absent) $\times$ 5 (repetitions: 1, 2, 4, 8, 16) $\times$ 4 (stimulus duration: 125 ms, 60 ms, 30 ms, 15 ms) $\times$ 10 (day: 1-10) repeated measures ANOVA was conducted on change detection accuracy. There was no main effect of day nor were there any statistically significant interactions involving day (all $F < 2.17$, $p > .058$), with the sole exception of a two-way interaction between day and stimulus duration ($F(27, 81) = 1.86$, $p = .018$, partial $\eta^2 = .382$). Thus, the effects of the remaining manipulated variables (stimulus duration, repetitions, and ISI) presumably reflect relatively stable performance across the 10 days of each condition, and can be seen in Figure 2.

The main effect of ISI was significant, $F(1, 3) = 18.67$, $p = .023$, partial $\eta^2 = .862$, indicating that accuracy was better on trials that did not have an ISI ($M = 62.52\%$) than on trials that did ($M = 51.54\%$). This is a replication of the basic change-blindness effect
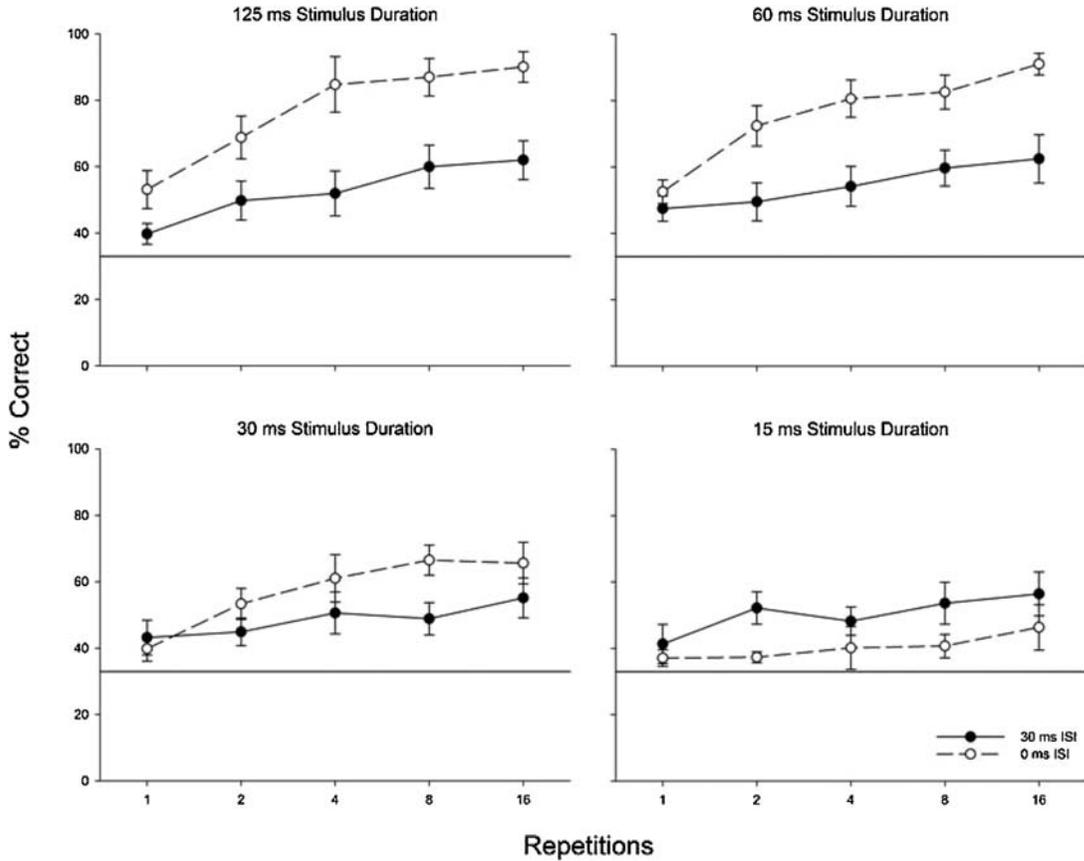
Fig. 2. Accuracy on ISI trials (solid lines) and no-ISI trials (dashed lines) on trials with varying numbers of repetitions during each of the four conditions in Experiment 1. Error bars span ± 1 standard error. Top left: 125-ms stimulus duration. Top right: 60-ms stimulus duration. Bottom left: 30-ms stimulus duration. Bottom right: 15-ms stimulus duration.

seen in previous implementations of the flicker task in both humans and pigeons. The main effect of repetitions was also significant, $F(4, 12) = 33.76$, $p < .001$, partial $\eta^2 = .918$, indicating that accuracy increased with more repetitions from an overall mean of 44.28% at one repetition to 66.10% at 16 repetitions. The main effect of stimulus duration was significant, $F(3, 9) = 20.82$, $p < .001$, partial $\eta^2 = .874$, with overall accuracy increasing with longer presentation times, from 45.34% at 15 ms to 64.70% at 125 ms. These last two findings are both consistent with the location-by-location search process reported by Herbranson et al., (2014): When given more time to search (either through additional repetitions or longer presentation times), pigeons could consider more locations and thus should be more likely to detect changes.

In addition, all interactions involving these manipulated factors were significant: ISI × repetition, $F(4, 12) = 6.34$, $p = .006$, partial $\eta^2 = .679$; ISI × duration, $F(3, 9) = 197.75$, $p < .001$, partial $\eta^2 = .985$; repetition x duration, $F(12, 36) = 3.56$, $p = .002$, partial $\eta^2 = .543$; and ISI × repetition × duration, $F(12, 36) = 3.33$, $p = .003$, partial $\eta^2 = .526$. At the core of the essential 3-way interaction is a pattern in which accuracy on ISI trials (points on the solid lines in Fig. 2) remained relatively stable across the four stimulus duration conditions, whereas accuracy on no-ISI trials (points on the dashed lines) decreased as stimulus durations were shortened in subsequent conditions. Note also that at the shortest duration (15 ms), the change blindness effect was reversed, in that accuracy on ISI trials was better than accuracy on no-ISI trials.

Effect of stimulus duration on ISI and no-ISI trials. To further explore the interactions involving ISI and stimulus duration, two separate 5 (repetitions: 1, 2, 4, 8, 16) × 4 (stimulus duration: 125 ms, 60 ms, 30 ms, 15 ms) ANOVAs were computed on just ISI and just no-ISI trials (solid lines and dashed lines in Fig. 2, respectively). For trials with an ISI, there was a main effect of repetitions, $F(4, 12) = 26.70$, $p < .001$, partial $\eta^2 = .899$, but no main effect of stimulus duration, $F(3, 9) = 1.92$, $p = .198$, partial $\eta^2 = .390$. There was also no interaction between repetitions and stimulus duration, $F(12, 36) = 1.60$, $p = .137$, partial $\eta^2 = .347$. This result indicates that the effect of repetitions on accuracy during ISI trials was consistent and insensitive to the different stimulus durations across conditions (i.e., the solid lines in each panel of Fig. 2 are effectively indistinguishable from one another). For trials with no ISI, a different pattern emerged. There was a main effect of repetitions, $F(4, 12) = 23.88$, $p < .001$, partial $\eta^2 = .888$, as well as a main effect of stimulus duration, reflecting poorer performance at shorter durations, $F(3, 9) = 51.14$, $p < .001$, partial $\eta^2 = .945$. In addition, there was an interaction between repetition and stimulus duration, $F(12, 36) = 4.72$, $p < .001$, partial $\eta^2 = .611$. This result indicates that the effect of repetition on accuracy during no-ISI trials was not consistent across conditions (i.e., the dashed lines in each panel of Fig. 2 become flatter as stimulus durations become shorter).

Discussion

Pigeons' ability to detect changes in Experiment 1 was influenced by all three manipulated variables: ISI, repetition, and stimulus duration. The former two coincide with previous research on both pigeons (Herbranson et al., 2014) and humans (Pashler, 1988). The latter factor (stimulus duration), however, had not yet been investigated in pigeons; its effect appears to directly contrast with the extant human data. The obtained ISI effect—presence of an ISI impairs change detection relative to no ISI—is a basic confirmation of the operational definition of change blindness, without which no other results could be interpreted. The repetition effect, while not a new finding, confirms that change detection happens across time: Additional repetitions afford more chances, and more time, to identify a changing feature. The primary novel result of Experiment 1 was that change detection was impaired by reduced stimulus duration; the impairment occurred exclusively on trials without an ISI. As a consequence, the magnitude of the change blindness effect (accuracy on no-ISI trials minus accuracy on ISI trials) became smaller as stimulus durations became shorter.

Recall that Herbranson et al. (2014) demonstrated that change detection accuracy decreased as ISI duration was lengthened, indicating that at least one aspect of stimulus timing has a consistent influence on change blindness. The present results concern a different aspect of stimulus timing: that of the displays themselves, rather than the ISIs between displays. Note that the pattern of results is importantly different: Shorter ISIs resulted in better performance (specifically on ISI trials), whereas shorter stimulus durations resulted in worse performance (specifically on no-ISI trials).

This disparity provides an important caveat to previous findings. Whereas the beneficial effects of repetition and of longer stimulus durations might be explained by the additional search time they afford, ISI effects cannot be explained in the same way. That is, the presence of an ISI by definition *lengthens* a trial, yet *impairs* accurate change detection. Furthermore, longer ISIs lengthen trials more than shorter ones, yet result in poorer performance.

Finally, note that the presence of an ISI does not change the cumulative amount of time that the original and modified displays are visible (i.e., time when the displays can actually be searched). Thus, there is something other than search time (possibly the temporal gap between displays) that negatively influences performance on ISI trials, making them importantly different from no-ISI trials.

This last point is underscored by the surprisingly low accuracy on no-ISI trials with 15-ms stimulus durations. Performance on these trials was actually lower than on parallel trials with an ISI (meaning that the normal change blindness effect was reversed). The reversal could be caused by the absolute magnitude of the stimulus duration (15 ms). By this explanation, the total stimulus presentation time of 30 ms (15 each for the original and modified display) times the number of repetitions was simply too

brief to accommodate all of the requirements for an accurate response.

At this point, those specific requirements cannot be specified, but they might include both central (e.g., attention) and peripheral (e.g., physical orienting or inter-key travel) processes. Only with the inclusion of an ISI was there adequate time to support accurate change detection. This explanation though, would not seem to account for all of the data.

Consider that the 480 ms required to display 16 repetitions (15 ms each for the original and modified displays times 16 repetitions) would be equal to the time required for 5.33 repetitions with a 15 ms display and 30 ms ISI, 4.00 repetitions with a 30 ms display and 30 ms ISI, 2.67 repetitions with a 60 ms display and 30 ms ISI, or 1.55 repetitions with a 125 ms display and 30 ms ISI. Figure 2 shows that in each case, one would predict better accuracy for an ISI trial with the same overall duration. Alternatively, consider that the 15 ms stimulus duration condition was the only one in which the ISI duration was longer than the stimulus duration. If duration relative to the ISI (as opposed to absolute duration) of stimulus displays is important, then one might expect impaired performance any time a greater proportion of the total display time is allocated to the ISI than to the stimulus display. That is, accuracy might be based not just on the available time to search a display, but also on how that search time is divided into stimulus display and ISI durations.

It has previously been shown that the relative durations of a stimulus and the interval between stimuli has a powerful effect on the acquisition and extinction of autoshaped key-pecks (Gibbon, Baldock, Locurto, Gold, & Terrace, 1977). Thus, there is at least one aspect of behavior that is sensitive to the relative durations of stimuli and interstimulus intervals. However, the specific mechanisms at work may be quite different in this case, given the differences in procedure (change detection versus autoshaping) and the time scales involved (15-125 ms versus 1-384 s). Experiment 2 further explores the importance of the ratio of stimulus duration to ISI.

Experiment 2

In order to see if pigeons' behavior is affected by the relative durations of the ISI

Table 1

Summary of Conditions in Experiment 2

Condition	Stimulus Duration (ms)	ISI Duration (ms)	Time per Iteration (ms)
1	15	30	45
2	30	15	45
3	30	60	90
4	60	30	90
5	60	125	185
6	125	60	185

Note. the final column (time per repetition) is the sum of the previous two (Stimulus Duration and ISI Duration).

and stimulus presentations, six additional conditions were run (see Table 1). The first two conditions consisted of trials having a 15-ms stimulus duration and 30-ms ISI (identical to the fourth condition of Experiment 1), and trials having a 30-ms stimulus duration and 15-ms ISI, respectively. Notice that because the durations of the stimulus presentations and ISIs are simply reversed between conditions, the corresponding trials from each would span the same overall amount of time, and allow pigeons exactly the same total search time (45 ms per presentation). The remaining four conditions used the same logic, but with longer total search times: 90 ms per presentation for the third and fourth conditions (combining 30-ms and 60-ms stimulus durations and ISI values), and 185 ms per presentation for the fifth and sixth conditions (combining 60-ms and 125-ms stimulus duration and ISI values). If overall search time is the critical factor, then pairs of conditions having the same search time per repetition should produce identical accuracy on ISI trials. On the other hand, if the relative timing of the stimulus displays and ISIs is important, then there should be differences between pairs of conditions.

The conditions defined here contain some additional comparisons of interest. Note that there are two instances of conditions with 30-ms stimulus durations. One is presented in the context of a shorter ISI (15 ms in Condition 2) and the other is in the context of a longer ISI (60 ms in Condition 3). Similarly, two instances of 60-ms stimulus durations are presented in the context of a shorter ISI (30 ms in Condition 4) and a longer ISI (125 ms in Condition 5). If stimulus duration alone determines accuracy (presumably just on no-ISI trials, based on the results from Experiment 1), then we should see

identical accuracy between these pairs of conditions featuring the same stimulus durations.

The same logic just applied to 30- and 60-ms stimulus durations can also be applied to 30- and 60-ms ISIs. ISIs of 30 ms appear twice: in the context of both a shorter stimulus duration (15 ms in Condition 1) and a longer stimulus duration (60 ms in Condition 4). Finally, a 60-ms ISI appears in the context of a shorter stimulus duration (30 ms in Condition 3) and a longer stimulus duration (125 ms in Condition 6). If the absolute magnitude of the ISI is important (presumably on ISI trials), then we should see similar accuracies for these critical pairs of trials, even though the ISI constitutes a larger or smaller proportion of a trial's total duration.

Method

Animals. The same four white Carneau pigeons used in Experiment 1 were used in Experiment 2.

Apparatus. The same four operant chambers used in Experiment 1 were used in Experiment 2.

Stimuli. Stimuli were generated and presented in the same manner as in Experiment 1, but with different stimulus and ISI durations, depending on the condition (see procedure and Table 1).

Procedure. The experiment was conducted over six blocks of 10 days each (60 days total). Daily sessions consisted of 120 trials, identical to the trials in Experiment 1, with the exception of different stimulus durations and different ISIs (on trials with an ISI) depending on condition.

Conditions. The six conditions were defined by the durations of stimulus presentations and ISIs (if present) during a trial (see Table 1). Note that Conditions 1 and 4 replicate conditions from Experiment 1.

Results

Factors influencing accuracy. A 2 (ISI: present, absent) × 5 (repetitions: 1, 2, 4, 8, 16) × 2 (relative duration: ISI > Stimulus Duration, Stimulus Duration > ISI) × 3 (total duration: 45 ms, 90 ms, 185 ms) × 10 (day: 1-10) repeated-measures ANOVA was computed on change detection accuracy. The main effect of day was not significant, $F(9, 27) = 1.06$, $p = .423$, partial $\eta^2 = .261$, indicating that performance was

relatively stable across the days that constituted each condition. All interactions involving day were not significant ($F < 1.54$, $p > .111$), with three exceptions: day × total duration, $F(18, 54) = 1.80$, $p = .050$, partial $\eta^2 = .375$; day × relative duration, $F(9, 27) = 2.63$, $p = .025$, partial $\eta^2 = .467$; and day × total duration × relative duration, $F(18, 54) = 2.10$, $p = .019$, partial $\eta^2 = .412$. Thus, there may have been some subtle changes during the 10 days that constituted each condition. Nevertheless, the effects of the remaining manipulated variables (ISI, repetitions, relative duration, and total duration) were consistent and can be seen in Figure 3.

The main effect of ISI was significant, $F(1, 3) = 39.47$, $p = .008$, partial $\eta^2 = .929$, indicating that accuracy was better on trials that did not have an ISI ($M = 64.10\%$) than on trials with an ISI ($M = 48.70\%$). This result is a replication of the basic change-blindness effect seen in previous implementations of the flicker task and in Experiment 1. The main effect of repetitions was also significant, $F(4, 12) = 27.56$, $p < .001$, partial $\eta^2 = .902$, and Figure 3 shows that as in Experiment 1 accuracy generally increased with additional repetitions (from 43.17% with one repetition to 65.14% with 16 repetitions). The main effect of relative duration was significant, $F(1, 3) = 99.95$, $p = .002$, partial $\eta^2 = .971$, with overall accuracy higher on trials when stimulus duration was longer than the ISI (panels on the right of Fig. 3) than when stimulus duration was shorter than the ISI (panels on the left of Fig. 3): 63.901% versus 48.89%. The main effect of total duration was not significant, $F(2, 6) = 0.75$, $p = .512$, partial $\eta^2 = .200$, indicating that longer cumulative time per repetition did not increase accuracy. That is, there was no overall difference between the top, middle, and bottom rows in Figure 3 (56.05%, 58.14%, and 55.01%, respectively). These last two results imply that the relative durations of the stimuli and ISI are more important than the overall duration of each repetition.

Though not pertinent to any specific hypotheses, several interactions were also obtained, and their effects can be seen in Figure 3: ISI × repetitions, $F(4, 12) = 13.72$, $p < .001$, partial $\eta^2 = .821$; ISI × relative duration, $F(1, 3) = 48.27$, $p = .006$, partial $\eta^2 = .941$; ISI × total duration, $F(2, 6) = 51.67$, $p < .001$, partial $\eta^2 = .945$; repetitions × relative duration, $F(4, 12) = 5.65$, $p = .009$, partial $\eta^2 = .653$; relative duration ×
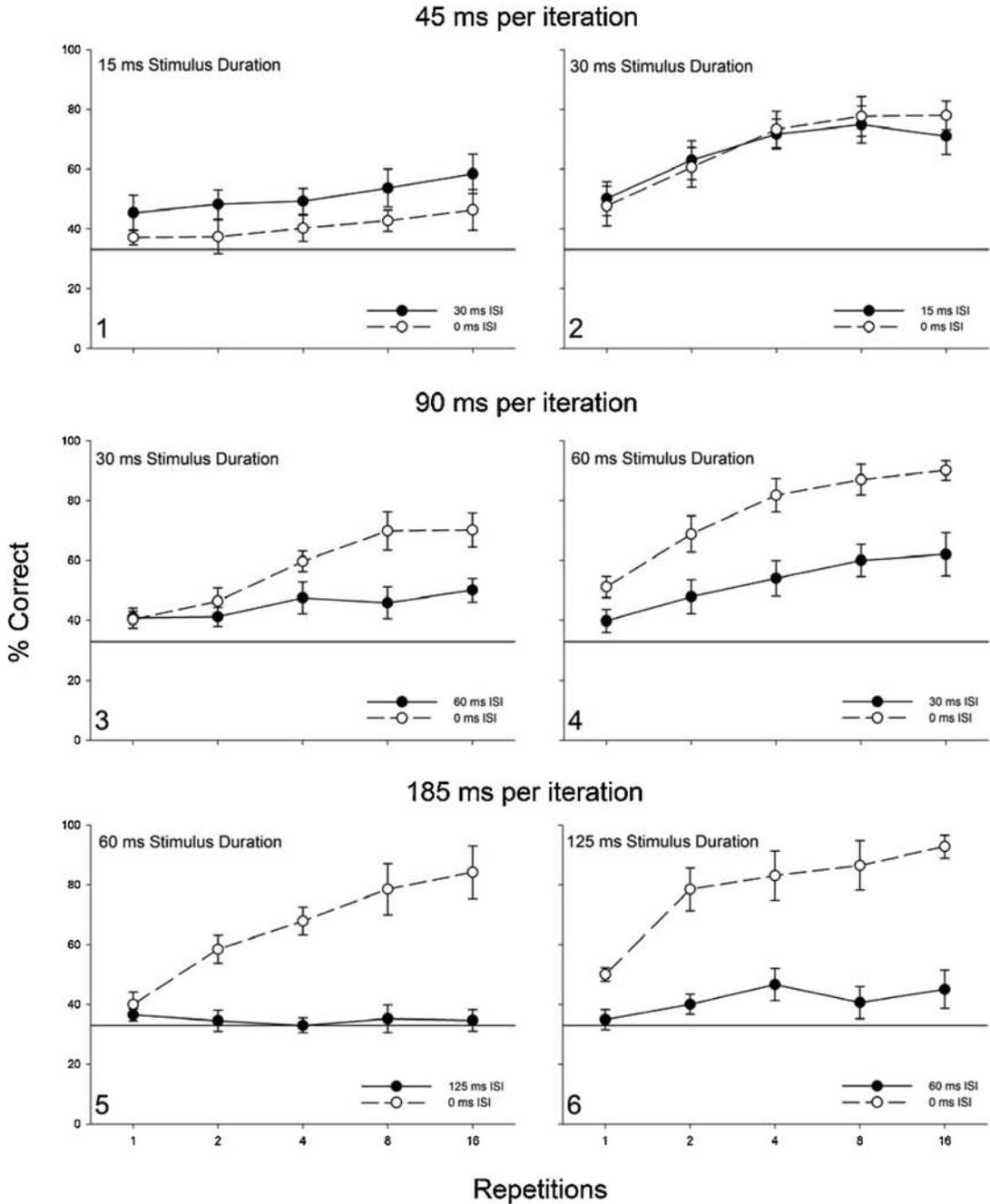
Fig. 3.   Performance on ISI (solid lines) and no-ISI (dashed lines) trials during the six conditions of Experiment 2 (indicated by the index number in the lower left of each panel). Note that side-by-side panels represent conditions with the same total duration (display plus ISI) per repetition of an ISI trial. Panels on the left represent conditions in which the ISI was longer than the display duration. Panels on the right represent conditions in which the ISI was shorter than the display duration. Error bars span ± 1 standard error.

total duration, $F(2, 6) = 16.68$, $p = .004$, partial $\eta^2 = .848$; and ISI × repetitions × total duration, $F(8, 24) = 13.07$, $p < .001$, partial $\eta^2 = .813$. The remaining interactions were not significant, $F < 2.33$, $p > .052$.

Allocation of search time to display and ISI durations. One of the primary questions leading to this experiment was whether allocation of time on ISI trials (to either stimulus durations or ISI durations) would affect accuracy. Figure 3 shows that performance at each total display duration (45 ms, 90 ms, and 185 ms) was generally better when the stimulus duration was longer than the ISI duration (dashed lines in the right panels of Fig. 3 reflect better performance than dashed lines in the left panels), and this is partially supported by post-hoc *t*-tests. To control for type I error, the alpha criterion was adjusted to .017 for each of the three comparisons, to yield a total alpha of .05. In Conditions 1 and 2 (45 ms per iteration), accuracy with an ISI was significantly higher when that ISI was shorter ($M = 66.12$, $SD = 11.25$) than when it was longer ($M = 50.34$, $SD = 10.67$), $t(3) = 17.450$, $p < .001$. In Conditions 3 and 4 (90 ms per iteration), accuracy was numerically higher when the ISI was shorter ($M = 54.60$, $SD = 10.68$) than when it was longer ($M = 44.50$, $SD = 7.554$), but not significantly so, $t(3) = 4.56$, $p = .020$. Similarly, in Conditions 5 and 6 (185 ms per iteration), accuracy was not significantly higher when the ISI was shorter ($M = 41.40$, $SD = 6.72$) than when it was longer ($M = 34.73$, $SD = 2.71$), $t(3) = 2.18$, $p = .117$.

Relative duration of stimulus presentations and ISIs. Recall that there were several paired conditions intended to reveal the importance of the relative durations of the stimulus presentations and ISIs. In particular, no-ISI trials with 30-ms stimulus durations appeared in the context of trials with both shorter (15-ms) and longer (60-ms) ISIs, and no-ISI trials with 60-ms stimulus durations appeared in the context of trials with both shorter (30-ms) and longer (125-ms) ISIs. A close inspection of those elements of Figure 3 (dashed lines in panels 2 and 3, and in panels 4 and 5, respectively) reveals that accuracy may be slightly better when a given stimulus duration is longer than the ISI duration. To test this possibility, a follow-up 2 (stimulus duration: 30 ms, 60 ms) × 2 (relative duration: ISI longer, ISI shorter) repeated measures ANOVA was run. Means from these pairs of conditions appear in the top panel of

Figure 4 (collapsed across repetitions to show only the effect of display timing). There was no main effect of stimulus duration: 61.47% versus 70.77% for the 30- and 60-ms conditions respectively, $F(1, 3) = 7.37$, $p = .073$, partial $\eta^2 = .711$. There was a significant main effect of relative duration: 61.47% versus 71.59% for conditions with longer ISIs and shorter ISIs respectively, $F(1, 3) = 80.69$, $p = .003$, partial $\eta^2 = .964$. There was no interaction between stimulus duration and relative duration, $F(1, 3) = 0.01$, $p = .944$, partial $\eta^2 = .002$. This confirms that performance did not vary as a consequence of absolute stimulus duration, but rather as a consequence of relative stimulus duration (in comparison to the ISI duration).

Similarly, 30-ms ISI trials appeared in the context of both shorter (15-ms) and longer (60-ms) stimulus durations, and 60-ms ISI trials appeared in the context of both shorter (30-ms) and longer (125-ms) stimulus durations. A follow-up 2 (ISI duration: 30 ms, 60 ms) × 2 (relative duration: ISI longer, ISI shorter) repeated measures ANOVA was run. Means from these pairs of conditions can be seen in the bottom panel of Figure 4. There was a main effect of ISI duration indicating higher accuracy on 30-ms ISI trials (52.48%) than on 60-ms ISI trials (43.20%), $F(1,3) = 16.61$, $p = .027$, partial $\eta^2 = .847$. There was no main effect of relative duration, $F(1, 3) = 0.10$, partial $\eta^2 = .769$, and there was no interaction between ISI duration and relative duration, $F(1, 3) = 2.96$, $p = .184$, partial $\eta^2 = .497$. In this case, accuracy on ISI trials was determined exclusively by the absolute duration of the ISI, and its relation to the stimulus duration had no effect.

Discussion

This experiment was designed to determine if change detection is affected by the relative durations of stimulus presentations and ISIs, motivated by the observation of a reversed change blindness effect in the only condition of Experiment 1 that featured an ISI that was longer than the stimulus presentation time. Although no other conditions produced the reversed change blindness effect seen in Experiment 1, the relative durations of the stimulus components did have a systematic effect on performance: Change detection was worse when the ISIs were longer than the
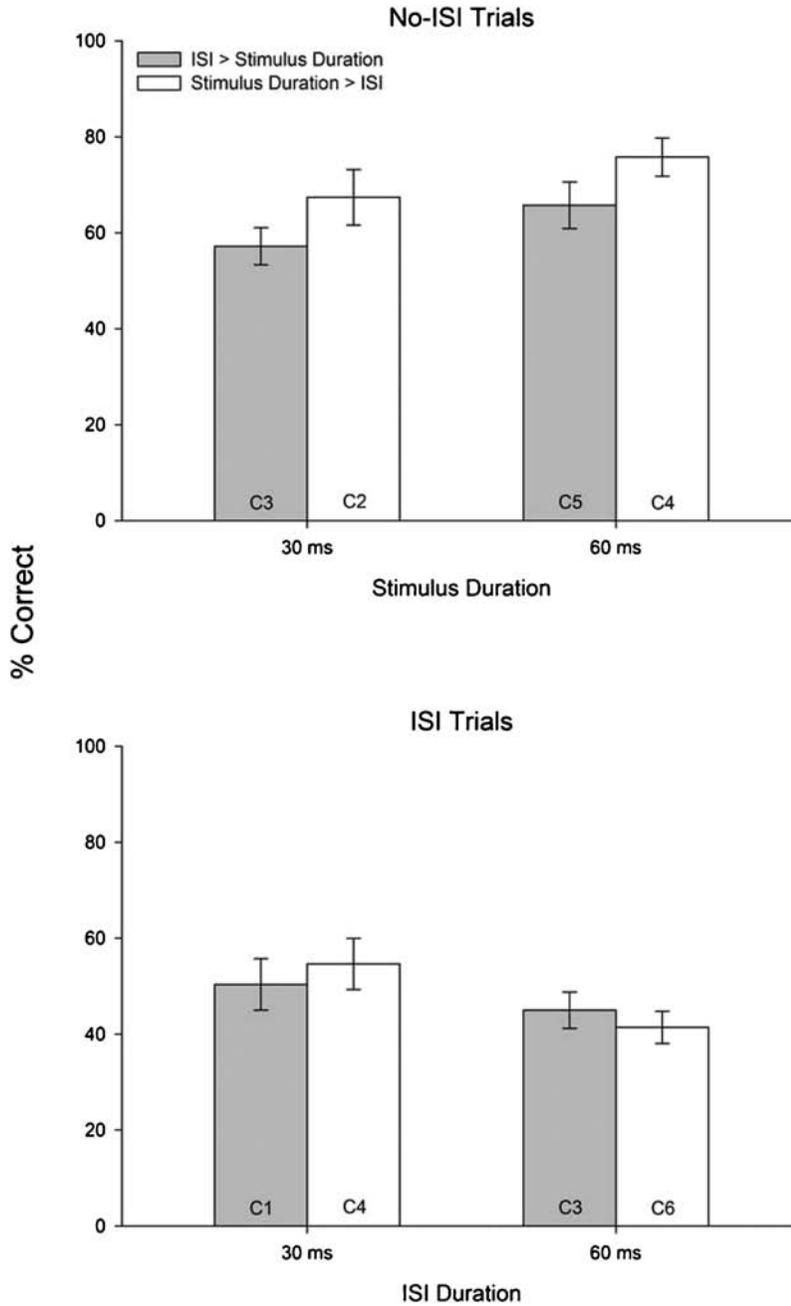
Fig. 4. Comparison of parallel trial types across conditions in Experiment 2. Top panel: ISI trials in different conditions having the same ISI duration, and either longer than or shorter than the stimulus display duration. Bottom panel: no-ISI trials in different conditions having the same display duration, and either longer than or shorter than the ISI duration. Error bars span ± 1 standard error. C1, C2, C3, C4, C5, and C6 refer to conditions 1-6 as defined in Table 1.

intervening stimulus presentations. Furthermore, the two types of trials in the change detection task (ISI trials and no-ISI trials) were influenced by different factors.

Accuracy on trials with an ISI was enhanced by repetitions and by shorter overall time per repetition. The latter effect, however, was entirely due to the ISI component of a

presentation: When the same ISI durations were presented in the context of different overall durations, there was no difference in accuracy. Instead, accurate change detection seemed to rely on the time interval across which two consecutive displays needed to be compared: Longer ISIs produced a greater impairment. Thus, although the main effect of overall time initially seems counterintuitive (with less search time paradoxically producing better accuracy), the effect is merely a consequence of trials with shorter ISIs by definition providing less time before a display terminates.

No-ISI trials, like ISI trials, were also influenced by repetitions and by overall time per iteration (although in this case, shorter durations were associated with worse accuracy). Interestingly, the effect of total duration was influenced by ISI duration: When no-ISI trials having the same stimulus durations were presented in the context of conditions having different ISI durations, performance was better when the stimulus durations were longer than the ISI duration. This is surprising, given that no-ISI trials did not themselves feature an ISI! Presumably there was some within-session, intertrial effect of ISI trials on no-ISI trials that influenced pigeons' search strategies or perception. More important than the absolute stimulus duration (which did not have a significant effect on performance) was its duration relative to the ISI.

General Discussion

These experiments add to a growing body of evidence that change detection in pigeons parallels human change detection in many ways: Both are influenced, and in roughly the same way, by the presence of an ISI (the basic operational definition of change blindness), stimulus repetitions, change salience, and ISI duration (Herbranson et al., 2014, Herbranson, 2015). Experiment 1 initially pointed toward a possible interspecies difference by showing that pigeons (unlike humans) are influenced by stimulus duration; Experiment 2 more precisely defined the nature of that difference. In particular, the effect of stimulus presentation times can be largely, if not completely accounted for by overall search time: Longer stimulus presentation times yield longer trials with more time to identify changes.

The lack of a display duration effect in human change blindness experiments (e.g., Pashler, 1988) is presumably due to slightly different procedures. The methods employed in human versions of the flicker task usually involve presentation of the alternating stimuli until the change is detected. The pigeon methods reported here instead presented alternating stimuli for a fixed number of iterations. This was done to force pigeons to see a specific number of repetitions, thus facilitating analysis of the repetition factor and minimizing the effects of random, impulsive responses early in a trial (presumably before a change was detected). Thus, human results more clearly show the effects of the powerful search time factor (in the form of response times), whereas the pigeon results confound search time with the number of forced iterations during a given trial.

The importance of search time matches the conclusions of previous studies and is consistent with the notion that pigeons (like humans) engage in a location-by-location search during the repeating stimulus display. Several components of a trial increase search time and can consequently increase accuracy, including repetitions, stimulus duration, and even ISI duration. The last of these, however, is importantly different from the others because inclusion of an ISI also interrupts the transition between consecutive stimulus displays, thereby impairing performance. In most cases, that impairment is strong enough to overcome the benefit of the additional search time.

An important implication of such a location-by-location search is that one might fail to detect a change for different reasons. First, the change might be obscured by some aspect of the stimulus display (such as by the presence of an ISI). Alternatively, change detection might fail because the location of the change was never observed in the first place. Both of these likely occurred in the reported experiments. However, we argue that the former is the important contributor to change blindness, as operationally defined here (specifically, as the difference in accuracy between trials with and without ISIs). The latter possibility, although it likely led to some instances of incorrect trials (especially trials with short total durations), ought to apply equally to both ISI trials and no-ISI trials, and to produce no net difference in performance (if anything, the additional inter-key travel time

made available by the ISI should allow for observation of more, not fewer locations).

One of the surprising findings here was the discovery of a reversed change blindness effect in one condition (the final condition of Experiment 1), in which pigeons produced better accuracy on ISI trials than on no-ISI trials. Apparently, in some circumstances (presumably those involving very fast presentations), the additional time provided by an ISI can overcome the negative effect of ISI presence and yield a net increase in accuracy. It is important to consider at this point how an ISI might have such an effect on change detection.

Because neither stimulus display is actually visible during the ISI, it does not initially seem as though an ISI could possibly enhance accuracy at detecting changes to those (non-visible) displays. However, consider two possible factors. First, given the spacing between keys, it would have been difficult for pigeons to observe more than one key at a time, and virtually impossible to view the entire three-key display at once. Consequently, pigeons would need to physically move to observe all of the possible change locations. An ISI could provide more time to travel from key to key without missing repetitions that could be used to identify the change. Although responses were not recorded during stimulus presentation, some pigeons tended to peck response keys while the displays were present. Thus, future research might benefit from recording those responses and their timing as indicators of inter-key travel. Alternatively, Herbranson et al. (2014) and Herbranson (2015) analyzed accuracy as a function of key preference and repetitions to infer this kind of serial search behavior, even in those birds that did not peck the ongoing stimulus displays.

Second, note that there was no masking stimulus presented during the ISI, as is sometimes done in change detection research. Consequently, it is possible that a sensory memory trace made stimulus displays persist briefly during the ISI. If so, then under some circumstances, an ISI would have the potential to increase accuracy in the same manner as longer stimulus durations do. In general, however, the impairment associated with ISI presence appears to be much more powerful than the possible benefit of additional search time that it might provide.

The crux of the flicker task is the difference between ISI and no-ISI trials, with the presence of an ISI increasing considerably the difficulty of change detection. On ISI trials, pigeons had to compare stimulus displays that were separated in time by the ISI, a feat requiring some form of memory. Given the methods and results of these experiments, it seems likely that pigeons were using sensory memory to span the ISI and compare stimuli. Although pigeons have proven capable of utilizing short-term memory (STM; or working memory) in similar kinds of change detection tasks (Gibson et al., 2011; Wright et al., 2010), some aspects of the current procedure and results would seem to better fit the characteristics of sensory memory than short-term memory (see Phillips, 1974).

First, sensory memory is considered to be a high capacity store, whereas STM has strict limitations on capacity. The stimulus displays in these experiments (consisting of as many as 24 individual features) would pose a tremendous challenge for the stringent limits on STM, but not sensory memory. Second, the results of trials with ISIs of different durations show that performance quickly degrades with ISI length, eventually reaching chance at the longest value tested in Experiment 2 (125 ms). This also fits closely the short duration of sensory memory, but not the relatively more durable characteristics of STM, which could presumably span intervals of several seconds or more. If pigeons were using STM, then performance could be easily maintained well beyond 125 ms (Shimp & Moffitt, 1977; Diekamp, Kalt, & Güntürkün, 2002).

Finally, although change blindness may be initially surprising and counterintuitive, the fact that it can be reliably produced in multiple species and is influenced by a host of variables, suggest that it may be an important consequence of a general cognitive process such as selective attention (Zentall, 2005).

Herein lie some of the primary virtues of the comparative approach. Although theories (such as selective attention) and methods (such as the flicker task) imported from cognitive psychology might or might not yield the expected results in nonhuman animals, they do motivate novel research that might not otherwise ever be pursued (Zentall, 2013). With modern technology (such as smart phones, tablet computers, text messaging, Twitter, and the like) seemingly stretching human attention toward its limits, a deep understanding of the limitations of

attention (such as change blindness and inattentional blindness) are essential if we are to effectively cope with those limitations. Comparative psychology in particular, is in a unique position to contribute not just to an understanding of the proximate causal factors that contribute to change blindness, but also the ecological relevance and ultimate causes of change blindness. The last of these may be particularly important, given much of psychology's disturbing reluctance to embrace evolutionary theory (Mesoudi, Veldhuis, & Foley, 2010). Although there remain some points of contention between modern comparative psychologists and behavior analysts, one would hope that both can agree on the relevance of animal behavior to human psychology, and on the need for a strong empirical framework, grounded in the natural sciences (a framework that includes a central role for Darwinian evolution).

References

Blough, D. S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior*, *2*(2), 151–160.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*(2), 149–178.

Cook, R.G., Kelly, D.M. & Katz, J.S. (2003). Successive two-item same-different discrimination and concept learning by pigeons. *Behavioural Processes*, *62*, 125–144.

Diekamp, B., Kalt, T., & Güntürkün, O. (2002). Working memory neurons in pigeons. *The Journal of Neuroscience*, *22*, 1–5.

Gibbon, J., Baldock, M., Locurto, C., Gold, L., & Terrace, H.S. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(3), 264–284.

Gibson, B. M., Wasserman, E. A., & Luck, S. J. (2011). Qualitative similarities in the visual working memory of pigeons and people. *Psychonomic Bulletin & Review*, *18*, 979–984.

Green, D.M., & Swets J.A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception (Vancouver studies in cognitive science)* (pp. 89–110). New York: Oxford University Press.

Herbranson, W.T. (2015). Change blindness in pigeons (*Columba livia*): The effects of change salience and timing. *Frontiers in Psychology*, *6*, 1109. Doi: 10.3389/psyg.2015.01109.

Herbranson, W. T., & Stanton, G. L. (2011). Flexible serial response learning by pigeons (*Columba livia*) and humans (*Homo sapiens*). *Journal of Comparative Psychology*, *125*(3), 328.

Herbranson, W. T., Trinh, Y. T., Xi, P. M., Arand, M. P., Barker, M. S. K., & Pratt, T. H. (2014). Change detection and change blindness in pigeons (*Columba livia*). *Journal of Comparative Psychology*, *128*(2), 181–187.

Lazareva, O. F., & Wasserman, E. A. (2015) No evidence for feature binding by pigeons in a change detection task. *Behavioural Processes*. Advance online publication. doi: 10.1016/j.beproc.2015.09.007.

Levin, D. T., Momen, N., Drivdahl IV, S. B., & Simons, D. J. (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*, *7*(1–3), 397–412.

Mesoudi, A., Veldhuis, D., & Foley, R.A. (2010). Why aren't the social sciences Darwinian? *Journal of Evolutionary Psychology*, *8*(2), 93–104.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*, 369–378.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290.

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*(3), 423.

Rensink, R. A., O'Regan, J. K., and Clark, J. J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.

Roitblat, H. L. (1980). Codes and coding processes in pigeon short-term memory. *Animal Learning & Behavior*, *8*, 341–351.

Shimp, C. P. (1969). The concurrent reinforcement of two interresponse times: the relative frequency of an interresponse time equals its relative harmonic length. *Journal of the Experimental Analysis of Behavior*, *12*(3), 403–411.

Shimp, C. P., & Moffitt, M. (1977). Short-term memory in the pigeon: delayed-pair-comparison procedures and some results. *Journal of the Experimental Analysis of Behavior*, *28*(1), 13–25.

Simons, D. J., & Ambinder, M. S. (2005). Change blindness theory and consequences. *Current Directions in Psychological Science*, *14*(1), 44–48.

Stagner, J. P., Michler, D. M., Rayburn-Reeves, R. M., Laude, J. R., & Zentall, T. R. (2013). Midsession reversal learning: why do pigeons anticipate and perseverate? *Learning & Behavior*, *41*(1), 54–60.

Wasserman, E. A., Kiedinger, R. E., & Bhatt, R. S. (1988). Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*(3), 235.

Wright, A.A., Katz, J.A., Magnotti, J., Elmore, L.C., Babb, S., & Alwin, S. (2010). Testing pigeon memory in a change detection task. *Psychonomic Bulletin & Review*, *17*, 243–249.

Zentall, T. R. (2005). Selective and divided attention in animals. *Behavioural Processes*, *69*(1), 1–15.

Zentall, T.R. (2013). Comparative cognition: An approach whose time has come. *Journal of the Experimental Analysis of Behavior*, *100*, 257–268.