Sensory and working memory in a spatial change-detection task by pigeons and humans

Kenneth Leising[a,*], Justin Jacqmain[a], Cheyenne Elliott[a], Joshua Wolf[b], James Taylor[c],
Lauren Cleland[a], Ronnie Lee[a], John F. Magnotti[d], Anthony A. Wright[e]

[a] Department of Psychology, Texas Christian University, Fort Worth, TX, United States
[b] Department of Psychology, Carroll University, Waukesha, WI, United States
[c] Department of Psychology, Utah Valley University, Orem, UT, United States
[d] Department of Neurosurgery, Baylor College of Medicine, Houston, TX, United States
[e] Department of Neurobiology and Anatomy, University of Texas Health Science Center, Houston, TX, United States

ARTICLE INFO

ABSTRACT

Judgements of items viewed less than 100 ms prior are predominantly supported by a sensory, or iconic, memory system. Iconic memory is of high-capacity, but is also volatile and limited in duration. Judgements after longer delays increasingly rely on a working memory system, which is lower in capacity and volatility than sensory memory, but is longer in duration. In four experiments, several factors (e.g., length of delay, number of items, time to view items, presence of a visual mask) were manipulated during a spatial change-detection task conducted with humans and pigeons. Both species were exposed to trials with an array of colored circles (2, 3, and 4 circles in Experiment 1 and 2a; 4, 6, and 8 circles in Experiment 2b) followed by a brief delay (0, 50, and 100 ms in Experiment 1a; 0, 100, and 1000 ms in Experiments 1b and 2), and then were presented with a test display in which the position of one of the items had changed. Pigeons, like humans, were less accurate in selecting the changed item with more items in the display and after longer delays. Pigeons were equally accurate on trials with 0 and 100-ms delays, but worse on trials with a 1000-ms delay; whereas, humans were equally accurate on 100-ms and 1000-ms delays, but better on 0-ms delay trials. Accurate change detection was disrupted in both species when a visual mask was inserted between the sample and test display after a short (100), but not a long (1000 ms) delay. The results support similarity between species in the functional relationships between delay and memory systems, despite time course differences related to sensory memory.

1. Sensory and working memory in a spatial change-detection task with pigeons and humans

Detecting a moving object has great survival value. To do so an animal may attend to areas of the visual field in which the pattern of light falling on the retina (retinotopic pattern) has changed from one moment to the next. Alternatively, the nervous system of the animal may be organized in such a way as to "store" the perceptual experience thereby enabling a kind of comparison of the current perceptual experience to one that occurred a moment ago. The former mechanism is referred to as iconic memory and the latter as working memory. Iconic memory is sustained activation within the areas of the brain that detected the stimulation. This sensory, or iconic (in the case of visual stimulation) memory can be thought of as a brief afterimage (but see Coltheart, 1980), containing the retinotopic coordinates of all of the

visual stimulation within the visual field at a given moment (i.e., the pattern of light as it fell on the retina). In monkeys and humans, current perceptual experience rapidly replaces past experience in iconic memory (< 1000 ms), rendering the prior memory inaccessible (Phillips, 1974; Truppa et al., 2014). However, coordinated activity within the prefrontal cortex, parietal cortex, and sensory cortex (and other recruited areas) can enable the perceptual experience to remain accessible in a neural state referred to as working memory (Christophel et al., 2017; Eriksson et al., 2015).

Iconic and working memory have been compared based on how response accuracy is influenced by the number of items present (capacity), the delay between successive displays or a display and the opportunity to respond (duration), and interference by task-irrelevant stimulation (volatility). Iconic memory has a higher capacity than working memory because it includes not just the items attended to, but

all of the items in the visual field (Öğmen, and Herzog, 2016). However, the eyes make frequent and rapid movements causing the retinotopic iconic memory to be updated (e.g., every 500 ms). Consequently, the accuracy of judgments based on iconic memory (e.g., partial report) is higher than working memory when they occur shortly after viewing the items (Wheeler and Treisman, 2002; Pashler, 1988; Phillips, 1974; Sperling, 1960). These data largely come from experiments with humans, though comparable data exist with monkeys (Keysers et al., 2005; Truppa et al., 2014), and to a lesser extent, pigeons (White, 1974). Any manipulation (e.g., a movement or masking procedure) which results in a mismatch between the retinotopic pattern (of light) from a moment ago and the current pattern (i.e., blurring) would reduce the accuracy of judgements based on iconic memory, but have little effect on judgements of working memory (Bradley and Pearson, 2012; Öğmen, and Herzog, 2016; Phillips, 1974).

A common procedure for comparing iconic and working memory is to ask subjects to detect a change among one or more items (e.g., colored circles) within their visual field. At a minimum, these change-detection procedures include two displays of items presented across a delay (e.g., 50 ms), after which one or more of the items from the first display have been altered (e.g., by changing an item's shape or color). The task may involve responses to indicate a change or no change within the second display, or a response directly to the item that has changed. In the most iconic (pun intended) study of both memory systems, Phillips (1974) presented humans with a 12 × 12 array with a subset of the cells illuminated on each trial. After a delay, on some trials the number of lit cells would increase or decrease by one. The inter-stimulus intervals (ISI) between the presentation of the sample and test displays varied between 20, 100, 3000, and 9000 ms. The array size varied between four, six, and eight illuminated cells, with the same number (plus or minus one on the change trials) on both displays. It was expected that larger arrays would result in poorer accuracy in detecting a change, except when relying on iconic memory. Accuracy was similar across arrays of four, six, and eight items when the change occurred after a 20-ms delay (i.e., when dependent on iconic memory), but as expected, accuracy was worse with eight items compared to six and four items with longer delays (e.g., 1000 ms). The capacity of iconic memory is greater, but it should also be more volatile (i.e., subject to interference). In the next experiment, Phillips manipulated volatility by shifting the position of a 5-item or 8-item configuration of lit cells (plus or minus one on change trials) within the 12 × 12 array. This movement caused the greatest disruption of change detection at the shortest delays (20, 60, and 100 ms). Furthermore, the effect of delay interacted with array size, resulting in a greater decrement with the 8-item than 4-item array at the shortest delays. A more direct method of altering the contents of iconic memory is to present a visual mask immediately after viewing the sample stimulus. In Experiment 3, Phillips inserted a visual mask during the delay and reported greater decrements at the shorter delays (40, 80, and 120 ms). Reaction times were faster on correct change trials with movement (Experiment 2) and masking (Experiment 3). More recently, Vogel et al. (2006) used colored squares and reported masking effects at delays between 117 ms and 584 ms. Consistent with Phillips, the size of the masking effect increased with larger array sizes. These data, among those of others (Carter and Henning, 1971; de Valois and Switkes, 1983; Henning et al., 1981; Jitsumori and Ushitani, 2017; Legge and Foley, 1980; Losada and Mullen, 1995; Sekuler, 1965; Solomon, 2000; Stromeyer and Julesz, 1972; Wilson et al., 1983), indicate that masking effects are strongest after short delays (< 300 ms) when items are likely to be active in iconic memory.

Two other manipulations, the duration of training and the duration of the displayed items, have been examined with visual change-detection procedures. Visuospatial working memory is hypothesized to be a relatively stable system, both across time and with respect to performance across a variety of tasks. Consistent with this hypothesis, studies have reported little or no improvement in the accuracy of change detection with practice (e.g., Eng et al., 2005; Oliva et al., 2004; Owen

et al., 2010; Wolfe et al., 2002; but see also Morrison and Chein, 2011). In Eng et al. (2005) none of the participants' accuracies on a change-detection task significantly improved despite participants being presented with 5400 trials across 10 sessions. The duration of the sample display is another parameter that has been varied within a change-detection procedure. Pashler (1988) varied the duration of the sample display prior to a 67-ms delay between the sample and test display in a visual change-detection task with humans. Performance improved as the duration of the sample display increased from 100 to 500 ms. Similar evidence for the effect of increased sample duration was obtained with pigeons in a match to sample procedure (Grant, 1976) and change detection (Herbranson and Davis, 2016), but this effect has not been evaluated in a spatial change-detection procedure. Contrary to these findings, Coltheart (1980) concluded that no relationship exists between stimulus duration and the accuracy of judgments based on iconic memory.

We are not aware of any studies utilizing change detection to compare iconic memory between species, though many studies have compared other aspects of working memory. For most vertebrates, the success or failure of predator avoidance, feeding, mate selection, and parental behavior depends on their ability to detect and respond to changes in the visual properties of their environment (e.g., an egg removed from the nest). In particular, birds have sophisticated color perception and are hypothesized to be the most visually dependent of all vertebrate classes (Hodos et al., 2003). Utilizing procedures other than change detection, data from monkeys and pigeons indicate that delays of less than 1000 ms result in different behavioral and neural effects than longer delays. Truppa et al. (2014) measured accuracy and response time during a delayed matching-to-sample procedure with capuchin monkeys. They found no difference in accuracy or response time on trials with a 0 and 500 ms delay, whereas, accuracy decreased and response time increased with longer delays (1000, 2000, and 3000 ms). Keysers et al. (2005) displayed a sample target image to human participants for a fixed duration and then instructed the participants to identify whether the target image was present or absent from a subsequent series of images. On some trials the sample target image was displayed for 18 ms followed by a 75 ms gap and then the series of images (gap trials). On other trials, the target was displayed for 18 ms (short-nogap) or 93 s (long-nogap) and followed immediately by the series of images. Accuracy was the same on gap and long-nogap trials, which was better than on short-nogap trials. The researchers then recorded from neurons in the temporal cortex of macaque monkeys and in those neurons reacting to the target image they found sustained activity for the 75 ms gap following the 18 ms display of the sample target. These data support a kind of iconic memory in non-human animals that persists for at least 93 ms and as long as 500 ms.

Little research has been conducted with pigeons at intervals less than 1000 ms. White (1974) tested pigeons in a successive same-different task and found comparable levels of accuracy after 100, 500, and 1000 ms delays, but not with 2000 ms delays or longer. Herbranson and Davis (2016) used a flicker paradigm, which alternates between the sample and test displays. With only one cycle, the paradigm is similar to a typical change detection task. However, they found with pigeons that the effect of display duration (30, 60, and 125 ms) on performance differed when the delay was 0 ms compared to 30 ms, suggesting iconic memory in pigeons may be limited to even shorter durations than humans. In general, these data suggest similar memory mechanisms with comparable time courses in humans, monkeys, and pigeons.

Direct comparisons between species support qualitative similarity in how visual information is utilized by the visual working memory system. For example, Gibson et al. (2011) trained pigeons and humans on a visual change-detection task with an array of eight items. Across trials they varied the number of items within the array that changed (from zero to eight). Elmore et al. (2011) compared monkeys and humans in a change-detection task in which only one item changed but the number of unchanging items varied (from one to five). In both

experiments, accuracy was better when the ratio of changed to non-changed items in the array was higher, although humans were overall more accurate at detecting changes. In each experiment, qualitative similarity in the underlying mechanism was supported by good fits between data from both species and the same working-memory model. These studies revealed qualitative similarities in working memory and offered support for the feasibility of cross-species comparisons.

The majority of research on change detection, including the literature reviewed thus far, has involved detecting a change across a delay in an item's color, shape, or orientation. Baddeley (2003) distinguished between memory processes engaged by these visual features of an item (also see Courtney et al., 1996; Jonides et al., 2008 and Jiang et al., 2008 for reviews) and the item's spatial location. This distinction has also been supported by research with non-human animals (e.g., Kesner et al., 1996; Leising et al., 2013). Leising et al. (2013) trained and tested two pigeons and a rhesus monkey with a spatial change-detection procedure. The position of an item within the display changed from the sample to the test display. Both species achieved above 70% percent correct in responding to the item that changed position. Both species were also presented with trials in which the item did not change position, but instead changed color. One pigeon performed above chance on these color-change trials, whereas the other pigeon and monkey did not. Elmore et al. (2012) trained pigeons successfully on a visual change-detection task (i.e., color change) and found no transfer to the detection of location change during subsequent testing. These results suggested that detection of one change (e.g., location) does not reliably transfer to detection of other changes (e.g., color), further supporting a division between visual and spatial working memory. These findings also validate the use of the spatial change-detection procedure to study other properties of spatial working memory.

In the four experiments reported here, pigeons and humans were trained on the spatial change-detection task first reported by Leising et al. (2013). Pigeons (Experiments 1a & 1b) and humans (Experiments 2a & 2b) were trained to detect and respond to an item that was unchanged in terms of visual features (e.g., color, shape, and size), but was spatially repositioned from the sample to test display. Three manipulations known to distinguish the iconic and working memory systems (the duration of the retention delay, number of items in the display, and presence of a visual mask) were evaluated for their effects on percent correct detection of the changed item and latency to correctly select the changed item. In Experiment 1a, pigeons were tested with variations in the duration of the sample display, delay between items, and number of display items. The aim of Experiment 1b was to determine whether the addition of a visual mask would have the same effect on change detection performance following short versus long delays between sample and test displays. Experiment 2a used the same parameters in a spatial change-detection task with humans. In Experiment 2b, the number of items in the array was increased. Based on prior reports (e.g., Elmore et al., 2011; Gibson et al., 2011), we predicted spatial change-detection performance would reveal quantitative differences, but qualitative similarities, in iconic and working memory between species.

2. Experiment 1

Experiments 1a and 1b were conducted with pigeons completing a spatial change-detection procedure. Experiment 1a aimed to investigate the effects of factors known to influence change detection performance in humans. In Experiment 1b, a visual mask was inserted on some trials to determine the effect of the mask on spatial change detection across different durations of the delay between sample and test displays.

3. Experiment 1a

Perhaps the two most common manipulations in memory research involve the delay following exposure to a target item and variations in the number of to-be-remembered items. In Leising et al. (2013), the

delay and number of items in the display were held constant during training, whereas, both factors were varied in Experiment 1a. In addition, the duration of the sample stimulus (viewing time) was also varied. Previous research has reported modest, but reliable effects on change-detection performance depending on how long the sample display was viewed, with moderately better performance for longer view times (Grant, 1976; Pashler, 1988). There is current debate among cognitive psychologists regarding whether working memory improves with training (Morrison and Chein, 2011). In Experiment 1a, pigeons were trained with different numbers of items in the array (2, 3, or 4 items), durations of the sample (1000, 2000, or 3000 ms), and delays between sample and test displays (0, 50, or 100 ms) in each session of training. In addition to analyses of each main effect, training was included as a factor to determine whether performance improved across sessions.

3.1. Method

3.1.1. Subjects

Six pigeons (*Columba livia*), ranging in age from 1 to 2 years old, served as subjects. All subjects were acquired from Double T Farm (Glenwood, IA). Pigeons were tested 5–6 days per week. They were individually housed in a colony with a 12-hr light-dark cycle and had free access to water and grit. Pigeons were maintained at 83–85% of their free-feeding weights. Experimental procedures occurred during the light portion of the cycle. All animal procedures were approved by the Institutional Animal Care and Use Committee at Texas Christian University (TCU). All pigeons had previously participated in a task that involved pecking at colored circles 6.5 cm in diameter and positioned in the center of the screen (Stahlman & Leising, 2015). All of the colors and the procedure used in the current task were novel to the pigeons.

3.1.2. Apparatus

Training and testing occurred in a flat-black Plexiglas chamber (45 cm × 41 cm × 46 cm). All items were presented by computer on a color LCD monitor (L1750, HP, Palo Alto, CA) visible through a 33 × 40 cm viewing window in the middle of the front panel of the chamber. The bottom edge of the viewing window was 12 cm above the chamber floor. Pecks to the monitor were detected by an infrared touch screen (EZ-170-WAVE, ezscreen, Houston, TX) mounted on the front panel. A 28-V houselight was located in the ceiling of the box. A food hopper (Coulbourn Instruments, Allentown, PA) was located in the center of the front panel, its access hole flush with the floor. All experimental events were controlled and recorded by a computer. A video card controlled the monitor in the SVGA graphics mode (800 × 600 pixels). All experimental events were coordinated via Microsoft Visual Basic 6.0 software.

3.1.2.1. Stimulus displays. The ready stimulus was a white, 2.5 cm diameter circle with a black plus sign in the center. The training items consisted of eight colored circles that were 2 cm in diameter. RGB 24-bit values for the items were: Red: 255, 0, 0; Aqua: 0, 255, 255; Yellow: 255, 255, 0; Purple: 180, 0, 255; Blue: 0, 0, 255; Green: 0, 255, 0; Magenta: 255, 0, 255; and Orange: 255, 128, 0.

The colors of items in the sample display were selected without replacement for each bird. There were 336 possible two-item color combinations (e.g., red and green circle on the display). When the number of sample items exceeded two, two of the items in the array were selected without replacement from the possible 336 color configurations. The colors of the remaining items were selected psuedorandomly, such that they could not match the color of another item. The items were displayed within a framework of two large (imaginary) concentric circles (see Fig. 1). The outer circle was 13.50 cm in diameter with 12 positions in which the items could appear spaced 3.53 cm apart (center to center). The inner circle was 6.5 cm in diameter and 4 cm inside the outer circle. The inner circle consisted of 8 positions
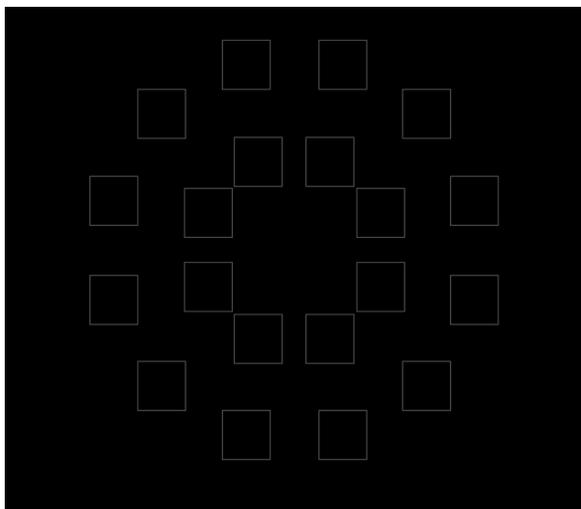
Fig. 1. A display of unfilled item positions. Twenty potential item positions were located along the circumference of two imaginary concentric circles. In the sample display the items could appear in any position, whereas during the test display the changed item could appear in any position except the two highest and lowest positions.

spaced 2.55 cm apart.

The distance of each position change from start to end point was calculated based on an 800×600 pixel display. The result was 23 possible distance-changes ranging from 98 pixels (from one position on the inner circle to the adjacent position) to 512 (from one position on the outer circle to its symmetrical opposite). There were 380 positional changes within the 23 distances. For programming purposes only, 7 of the distance changes were combined into 3 categories (e.g., 132 and 133 pixel distances were combined into one distance category) to create 19 total categories of distance changes (all 23 distances were still possible when building a session). The initial positions of the items and the positional change within and across the 19 categories were selected without replacement for each bird. Due to difficulties in requiring the birds to peck the top and bottom portions of the screen, the changed item did not appear in the top two or bottom two positions of the outer concentric circle. This left 304 possible positional changes within the 19 distance categories. Four trials from each of the 19 categories were selected for each session.

3.1.3. Procedure

3.1.3.1. Change-detection training. Sessions lasted for 76 trials or 60 min, whichever came first. If a session was not completed on a single day, then the pigeon would start on the next trial of the unfinished session on the following day. Each trial consisted of a ready signal, a sample display followed by a delay consisting of an empty screen, and then a test display. During preliminary training, only one item was used; the item changed position from sample to test display on every trial and remained until pecked. Following preliminary training, two items were presented but only one item changed location.

A trial began with the presentation of the ready signal, which remained until pecked. The sample display was then presented for 1000, 2000, or 3000 ms and then removed leaving only a black background for the delay interval that was either 0, 50, or 100 ms. After the delay expired, the test display was presented until the subject pecked one of the items in the test display. The correct item had been shifted from its position in the sample display. All other characteristics of the item (e.g., color and shape) and test display (e.g., the position, color, and shape of the remaining items) were identical to the sample display. A peck to the item that was changed in location was reinforced with 3 s access to mixed grain. Selection of the unchanged item resulted in the immediate termination of the trial without reinforcement, followed by a 3-s

flashing light. A 20-s inter-trial interval (ITI) separated the end of grain access or the flashing light from the next presentation of the ready signal. The houselight was on during a trial, but was turned off and the screen blank during ITI. A correction procedure was implemented such that incorrect trials were repeated up to five times before the next trial was presented. The correction procedure was only used during initial two-item training and removed one week prior to the introduction of three-item trials. Training with the correction procedure continued until the subjects achieved an average of $\geq 75\%$ correct across two blocks of ten sessions. The correction procedure was then removed and training continued until the subjects completed eight sessions and maintained an average percent correct of $\geq 70\%$.

Following Phase 1 sessions, which contained only two-item trials, all subjects received 30 sessions of equal numbers of 2-item and 3-item trials during Phase 2. On both kinds of trials, all items were present on the test display but one item was in a new position. Phase 3 consisted of 30 sessions of an equal number of 2-item, 3-item, and 4-item trials.

3.2. Analysis strategy

The primary goal of the analyses in this manuscript was to determine the impact of several variables on change detection performance. The measures included percent correct (sum of trials with a response to the changed item divided by total trials) and response time (time of response to an item on the test display minus time of onset of the test display). Because all manipulations were within-subject, the most common analysis method is the repeated-measures analysis of variance (RM ANOVA). For estimating changes in accuracy, however, RM ANOVA has serious drawbacks. The primary problem is the assumption that percent correct data follow a normal distribution with a range that includes all real numbers. Of course, percent correct data are proportions and are bounded between 0 and 100. This restricted range can make the effect sizes and *p*-values (and confidence intervals) from an RM ANOVA uninterpretable. To remedy this problem, we included the results from a binomial generalized linear mixed-effects model (GLMM) to compare to the results of the more common RM ANOVA. The binomial GLMM overcomes the range restriction by fitting the data in the log odds-ratio space. The impact of this transform is that the effect size of a change in raw accuracy changes across the scale. For example, a 10% increase in accuracy from 60% to 70% corresponds to a 3.6 increase in the odds ratio (0.7/0.3 *vs.* 0.6/0.4), but the same 10% increase from 80% to 90% corresponds to a 2.25 increase in the odds ratio (0.9/0.1 *vs.* 0.8/0.2). In general, changes near 50% have a smaller change in odds ratio than the same change nearer the boundary. A secondary problem of RM ANOVA is the inability to specify complex random-effects structures. Thus, even for log response time which may be assumed to be normally distributed, we applied linear mixed effects models to these data to take advantage of the flexibility of mixed effects models to specify complex random-effects structures.

3.3. Results and discussion

Although all pigeons were trained on the same days, some pigeons required more than one day to complete a session. Across blocks of training the pigeons separated into two groups and were tested as such (see Fig. 2, as well as a see Online Resource 1 for a video of the task). One pigeon never acquired the two-item task and was removed from training after 30 days. The first group of three birds completed an average of 96 sessions ($SD = 11.55$) and was tested 12 weeks before the second group of two birds ($M = 145$ sessions, $SD = 7.07$).

Fig. 3 displays percent correct on blocks of sessions across phases with increasing numbers of items. Single-sample *t*-tests against chance (chance level for 2-items trials was 50%, 33% for 3-item, and 25% for 4-item trials) were conducted for each of the array sizes during each phase in which they were present. Analyses revealed that, for 2-items trials, percent correct was significantly above chance (25%) levels for
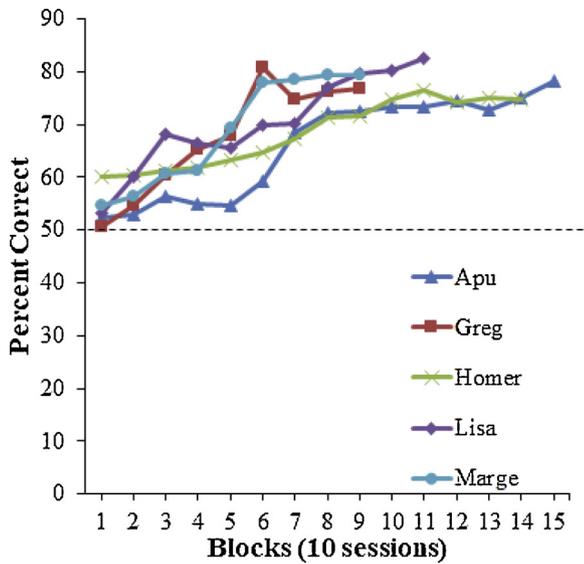
Fig. 2. Mean percent correct across sessions of 2-item trials in Experiment 1a with pigeons. Lisa, Marge, and Greg were tested at the same time during acquisition, whereas Apu and Homer required additional sessions before testing. The dotted line represents chance-level performance.
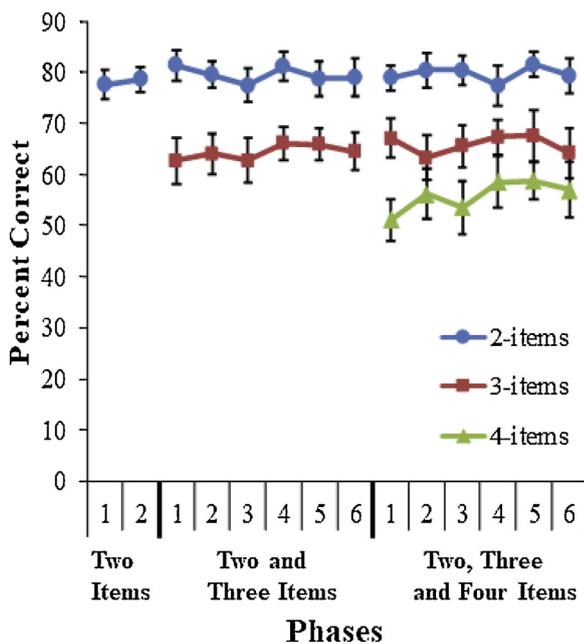

Fig. 3. Mean percent correct across phases (1, 2, and 3) and blocks (5 sessions) in Experiment 1a with pigeons. The number of items refers to size of the array during the sample and test display. The first two blocks are from Phase 1 sessions of two-item only trials, followed by six blocks of Phase 2 and Phase 3 sessions. Error bars represent standard error of the mean across subjects.

all phases (1, 2, and 3), ts (4) $\geq$ 13.50, $ps \leq$ .001. Similarly, percent correct was significantly above chance (33%) on 3-item trials in Phases 2 and 3, ts (4) $\geq$ 17.33, $ps \leq$ .001. Finally, results showed that subjects performed significantly above chance (25%) on 4-item trials in Phase 3, t (4) = 14.17, $p \leq$ .001.

Improvement was assessed by comparing acquisition across the first 30 days when each type of trial was introduced (Phase 2 for 3-item trials and Phase 3 for 4-item trials). We excluded the training that occurred with 2-item trials in Phase 1 and prior to it because pigeons were learning basic aspects of the task unrelated to working memory performance (e.g., correction procedure). A RM ANOVA was conducted

with Blocks (1–6) and Array Size (three and four) included as factors. The results indicated there was both a main effect of Block F(5, 20) = 3.91, p = .01, η^2 = .49, as well as a main effect of Array Size, F (1, 4) = 29.51, $p <$ .001, η^2 = .99, but no interaction, F(5, 20) < 1.0. Tukey's Honestly Significant Differences (HSD) post-hoc comparisons indicated that Block 1 (M = .56, SD = .08) differed from Block 4 (M = .62, SD = .05), p = .03, and Block 5 (M = .62, SD = .05), p = .03, but not Block 6 (M = .60, SD = .05), p = .18. A separate post-hoc test revealed percent correct was significantly better with 3 items (M = .64, SD = .05) than 4 items (M = .55, SD = .05), p = .006. We also used a GLMM (function *glmer* in R, with family set to *binomial*) with fixed effect of Array Size and allowed the effect of Session to vary across subjects by including it as a random effect. The model yielded results comparable to the RM ANOVA, which included a significant main effect of Array Size, X^2 (1) = 69.08, $p <$ .001, and no Array Size x Session interaction, X^2 (1) = 2.08, p = 0.15, but only a marginal effect of Session, X^2 (1) = 3.29, p = 0.07 (see Table 1).

The effects of Array Size (2, 3, and 4 items), Delay (0, 50, and 100 ms), and View Time (1000, 2000, and 3000 ms) on change detection were analyzed with a RM ANOVA conducted on percent correct during Phase 3 (see Fig. 4). This phase consisted of sessions of an equal number of 2-item, 3-item, and 4-item trials. The results indicated a main effect of Array Size F(2, 8) = 195.67, $p <$ .001, η^2 = .98, but no main effect of Delay, F(2, 8) < 1.0, or View Time, F(2, 8) = 1.22, p = .34, η^2 = .23. There was a marginally significant interaction between Delay and View Time, F(4, 16) = 2.85, p = .06, η^2 = .42, but no other significant two-way interactions, Fs(4,16) $\leq$ .35, $ps \geq$ .84, $\eta^2 s \leq$ .08. The three-way interaction between Array Size, Delay, and View Time was also non-significant, F(8, 32) < 1.0. Tukey's HSD post-hoc tests on the main effect of Array Size again revealed that as the number of items increased, percent correct decreased, $ps \leq$ .001. Tukey's HSD post-hoc tests on the marginal interaction revealed that no post-hoc comparisons differed significantly, $ps >$ .09, although percent correct with a 1-s View Time and a 50 ms delay was marginally lower than with a 2-s View Time, p = .09 (see Fig. 4). The GLMM analysis failed to find the significant Delay x View Time interaction, X^2 (1) = .66, p = 0.42 reported for the ANOVA, but all other results were consistent with the ANOVA (see Table 1).

The position of the changed item within the array and distance traveled by the changed item were confounded, such that some positions were not possible with some distance changes. Consequently, two separate RM ANOVAs evaluated the effects of Position (Inner-Inner, Inner-Outer, Outer-Inner, and Outer-Outer) and Distance (a "short" distance consisted of eight distance changes covering 98–205 pixels, "medium" was seven changes from 236 to 349 pixels, and "long" was eight distance changes from 362 to 512 pixels) on percent correct during Phase 3. The Position factor combined the position of the changed item within the two concentric circles on the sample (inner or outer) and test displays (inner or outer). Fig. 5 displays mean percent correct as a function of Position and Array Size. Mean percent correct on trials with the changed item in the inner position during the sample or test (or both) display appeared higher than outer-outer trials regardless of array size, but the pattern across positions changed with array size. The ANOVA with Position and Array Size as repeated measures found a main effect of Array Size, F(2, 8) = 189.62, $p <$ .001, η^2 = .98, and a Position x Array Size interaction, F(6, 24) = 7.20, $p <$ .001, η^2 = .64, but no main effect of Position, F(3, 12) = 2.57, p = .10, η^2 = .39. Subsequent Tukey's HSD post-hoc tests of Position confirmed that Inner-Inner differed reliably from Outer-Outer for 2-item and 4-item trials, $ps <$ .05, but not 3-item trials, p = .99. Inner-Inner also differed reliably from Inner-Outer for 2-item trials, $p <$ .001, but no other comparisons were reliable, $ps >$ .06. The GLMM analysis found the same main effect and interaction as the ANOVA, but also found a main effect of Position, X^2 (3) = 21.34, $p <$ .001.

The ANOVA with distance found a main effect of Distance, F(2, 8) = 8.13, p = .01, η^2 = .67, and the main effect of Array Size

Table 1
GLMM Analysis for Experiments 1a and 1b.

	Percent Correct			Log RT		
	Chisq	Df	Pr(> Chisq)	Chisq	Df	Pr(> Chisq)
Experiment 1a						
Phases 2 and 3						
Abs.Session	3.29	1	0.07			
TrialSetSize	69.08	1	< .001*			
Abs.Session:TrialSetSize	2.08	1	0.15			
Phase 3						
ArraySize	476.19	2	< .001*	1.64	2	0.44
Delay	0.12	1	0.73	1.58	1	0.21
ViewTime	1.14	1	0.29	2.52	1	0.11
ArraySize:Delay	0.04	2	0.98	2.71	2	0.26
ArraySize:ViewTime	0.53	2	0.77	4.22	2	0.12
Delay:ViewTime	0.66	1	0.42	0.02	1	0.90
ArraySize:Delay:ViewTime	0.35	2	0.84	1.28	2	0.53
Abs.Session	2.98	1	0.08			
Distance	4.84	1	0.03*			
ArraySize:Distance	0.20	2	0.91			
Position	21.34	3	< .001*,†			
ArraySize:Position	19.58	6	< .01*			
Experiment 1b						
All trials (no 0 ms delay)						
ArraySize	381.05	2	< .001*	3.66	2	0.16
Delay	96.45	1	< .001*	27.08	1	< .001*
Mask	64.19	1	< .001*	99.82	1	< .001*
ArraySize:Delay	1.73	2	0.42	12.91	2	< .01*
ArraySize:Mask	3.54	2	0.17†	1.22	2	0.54
Delay:Mask	30.66	1	< .001*	0.71	1	0.40
ArraySize:Delay:Mask	0.32	2	0.85	1.08	2	0.58
No mask						
ArraySize	217.27	2	< .001*			
Delay	166.57	1	< .001*			
ArraySize:Delay	0.00	2	1.00			
Position	39.60	3	< .001*			
Distance	22.81	1	< .001*			
ArraySize:Distance	2.93	2	0.15			
ArraySize:Position	3.98	6	0.69			

* Pr < .05.
† Different from ANOVA.

previously reported, but no interaction of Distance x Array Size, $F(4, 16) < 1.0$ (see Supplementary Fig. 1). Subsequent Tukey's HSD post-hoc tests of Distance found that short ($M = 65.05$, $SEM = 1.75$) differed reliably from medium ($M = 68.72$, $SEM = 1.57$), $p = .01$, and marginally from long ($M = 67.50$, $SEM = 1.57$), $p = .07$, but medium and long did not differ from each other, $p = .42$. The GLMM analysis

produced identical results to the ANOVA (see Table 1).

The response times from correct trials during Phase 3 were log transformed (base 10) and analyzed with an ANOVA and GLMM with Array Size, Delay, and View Time as fixed effects. Neither analysis revealed any significant main effects or interactions, $Fs < 3.60$, $ps > .07$, $\eta^2 < .48$, the nearest of which was the main effect of View Time, F
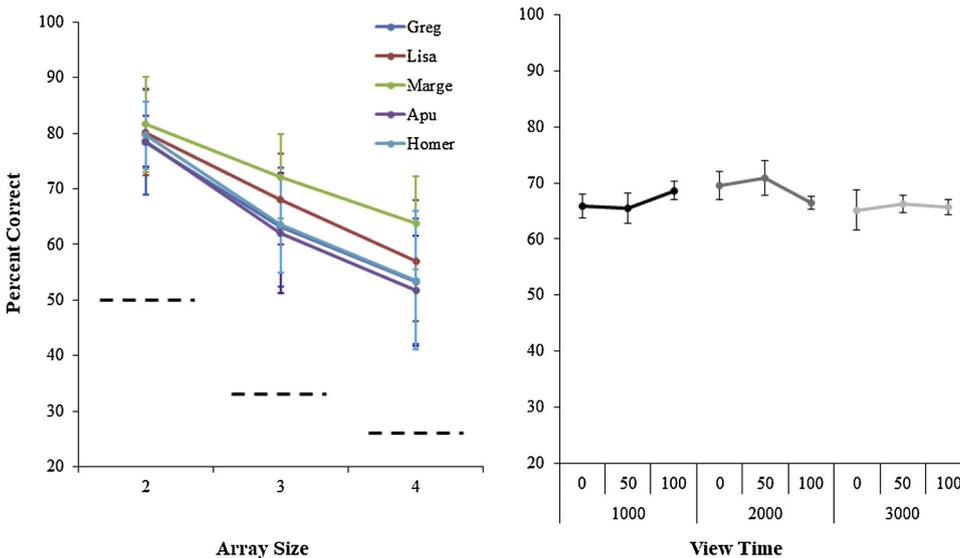

Fig. 4. Left panel: Mean percent correct of pigeons with different array sizes in Experiment 1a during the last 30 sessions of training (Phase 3). Dash lines represent chance performance for each array size. Error bars represent standard deviation. Right panel: Mean percent correct as a function of the duration of the sample stimulus (View Time; 1000, 2000, or 3000 ms) and Delay (0, 50, or 100 ms) between the sample and test display. Error bars represent standard error of the mean across subjects.
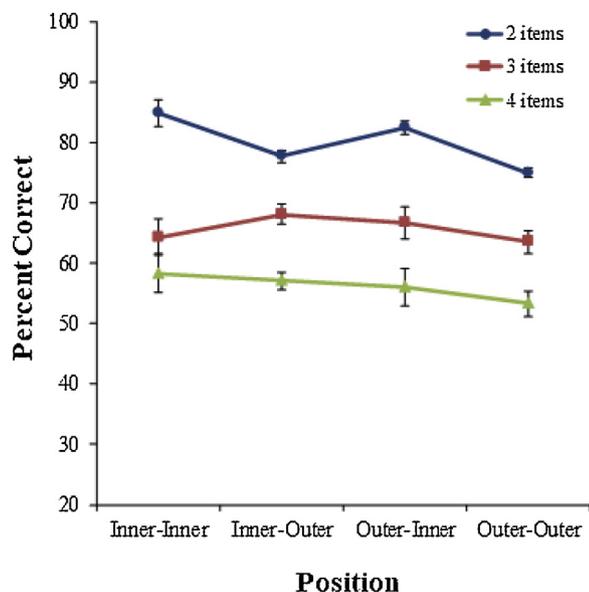
Fig. 5. Mean percent correct as function of array size and position of change in the display with pigeons in Experiment 1a. The position of the to-be-changed item within the two concentric circles is indicated by the word preceding the hyphen and the position of the changed item in the test display is indicated post hyphen. For example, "Inner-Outer" indicated an item located in the inner concentric circle array changed to another position within the outer concentric circle. Error bars represent standard error of the mean across subjects.

(2,8) = 3.60, $p = .08$, $\eta^2 = .47$. Descriptive statistics revealed that the mean response times were longer for View Time 1000 ms ($M = 1.48$, $SEM = .22$) than View Time 2000 ms ($M = 1.37$, $SEM = .19$) or 3000 ms ($M = 1.39$, $SEM = .19$).

Several of the findings in Experiment 1a with pigeons are consistent with the literature on working memory in humans. Firstly, there was a questionable (e.g., percent correct during Block 1 differed from 4 and 5, but not 6 based on the ANOVA and a marginal main effect of session in GLMM) improvement in visual working memory performance across 30-days of training. Secondly, as the number of items in the array increased, accurate change detection decreased. Thirdly, percent correct was higher when the changed item traveled a medium or long versus short distance. The finding of higher percent correct for inner-inner than outer-outer changes is reminiscent of an attentional effect reported by Sperling and Melchner (1978). Their study involved presenting arrays made up of several letters and two digits for 240 ms. Participants were required to name the location and identity of the two digits from each display. Participants were more likely to detect the digits if they were previously directed to attend to the area in which the digits were presented (inner versus outer square). In the present research, the start of a trial required the pigeons to peck a ready signal positioned at the center of the display (in the center of the smaller concentric circle), and change-detection performance was better when the changed item moved into or out of the inner circle. However, it is also possible that the pigeons simply tended to perseverate in the area near the center of the screen, which would increase the likelihood of observing the change of an item that started or ended in the inner portion of the display. Future studies could investigate whether moving the ready signal to different quadrants of the display would enhance detection in that quadrant.

In contrast to previous literature, longer view times did not enhance memory performance. Interestingly, the delay between displays had no effect on percent correct. Research from humans indicates performance is significantly worse at durations > 60 ms for most change-detection tasks (Pashler, 1988; Phillips, 1974; Rensink et al., 1997), but for pigeons change detection has been found to be stable out to 1000 ms (Leising et al., 2013; Wright et al., 2010; but see Herbranson and Davis,

2016). The replicability of this finding and its relationship to masking effects were further explored in Experiment 1b.

4. Experiment 1b

In humans, masking effects occur for items in memory when those items are followed by variations in the current visual field shortly after the items were viewed (< 67 ms). Masking effects are the result of the dependence of iconic memory on retinotopic coordinates, which are disrupted by movement of the display (Phillips, 1974) or by subsequent rapid replacement of the display with an unrelated display (e.g., grey and black checkerboard pattern). Memory for items after longer delays are less susceptible to the effects of a visual mask (Pashler, 1988; Phillips, 1974; Wilson et al., 1983). Experiment 1b altered the delay durations to include trials with and without a mask following short (100 ms) and long (1000 ms) delays.

4.1. Method

4.1.1. Subjects

The five pigeons (*Columba livia*) that acquired the task in Experiment 1a were used as subjects. All other details are identical to Experiment 1a.

4.1.2. Apparatus

Same as in Experiment 1a.

4.1.2.1. Stimulus displays. A 40 × 54 black and gray checkerboard was used as the mask visual stimulus. The mask filled the entire display for the full duration of the delay. All other details were identical to Experiment 1a.

4.1.3. Procedure

4.1.3.1. Change-detection training. Training continued for 14 sessions for all subjects. The number of trials, view time, and delay differed from Experiment 1a. The total number of trials per session was increased to 150. The sample duration was 1000 ms for all trials. The 50-ms delay was replaced with a 1000-ms delay, resulting in an equal number of trials with 0, 100, and 1000 ms delays. The visual mask appeared on 10 trials per session of each array size (2, 3, and 4) with a 100 and 1000-ms delay (10 trials x 3 array sizes x 2 delays = 60 mask trials). The remaining 90 trials were randomly distributed between the other trial types (array sizes 2, 3, or 4 with a 0, 100, or 1000-ms delays). On mask trials the checkerboard visual mask was inserted for the entire duration of the delay. All other details were the same as in Experiment 1a.

4.2. Results and discussion

The two changes of greatest interest were the additions of a 1000-ms delay and the inclusion of a visual mask on some trials. Fig. 6 displays percent correct as a function of array size and delay. All pigeons performed worst at the longest delay. The new delay was analyzed using an RM ANOVA with Array Size (2–4) and Delay (0, 100 and 1000 ms) as repeated measures. This analysis revealed a main effect of Array Size, $F(2, 8) = 185.74$, $p < .001$, $\eta^2 = .98$, and Delay, $F(2, 8) = 30.82$, $p < .001$, $\eta^2 = .86$, but no interaction between Array Size and Delay, $F(4, 16) = .47$, $p = .76$, $\eta^2 = .11$. Tukey's HSD post-hoc tests revealed that each array size differed from the other, $ps < .001$. With respect to delays, the 0-ms and 100-ms delays both differed from 1000-ms, $ps < .001$, but did not differ from each other $p = .99$. A GLMM (function *glmer* in R, with family set to *binomial*) with fixed effects of Array Size and Delay and Subject as a random effect yielded the same main effects as the ANOVA, ($Xs^2 > 166.57$, $ps < .001$), and no Delay x Array Size interaction, $X^2 (2) < 1.0$ (see Table 1).

The inclusion of the mask during a delay prevented mask trials of 0-ms (since there was no delay in which to insert a mask). A second
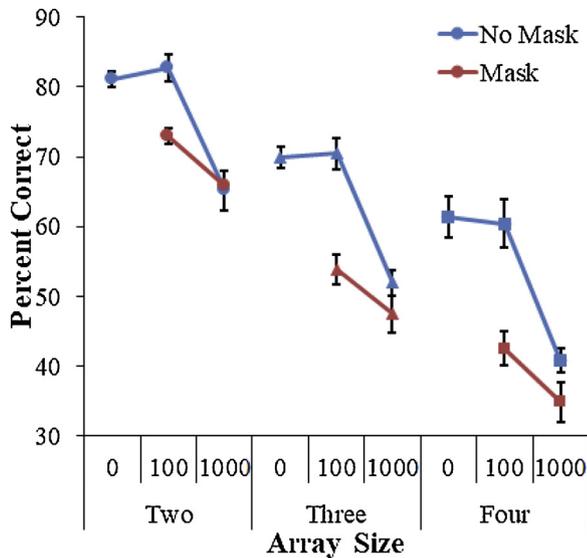
Fig. 6. Mean percent correct as a function of array size (two, three, or four items in the displays) and delay (0, 100, and 1000 ms) in Experiment 1b with pigeons. Mask refers to trials with the insertion of a visual mask between the sample and test displays. Error bars represent standard error of the mean across subjects.
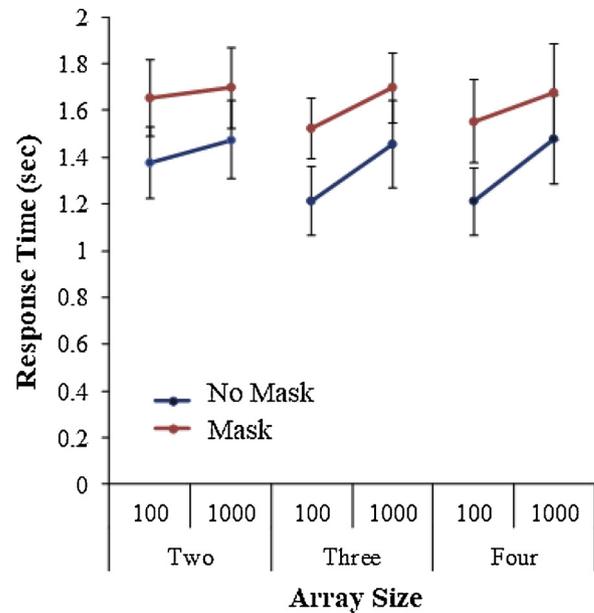

Fig. 7. Mean response times as a function of array size (two, three, or four items in the displays) and delay (0, 100, or 1000 ms) in Experiment 1b with pigeons. Mask refers to trials with the insertion of a visual mask between the sample and test displays. Error bars represent standard error of the mean across subjects.

analysis was conducted that included trials with or without a Mask, excluding trials with a 0-ms delay (see Fig. 6). This analysis revealed a main effect of Mask, $F(1, 4) = 40.18$, $p < .01$, $\eta^2 = .91$, Array Size, $F(2, 8) = 168.20$, $p < .001$, $\eta^2 = .98$, and Delay, $F(1, 4) = 41.61$, $p < .01$, $\eta^2 = .92$, as well as Mask x Array Size, $F(2, 8) = 5.95$, $p = .03$, $\eta2 = .60$, and Mask x Delay interactions, $F(1, 4) = 10.49$, $p = .03$, $\eta^2 = .72$. There were no other significant interactions, $Fs(2, 8) \leq 1.0$. Tukey's HSD post-hoc test of the Mask by Array Size interaction revealed that the difference between Mask and No Mask with 2-item trials was not statistically significant, $p = .14$, whereas the same comparisons with 3-item and 4-item trials were reliably different, $ps < .01$. Tukey's HSD post-hoc tests isolated the effects driving the Mask by Delay interaction. Percent correct on trials with a mask and a delay of 100 ms was lower than on trials of the same delay but no mask, $p < .05$, whereas performance was similar on trials with a delay of 1000 ms with or without a mask, $p = .61$. The GLMM yielded the same main effects and Delay x Mask interaction as the ANOVA, ($Xs^2 > 30.65$, $ps < .001$), but did not yield a significant Array Size x Mask interaction, $X^2 (2) = 3.54$, $p = 0.17$.

Fig. 7 displays mean response times from correct trials. Response times were longer for mask than no-mask trials, and longer following a delay of 1000 than 100 ms. The response times from trials with delays of 100 or 1000 (excluding trials with 0 delay) with were log transformed (base 10) and analyzed with a RM ANOVA with Array Size (2, 3, and 4), Delay (100 and 1000), and Mask (0 and 1) as repeated measures. This analysis revealed a main effect of Mask, $F(1, 4) = 132.64$, $p < .01$, $\eta^2 = .97$, Delay, $F(1, 4) = 20.30$, $p = .01$, $\eta^2 = .84$, as well as a Delay x Array Size interaction, $F(2, 8) = 16.70$, $p < .01$, $\eta^2 = .81$. There were no other significant main effects or interactions, $Fs(2, 8) \leq 2.30$, $ps \geq .16$, $\eta^2s \leq .37$. Post-hoc Tukey tests on the interaction revealed that mean response times following 100 and 1000 ms delays did not differ on 2-item trials, $p = .52$, whereas mean response times were longer for the 1000 ms delay for both of the larger array sizes, $ps < .01$. A GLMM with Array Size, Delay, and View Time as fixed effects and Subject as a random factor found identical results to the ANOVA (see Table 1).

These results revealed several important findings. Firstly, the data without the mask from trials with the 0 and 100 ms delays replicated the result from Experiment 1a of comparable performance at both

delays. Secondly, percent correct during both of these short delays was higher than on trials with the longer delay (1000 ms). From these data it could be hypothesized that whatever mechanism supported change detection performance at 0 ms was the same at 100 ms, but different than that operating at 1000 ms. Thirdly, this hypothesis was supported by different effects of the visual mask on trials with 100 ms and 1000 ms delays. A visual mask during the 100 ms delay disrupted change detection, whereas a mask during the 1000 ms delay had no effect on percent correct. Lastly, the effect of the visual mask on trials with the short delay was greater with larger array sizes. These data are consistent with previous research indicating the effect of a visual mask is confined to short durations following a stimulus and more disruptive with larger arrays (Phillips, 1974; Vogel et al., 2006). To our knowledge, this is the first demonstration of a masking effect with pigeons.

5. Experiment 2

Experiments 2a and 2b were conducted with humans completing a spatial change-detection procedure nearly identical to that used in Experiment 1 with pigeons. There is a rich history of comparisons of perception, learning, and cognition across species. These comparisons are often focused on revealing functional relationships (Wright, 2013). Subjects are exposed to several levels of an independent variable to reveal a systematic relationship between the levels and different kinds of responding. For example, pigeons and monkeys shown few exemplars of a concept typically fail to transfer the correct response to novel examples of that concept, whereas, a greater number of trained exemplars supports better transfer across several species (e.g., Wright and Katz, 2006). Experiment 2a used the parameters from Experiment 1b to investigate the same functional relationship between delay and visual mask in humans within a spatial change-detection procedure.

6. Experiment 2a

In Experiment 2a, the interest was in whether the functional relationship between delay duration and the masking effect would be present with humans using the exact same parameters from Experiment 1b (with pigeons). Human subjects were exposed to a single session of

trials with a procedure nearly identical to that of Experiment 1b. The primary differences were the mode of selecting the changed item, the ITI between trials, and the method of feedback following a selection (i.e., reinforcement).

6.1. Method

6.1.1. Participants

A total of 20 undergraduate psychology students (13 females and 7 males, and aged 18–24 years) at TCU participated as a partial fulfillment of course requirements. None of the students had any previous experience with the procedure, and all were uninformed as to the purpose of the experiment. Participants were trained and tested individually. Sessions consisted of 150 trials completed within 60 min. All research was conducted in accordance with TCU's Human Participant Ethics Committee and an approved Institutional Review Board (IRB) protocol.

6.1.2. Apparatus and stimuli

A Hewlett Packard Touchsmart computer was used to display items; the computer monitor dimensions were 48.26 × 30.48 cm.

6.1.2.1. Stimulus displays.
The same eight circles from Experiments 1a and 1b were used. The size of each item on the screen was stretched by the widescreen display to 3.00 × 2.20 cm for the ready signal and 2.20 × 1.60 cm for the training stimuli. Based on a typical viewing distance of 55 cm, this represented 2.29° x .1.67° visual angle.

6.1.3. Procedure
6.1.3.1. Change-detection training.
Participants were instructed to sit comfortably and rest one elbow on a mousepad in front of the computer. The participants were asked to maintain this basic position throughout the experiment. The subjects were provided with minimal information regarding the task.

> "Each trial will begin with a black cross at the center of a black screen. Please rest your eyes on the cross until you are ready to respond. Once ready, please touch the cross. The first display will be presented immediately but for only 1000 ms. After the sample display disappears, a short delay will follow and then a second display (test display) will be presented. One of the items in the second display will have changed position. Your goal is to touch the item that has moved position. Sometimes a checker pattern "mask" will appear during the delay between the first and second displays. Please continue to look at the screen, but the checker pattern doesn't provide any clues regarding the location of the changed item.
> The game moves quickly, but at any time you can take a break before touching the black cross and starting the next trial. **One last thing**: although your goal is to get as many <u>correct choices</u> as possible, we ask you to refrain from any strategies to make the game easier, such as pointing at the screen, or looking away during the checker pattern."

The same variations in the distance the changed item traveled, the Array Size, Delay, and the presence or absence of a visual mask from Experiment 1b were used. In contrast to previous experiments, trials were separated by a 2-s ITI. Following the ITI, the ready signal was presented and remained until touched. The sample display remained for 1000 ms and was followed by delay of 0, 100, or 1000 ms. After the interval expired, the test display was presented until the participant touched an item. The correct item had shifted in position with respect to the sample display. A touch to the item that had changed location was followed by a 1-s windows chime (windows\media\tada.wav); selection of the unchanged item resulted in the entire display changing red for 1-s. Four sessions were created with training colors and positional changes selected without replacement across sessions. To be consistent with the pigeons, the changed item was restricted from
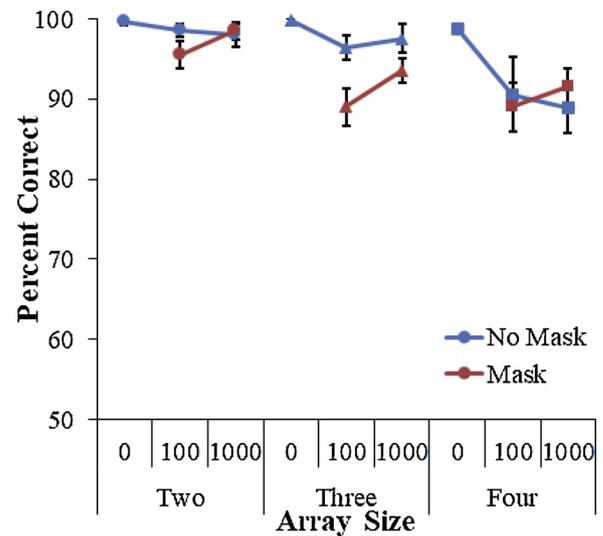

Fig. 8. Mean percent correct as a function of array size (two, three, or four items in the display) and delay (0, 100, and 1000 ms) in Experiment 2a with humans. Mask refers to trials with the insertion of a visual mask between the sample and test displays. Error bars represent standard error of the mean across subjects.

appearing in the top two or bottom two positions of the outer concentric circle. As in Experiment 1b, there were 150 trials per session with 10 trials per session of each array size (2, 3, and 4) with a 100 and 1000-ms delay and visual mask. The remaining trials were randomly distributed between the other trial types. On mask trials the checkerboard visual mask was inserted for the entire duration of the delay.

6.2. Results and discussion

As in Experiment 1b, the visual mask and delay effects were analyzed separately due to the absence of trials with a mask at the 0-s delay (see Fig. 8). Mean percent correct was above 88% in all conditions. Looking only at no-mask trials, percent correct appears lower with longer delays and larger array sizes, with the greatest effect of delay during 4-item trials. A RM ANOVA with Delay (0, 100, and 1000 ms) and Array Size (2, 3, and 4 items) as factors and conducted on percent correct revealed a main effect of Array Size, $F(2, 38) = 7.90$, $p \leq .01$, $\eta^2 = .30$, as well as a marginally non-significant main effect of delay, $F(2, 38) = 3.24$, $p = .05$, $\eta^2 = .15$, but no interaction, $F(4, 76) = 1.77$, $p = .14$, $\eta^2 = .09$. Tukey's HSD post-hoc tests of the main effect of Array Size showed that there was not a significant difference between percent correct on 2- and 3-items trials, $p = .88$; but both significantly differed from 4-item trials, $ps < .01$. Post-hoc comparisons of the main effect of Delay indicated that the difference in percent correct was marginally non-significant between 0 and 100 ms, $p = .10$, or 1000 ms, $p = .07$. In contrast, the difference in percent correct between 100 and 1000 ms was non-significant, $p = .98$. A GLMM analysis with Array Size and Delay as fixed factors and Subject as a random factor found the same main effect of Array Size, $X^2(2) = 17.67$, $p < .001$, and no interaction, but in contrast, the GLMM did uncover the effect of Delay, $X^2(1) = 16.84$, $p < .001$ (see Table 2).

A separate ANOVA with the 0-ms trials removed and Array Size, Delay (100 and 1000), and Mask (0 and 1) as repeated measure was conducted (see Fig. 8). This analysis revealed a main effect of Array Size, $F(2, 38) = 12.33$, $p < .001.$, $\eta^2 = .39$, a marginally non-significant effect of Mask, $F(1, 19) = 3.39$, $p = .08.$, $\eta^2 = .15$, and a marginally non-significant Mask x Array Size interaction, $F(2, 38) = 3.01$, $p = .06$, $\eta^2 = .14$. All other comparisons, including the Mask x Delay interaction, were non-significant, $Fs(2, 38) < 1.82$, $ps > .19$, $\eta^2 < .09$. A GLMM analysis with Array Size, Delay, and Mask

Table 2
GLMM Analysis for Experiments 2a and 2b.

	Percent Correct			Log RT		
	Chisq	Df	Pr(> Chisq)	Chisq	Df	Pr(> Chisq)
Experiment 2a						
All trials (no 0 ms delay)						
ArraySize	27.62	2	< .001[*]	83.16	1	< .001[*]
Delay	3.06	1	0.08	10.83	1	< .001[*]
Mask	2.40	1	0.12	31.24	1	< .001[*]
ArraySize:Delay	1.72	2	0.42	0.20	1	0.66
ArraySize:Mask	6.49	2	0.04[*,†]	0.14	1	0.71
Delay:Mask	1.66	1	0.20	9.26	1	< .01[*]
No mask						
ArraySize	17.67	2	< .001[*]			
Delay	16.84	1	< .001[*,†]			
ArraySize:Delay	1.33	2	0.51			
Distance	0.15	1	0.70			
ArraySize:Distance	3.82	2	0.15			
Position	3.23	3	0.36			
Experiment 2b	Percent Correct			Log RT		
All trials (no 0 ms delay)						
ArraySize	91.02	2	< .001	21.39	2	< .001[*]
Delay	4.35	1	0.04[*,†]	12.33	1	< .001[*]
Mask	1.99	1	0.16	11.23	1	< .001[*]
ArraySize:Delay	0.06	2	0.97	6.32	2	0.04[*]
ArraySize:Mask	0.29	2	0.86	10.89	2	< .01[*]
Delay:Mask	18.51	1	< .001[*]	19.25	1	< .001[*]
No mask						
ArraySize	19.17	2	< .001[*]			
Delay	136.22	1	< .001[*]			
ArraySize:Delay	5.34	2	0.07[†]			
Distance	0.58	1	0.45			
ArraySize:Distance	8.36	2	0.02[*]			
Position	10.70	3	0.01[*]			

[*] Pr < .05.
[†] Different from ANOVA.

as fixed factors and Subject as a random factor produced identical results, except the interaction between Array Size and Mask was significant, $X^2 (2) = 6.49$, $p = 0.04$ (see Table 2). The three-way interaction could not be added without affecting converging.

The human participants completed a single session, which reduced the number of trials per trial type. Separate RM ANOVAs were conducted on percent correct from all trials except those with a mask (to compare to Experiment 1a) with Array Size and Position (inner-inner, inner-outer, outer-inner, and outer-outer) or Distance (short, medium, and long) as repeated measures. The analysis with Position could not be completed because the mean across trials with Inner-Inner and Inner-Outer positions had no variance (M = 100%) on trials of Array Size 2 and 3. When collapsed across Array Size, the descriptive means indicate that Inner-Inner (M = 98.25, SEM = 1.32) and Inner-Outer (M = 98.06, SEM = 1.27) were higher than Outer-Inner (M = 96.45, SEM = 1.40) and Outer-Outer (M = 97.39, SEM = 1.59). A GLMM analysis could be conducted, but with Position as the only fixed factor and Subject as a random factor. This analysis failed to find a significant effect of position, $X^2 (3) = 3.23$, $p = 0.36$. A separate ANOVA with distance revealed the main effect of Array Size reported previously, but no other main effects or interactions, $Fs < 1.19$, $ps > .32$, $\eta s^2 < .08$, which matched the results of the GLMM (see Table 2).

Fig. 9 reveals that mean response times from correct trials were longer for trials with a Mask than No Mask, and longer following a Delay of 1000 than 100 ms. However, these two factors also appear to interact. The response times from correct trials with delays of 100 or 1000 (excluding trials with 0 delay) with were log transformed (base 10) and analyzed with a RM ANOVA with Array Size (2, 3, and 4), Delay (100 and 1000), and Mask (0 and 1) as repeated measures. This analysis revealed main effects of Array Size, $F(2, 38) = 26.61$, $p < .001$, $\eta^2 = .58$, Mask, $F(1, 19) = 37.14$, $p < .001$, $\eta^2 = .66$, Delay, $F(1, 19) = 14.43$, $p < .01$, $\eta^2 = .43$, as well as a Delay x Mask
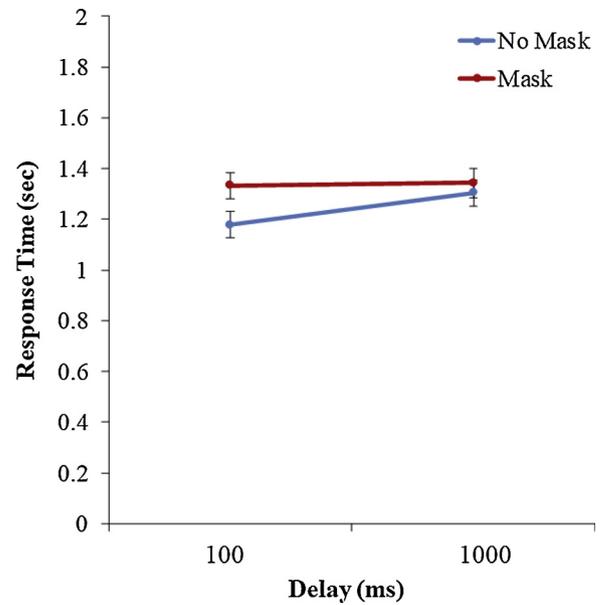

Fig. 9. Mean percent correct across training in Experiment 2a with humans. Mask refers to trials with the insertion of a visual mask between the sample and test displays. Error bars represent standard error of the mean across subjects.

interaction, $F(1, 19) = 20.38$, $p < .001$, $\eta2 = .52$. There were no other significant interactions, $Fs(2, 38) \leq 1.57$, $ps > .22$, $\eta2 s \leq .08$. Tukey post-hoc tests of the main effect of Array Size confirmed that mean response time during 2-item trials (M = 1.14, SEM = .04) was faster than 3-item (M = 2.22, SEM = .04), $p < .001$, and 4-item trials (M = 1.31, SEM = .05), $p < .001$, and mean response time on 3-item trials was faster than 4-items trials, $p < .01$. Post-hoc Tukey tests on the interaction revealed that the mean response time on trials without a mask and following a 100 ms delay was shorter than with a mask, $p < .001$, as well as shorter than the other two conditions, $ps < .001$, but mean response time did not differ between trials with or without a mask following a 1000 ms delay, $p = .37$. A GLMM with Array Size, Delay, and Mask as fixed effects and Subject as a random factor found identical results to the ANOVA (see Table 2). The three-way interaction could not be added without affecting converging.

In contrast to the pigeons in Experiments 1a and 1b, percent correct for humans after a 100-ms delay was more similar to trials with a 1000-ms delay than a 0-ms delay. Regarding array size, unlike pigeons, percent correct in Experiment 1b was similar with 2 and 3-item displays, but lower with 4-item displays. There was also no effect on percent correct of the position or distance of the changed item. One reason for this may be the viewing position of the subject. The pigeons were required to peck the item, and thus may have perceived only a portion of the display. The ready signal was placed in the center of the display, and thus changes that involved an item in the inner circle may have been better detected than changes on the outer circle. Humans, however, viewed the items from .55 m away from the monitor, which may have allowed them to perceive more of the display. Furthermore, a reliable interaction between delay and mask was not found. There was a marginal interaction between array size and mask. In reviewing Fig. 8, percent correct was actually better on trials with the mask than without the mask following a 1000-ms delay when four items were present. Interestingly, the mask slowed response times but did not lower percent correct on trials with a 100 ms delay, and did not affect either measure with a 1000 ms delay. This suggests the mask disrupted how fast, but not how well, participants could detect the changed item differently across the short and long delays. In Experiment 1b, the mask slowed response times in pigeons equally at the 100 and 1000 ms delays, but change detection accuracy differed across the short and long delays. In other words, the mask disrupted how well, but not how fast, the pigeons

could respond differently across the short and long delays.

Experiment 2b attempted to exploit the mask by array size interaction found in Experiment 1b and 2a, as well as reported by Vogel et al. (2006), by increasing the array size to magnify any effect of the mask.

7. Experiment 2b

Experiment 2a failed to replicate the mask by delay effect found with pigeons. The parameters from the pigeon experiment were used intentionally to try and replicate the effects, however, it is often the case that functional relationships are revealed using different levels of the manipulation across species (Wright, 2013). For example, Wright et al. (1985) were interested in the effects of a retention delay on primacy and recency effects in pigeons. A similar pattern was observed in pigeons, monkeys, and humans; such that short retention delays resulted in a large recency effect and no primacy effect, whereas longer delays produced the inverse pattern. Relevant to the current research, the absolute values of the retention delays differed greatly across species. This suggests that in the current task more variation in the array size may be necessary to reveal the functional relationship between delay duration and mask effects in humans. Phillips (1974) and Vogel et al. (2006) reported mask by array size interactions, such that the visual mask disrupted performance more with larger set sizes. In Experiment 2b, the procedure was revised to include trials with array sizes of 4, 6, and 8 items.

7.1. Method

7.1.1. Participants

A total of 20 undergraduate psychology students (15 females and 5 males, and aged 18–24 years) at TCU participated as a partial fulfillment of course requirements. None of the students had any previous experience with the procedure, and all were uninformed as to the purpose of the experiment. Participants were trained and tested individually; the experiment's duration was 60 min. All research was conducted in accordance with TCU's Human Participant Ethics Committee and an approved IRB protocol.

7.1.2. Apparatus and stimuli

The same Hewlett Packard Touchsmart computer was used.

7.1.2.1. Stimulus displays. The same eight circles and visual mask were used.

7.1.3. Procedure
7.1.3.1. Change-detection training. All experimental details were the same as in Experiment 2a, except for the number of items present in the display. The trial began once the ready stimulus positioned in the center of the screen was touched. A sample display appeared for 1000 ms consisting of 4, 6, or 8 items followed by a delay before the test display. As before, during 60 trials the black and gray chessboard-patterned visual mask appeared during the delay, otherwise the display was black and cleared of all items. After the delay, the test display appeared with one item that had changed position. A session included 150 trials.

7.2. Results and discussion

Fig. 10 displays percent correct as a function of delay and array size. Performance appears worse with more items in the array and at longer delays, with the exception of higher percent correct on trials with mask and a 1000-ms than 100-ms delay. We first evaluated the effect of the larger arrays sizes with a RM ANOVA conducted on percent correct with Array Size (4, 6, and 8) and Delay (0, 100, and 1000 ms) as repeated measures on all trials without a mask. This analysis revealed a
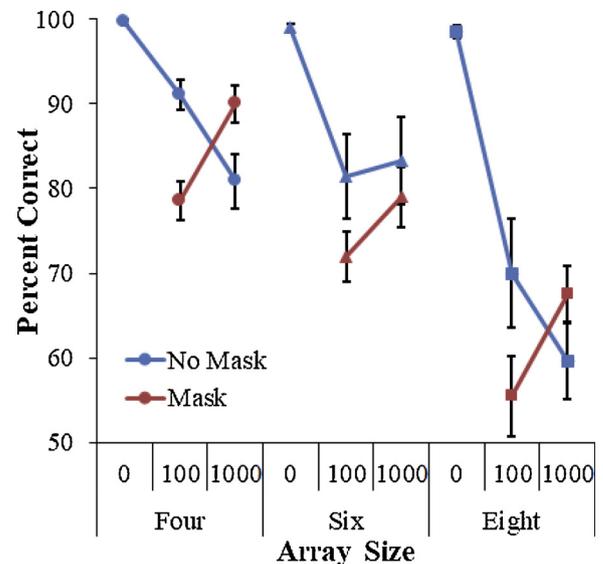

Fig. 10. Mean percent correct as a function of array size (four, six, or eight items in the display) and delay (0, 100, and 1000 ms) in Experiment 2b with humans. Mask refers to trials with the insertion of a visual mask between the sample and test displays. Error bars represent standard error of the mean across subjects.

main effect of Array Size, $F(2, 38) = 16.79$, $p \leq .001$, $\eta^2 = .47$, as well as a significant main effect of Delay, $F(2, 38) = 32.24$, $p < .001$, $\eta^2 = .63$, and the interaction, $F(4, 76) = 5.35$, $p < .001$, $\eta^2 = .22$. Tukey's HSD post-hoc tests of the main effect of Array Size showed that there was not a significant difference between performance on 4-item and 6-item trials, $p = .60$; although, both significantly differed from 8-item trials, $ps < .001$. Post-hoc comparisons of the main effect of Delay indicated that performance differed between trials with 0 and 100 ms delays, $p < .001$, and between 0 and 1000 ms delays, $p < .001$. In contrast, performance between 100 and 1000 ms was non-significant, $p = .13$. The post-hoc tests for the interaction did not reveal a reliable pattern, but as Fig. 10 indicates, the delay appears more disruptive with larger numbers of items. A GLMM analysis was conducted with Array Size and Delay as fixed factors and Subject as a random factor. The GLMM revealed the same main effects as the ANOVA, $Xs^2 > 19.17$, $ps < .001$, but in contrast, detected only a marginally non-significant Array Size x Delay interaction, $X^2 (2) = 5.34$, $p = 0.07$.

A separate ANOVA was conducted with Array Size (4, 6, and 8), Mask (0 and 1) and Delay (100 and 1000-ms, removing the 0-ms trials) on percent correct. This analysis revealed a main effect of Array Size, $F(2, 38) = 27.71$, $p < .001$, $\eta^2 = .59$, no effect of Mask, $F(1, 19) = 2.56$, $p = .13$, $\eta^2 = .12$, but critically, a reliable Mask x Delay interaction, $F(1, 19) = 25.31$, $p < .001$, $\eta^2 = .57$. All other comparisons, including a Mask x Array Size interaction were not significant, $Fs(2, 38) < 2.18$, $ps > .12$, $\eta s^2 < .11$. Post-hoc tests conducted on the Mask by Delay interaction revealed that percent correct on trials with a 100-ms delay differed reliably when there was mask compared to no mask, $p < .001$, however, trials with a 1000-ms delay did not differ reliably with or without a mask $p = .28$ (see Fig. 10). Interestingly, percent correct with the mask and 1000-ms delay was reliably better than with a mask and 100-ms delay, $p < .01$. A GLMM revealed the same main effect of Array Size and Mask by Delay interaction as the ANOVA, but also a main effect of Delay, $X^2 (1) = 4.35$, $p = .04$ (see Table 2). The three-way interaction could not be added without affecting converging.

Position and distance traveled by the changed item were analyzed with two separate RM ANOVAs with Array Size and Position (Inner-Inner, Inner-Outer, Outer-Inner, and Outer-Outer; see Fig. 11) or Distance as the repeated measure. The ANOVA with position found main
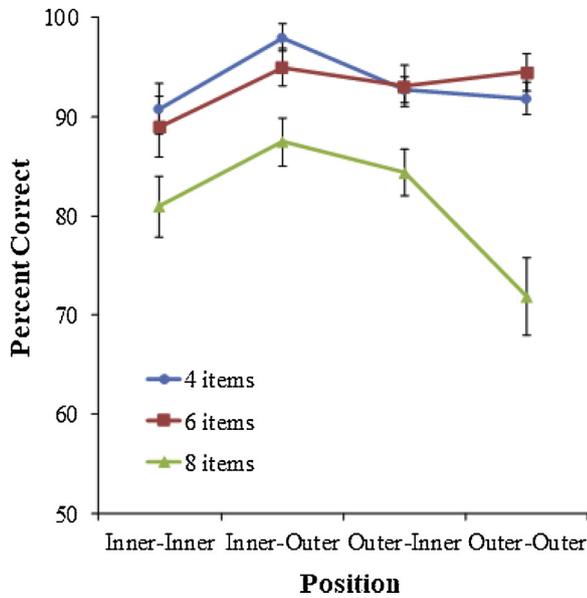
Fig. 11. Mean percent correct as function of array size and location of change in the display with human participants in Experiment 2b. The location of the to-be-changed item within the two concentric circles is indicated by the word preceding the hyphen and the location of the changed item in the test display is indicated post hyphen. For example, "Inner-Inner" indicated an item located in the inner concentric circle array changed to another position within the inner concentric circle. Error bars represent standard error of the mean across participants.
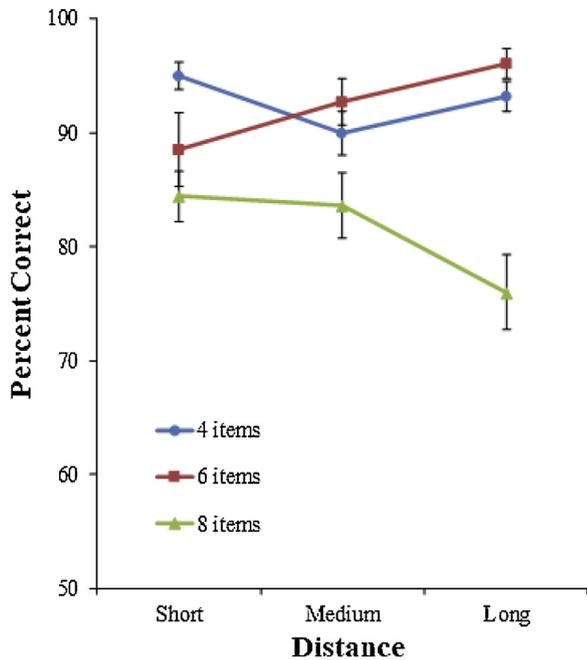

Fig. 13. Mean response time as a function of array size (four, six, or eight items in the display) and delay (0, 100, and 1000 ms) in Experiment 2b with human participants. Mask refers to trials with the insertion of a visual mask between the sample and test displays. Error bars represent standard error of the mean across subjects.

Fig. 12. Mean percent correct as function of array size and distance the changed item traveled from sample to test display with human participants in Experiment 2b. Error bars represent standard error of the mean across participants.

effects of Array Size, $F(2, 38) = 39.47$, $p < .001$, $\eta^2 = .68$, and Position, $F(3, 57) = 8.69$, $p < .001$, $\eta^2 = .31$, and a Position x Array Size interaction, $F(6, 114) = 3.51$, $p < .001$, $\eta^2 = .16$. A previous post-hoc test evaluated the main effect of Array Size, but Tukey's HSD post-hoc tests of Position revealed that percent correct was higher with Inner-Outer positions than Inner-Inner, $p < .001$ or Outer-Outer, $p < .001$, but not different from Outer-Inner, $p = .16$. The percent correct for
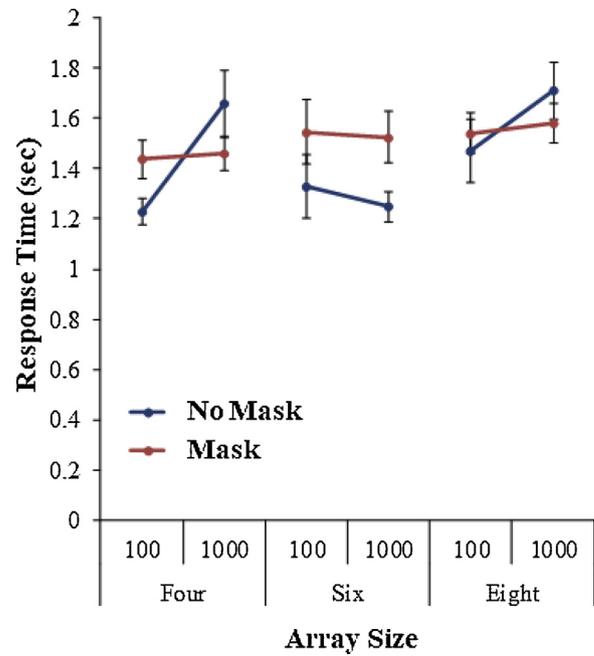
Outer-Inner was marginally higher than Outer-Outer, $p = .07$, all other comparisons were non-significant, $ps > .22$. Tukey's HSD post-hoc tests of the Position by Array Size interaction found no differences between positions for 4-item or 6-item trials, but on 8-item trials Inner-Inner differed marginally from Outer-Outer, $p = .08$, and Inner-Outer and Outer-Inner differed reliably from Outer-Outer, $ps < .01$. The GLMM analysis found the same main effect of Position, $X^2 (3) = 10.70$, $p < .05$, but was unable to test the interaction. The ANOVA with distance found the same main effect of Array Size, and an Array Size x Distance interaction, $F(4, 76) = 5.44$, $p < .001$, $\eta^2 = .67$, but no main effect of Distance, $F(2, 38) < 1.0$. Fig. 12 displays nearly symmetrically opposite patterns of percent correct across distance during 6-item and 8-item trials. The GLMM analysis produced identical results to the ANOVA (see Table 2).

The mean response times from correct trials displayed in Fig. 13 were longer for trials with a mask than no mask, and longer following a delay of 1000 than 100 ms. However, Array Size, Delay, and Mask all interacted with one another. The response times from correct trials with delays of 100 or 1000 (excluding trials with 0 delay) were log transformed (base 10) and analyzed with a RM ANOVA with Array Size (2, 3, and 4), Delay (100 and 1000), and Mask (0 and 1) as repeated measures. This analysis revealed all main effects were significant, $Fs > 9.68$, $ps < .01$, $\eta s^2 > .33$, as well as all interactions, $Fs > 4.37$, $ps < .05$, $\eta s^2 > .18$. Participants were slower to respond with a mask than no mask, and slower on trials with a longer delay (see Fig. 13). Tukey's HSD post-hoc tests of the main effect of Array Size revealed the same results found with percent correct; that is, no significant difference between 4-item and 6-item trials, $p = .46$, but response times for both were significantly shorter than for 8-item trials, $ps < .001$. The only consistent pattern in the remaining post-hoc tests was that response times were shorter for no mask, array size of four items, and a delay of 100 than all others, $ps < .05$, except for no mask, array size of six items at both delays, $ps > .98$. A GLMM with Array Size, Delay, and Mask as fixed effects and Subject as a random factor found identical results to the ANOVA (see Table 2). The three-way interaction could not be added without affecting converging.

The effect of the mask varied systematically across the 100 and 1000-ms delays. Consistent with the pigeons, the mask reduced percent correct at shorter delays but did not disrupt percent correct at the longer delay. No main effect of the mask was observed. As in Experiment 2a, percent correct on mask trials was higher than on non-mask trials with a 1000 ms delay for 4-item and 8-item trials. In fact, a post-hoc comparison test found percent correct was higher at 1000-ms than 100-ms when the mask was present. This was not found in Experiment 1b with pigeons. The main effect of array size revealed that performance during 4-item trials was similar to 6-item trials, which were both reliably different from 8-item trials. In contrast to Experiment 2a, the mask disrupted how fast (response times) and how well participants detected the changed item differently across the 100 and 1000 ms delays.

The increase in array sizes resulted in a mask effect for humans using similar parameters to that of Experiment 1b with pigeons. The motivation for this change from Experiment 2a was the previous findings that array size interacted with the mask effect. Interestingly, the data from Experiment 2b did not indicate this interaction. We may have reached the ceiling of this effect, which was present with fewer items in Experiment 2a.

8. General discussion

Considerable research has been conducted on the behavioral effects of delays in change-detection tasks. However, the majority of this research utilized one species (humans), and changes in the visual features of an item (e.g., color, shape, or size). The processes involved in visuospatial working memory are thought to be separable into visual, which involves the features of items, and spatial, which involves the positions of items (Jiang et al., 2008). The current research investigated whether pigeons would demonstrate effects of view time, array size, delay duration, training, and a visual mask comparable to that of human participants during a spatial change-detection task.

Experiments 1 and 2 reported effects of array size, delay, and a visual mask largely consistent with previous research using visual change-detection tasks. In Experiment 1a with pigeons, change detection accuracy decreased with more items in the array and longer delays, and training produced only a modest improvement. In Experiments 2a and 2b with humans, change detection accuracy also decreased with larger arrays, except array size was observed to have no effect on performance at the 0-s delay (c.f., Phillips, 1974). This reflects the property of high capacity within the iconic memory system. In pigeons, however, change detection accuracy decreased with larger arrays even on trials with a 0-s delay. This is unlikely to be a capacity or processing speed problem, given the dependence of birds on visual information during flight. This may reflect differences in how the displays are being processed due to the viewing angle of the subject. Pigeons are nearer to the display and may be unable to process the entire display, whereas humans view the display from a short distance that allows them access to the entire display.

In Experiments 1b and 2b the functional relationship between the effects of the mask and the delay was consistent across species and with extant literature, although humans required larger array sizes than pigeons to demonstrate the effect. Adding a visual mask decreased accurate detection of the changed item in both species only with the short delay. The data from pigeons indicated no reliable disruption at the longer delay, whereas for humans, the presence of the mask facilitated change detection accuracy in two of the three array sizes. This may be the result of the mask selected for this experiment. We utilized a gray and black checkerboard mask, similar to that of Phillips (1974). However, other mask types have been used, and it could be the case that the use of a checkerboard pattern provided a grid pattern on which participants could superimpose the colored circles, thereby facilitating subsequent detection of positions of the items with a longer delay.

One finding that was inconsistent with the literature for pigeons, but not humans, was the effect of delay on percent correct. In Experiment 1a, pigeons exhibited equivalent change detection accuracies with 0, 50, and 100-ms delays. To our knowledge, no published data from change-detection procedures with humans has reported this equivalence. However, similar findings have been reported with visual change-detection with pigeons (Wright et al., 2010) and the only other published study on spatial change-detection with pigeons (Leising et al., 2013). In the former study, performance actually improved from 50 to a 100-ms delay. Wright et al. argued that the gradual decline in accuracy from 50 to 400 ms was evidence against the influence of iconic memory in visual change-detection performance. The current data, however, suggests that whatever processes are supporting spatial change detection at 50 ms are the same as those operating at 0 and 100 ms, and perhaps out to 400 ms. Without additional data, it is unclear within what range of delays performance transitions from predominantly based on iconic memory to working memory, but the current data indicate it must be somewhere between 100 and 1000-ms.

It is also worth reiterating here the contrast between the effect of delay on percent correct in pigeons and humans. In contrast to the pattern found with pigeons, the data from humans indicate that performance after the 0-ms delay was reliably better than after a 100 ms delay, which did not differ from trials with the 1000-ms delay. This was the expected finding, with most studies finding masking effects with delays less than or equal to 100-ms (Phillips, 1974; Vogel et al., 2006). This finding suggests that while iconic memory is operating at 100 ms in humans (i.e., the masking effect in Experiment 2b), it plays less of a role for humans in detecting change than at the 0-s delay. This is consistent with another finding from Experiment 2b; mean percent correct at the 0-delay, but not 100-ms delay, did not decrease with larger array sizes.

It is unclear why we failed to find effects of view time in pigeons. Although these effects in humans are typically small (e.g., Pashler, 1988), they are reliable. There are data from pigeons demonstrating view time effects at shorter durations (Herbranson and Davis, 2016), such as between 30 and 200 ms, so the longer view times used in the current experiments (1000–3000 ms) may have been outside the range for an effect. Coltheart (1980) concluded that view time does not influence accuracy judgements based on iconic memory, but more data is needed to evaluate conflicting results in the literature. Lastly, the lack of a view time effect in the current research may again be the result of the viewpoint of the pigeon in the task. Whereas additional time for humans may allow for rapid shifts to other areas of the display, pigeons must shift their entire head to view another portion of the screen, and may process only small sections at a time. It might also be the case that pigeons simply do not shift their attention during the additional view time, but rather, perseverate on the same section of the screen thereby gleaning no benefit from the additional time.

Across Experiments 1b, 2a and 2b, both species were slower to respond correctly to the changed item on trials with a longer delay (1000 ms), and slower to respond on trials with a visual mask present. In Experiments 2a and 2b, the presence of the mask interacted with at least one other factor. Human participants were slower to respond on trials with a mask than trials with no mask only at the shorter delay (100 ms), but pigeons were slower to respond on trials with a mask at both delays (100 and 1000 ms). Interestingly, as the array size increased, the latency to respond increased for humans, but not for pigeons. There was, however, a delay by array size interaction for pigeons in Experiment 1b, such that the effect of delay on response time was greater with larger arrays. Response time appears to be a more sensitive measure for delay and mask effects in humans, as evidenced by main effects of both factors in Experiments 2a and 2b based on response time but not percent correct.

Cross-species evaluations of functional relationships are critical for claims of equivalence between species. Though the absolute value of parameters and dependent measures may vary, the relationship between the levels of the manipulation and behavior must be consistent.

In the current experiments, we tested how a variety of factors would influence the change detection performance of pigeons and humans. Most of the effects were consistent between species. The most central to the current manuscript, was the finding that a visual mask disrupts change detection accuracy for both species when a 100-ms delay occurs between the sample and test display, but does not disrupt performance following trials with a 1000-ms delay. This finding suggests similar perceptual mechanisms between species in the detection of change following a short delay (100 ms). In addition, these data also suggest that while the behavior of humans is less dependent on sensory, or iconic, memory at the shorter delay than with no delay, pigeons are equally dependent on the same perceptual processing mechanism during both no delay and 100 ms delay trials. Future research is needed to further investigate this hypothesis and identify the transition point between the two systems of memory in pigeons.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.beproc.2019.103957.

References

Baddeley, A., 2003. Working memory: looking back and looking forward. Nat. Rev. Neurosci. 4 (10), 829.

Bradley, C., Pearson, J., 2012. The sensory components of high-capacity iconic memory and visual working memory. Front. Psychol. 3, 355.

Carter, B.E., Henning, G.B., 1971. The detection of gratings in narrow-band visual noise. J. Physiol. 219 (2), 355–365.

Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., Haynes, J.D., 2017. The distributed nature of working memory. Trends Cogn. Sci. 21 (2), 111–124.

Courtney, S.M., Ungerleider, L.G., Keil, K., Haxby, J.V., 1996. Object and spatial visual working memory activate separate neural systems in human cortex. Cereb. Cortex 6 (1), 39–49.

Coltheart, M., 1980. Iconic memory and visible persistence. Percept. Psychophys. 27, 183–228.

De Valois, K.K., Switkes, E., 1983. Simultaneous masking interactions between chromatic and luminance gratings. JOSA 73 (1), 11–18.

Elmore, L.C., Ma, W.J., Magnotti, J.F., Leising, K.J., Passaro, A.D., Katz, J.S., Wright, A.A., 2011. Visual short-term memory compared in rhesus monkeys and humans. Curr. Biol. 21 (11), 975–979.

Eng, H.Y., Chen, D., Jiang, Y., 2005. Visual working memory for simple and complex visual stimuli. Psychon. Bull. Rev. 12 (6), 1127–1133.

Eriksson, J., Vogel, E.K., Lansner, A., Bergström, F., Nyberg, L., 2015. Neurocognitive architecture of working memory. Neuron 88 (1), 33–46.

Gibson, B., Wasserman, E., Luck, S.J., 2011. Qualitative similarities in the visual short-term memory of pigeons and people. Psychon. Bull. Rev. 18 (5), 979–984.

Grant, D.S., 1976. Effect of sample presentation time on long-delay matching in the pigeon. Learn. Motiv. 7 (4), 580–590.

Henning, G.B., Hertz, B.G., Hinton, J.L., 1981. Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. Noise masks. JOSA 71 (5), 574–581.

Herbranson, W.T., Davis, E.T., 2016. The effect of display timing on change blindness in pigeons (Columba livia). J. Exp. Anal. Behav. 105 (1), 85–99.

Hodos, W., Potocki, A., Ghim, M.M., Gaffney, M., 2003. Temporal modulation of spatial contrast vision in pigeons (Columba livia). Vision Res. 43 (7), 761–767.

Jiang, Y.V., Makovski, T., Shim, W.M., 2008. Visual memory for features, conjunctions, objects, and locations. The Visual World in Memory. Routledge, pp. 45–77.

Jitsumori, M., Ushitani, T., 2017. Rapid visual processing of picture stimuli by pigeons in an RSVP (rapid serial visual presentation) task. J. Exp. Psychol. Anim. Learn. Cogn. 43 (2), 127–138.

Jonides, J., Lewis, R.L., Nee, D.E., Lustig, C.A., Berman, M.G., Moore, K.S., 2008. The mind and brain of short-term memory. Annu. Rev. Psychol. 59, 193–224.

Kesner, R.P., Hunt, M.E., Williams, J.M., Long, J.M., 1996. Prefrontal cortex and working memory for spatial response, spatial location and visual object information in the rat. Cereb. Cortex 6 (2), 311–318.

Keysers, C., Xiao, D.K., Földiák, P., Perrett, D.I., 2005. Out of sight but not out of mind: the neurophysiology of iconic memory in the superior temporal sulcus. Cogn. Neuropsychol. 22 (3-4), 316–332.

Legge, G.E., Foley, J.M., 1980. Contrast masking in human vision. JOSA 70 (12), 1458–1471.

Leising, K.J., Elmore, L.C., Rivera, J.J., Magnotti, J.F., Katz, J.S., Wright, A.A., 2013. Testing visual short-term memory of pigeons (Columba livia) and a rhesus monkey (Macaca mulatta) with a location change detection task. Anim. Cogn. 16 (5), 839–844.

Losada, M.A., Mullen, K.T., 1995. Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking. JOSA A 12 (2), 250–260.

Morrison, A.B., Chein, J.M., 2011. Does working memory training work? The promise and challenges of enhancing cognition by training working memory. Psychon. Bull. Rev. 18 (1), 46–60.

Öğmen, H., Herzog, M.H., 2016. A new conceptualization of human visual sensory-memory. Front. Psychol. 7, 830.

Oliva, A., Wolfe, J.M., Arsenio, H.C., 2004. Panoramic search: the interaction of memory and vision in search through a familiar scene. J. Exp. Psychol. Hum. Percept. Perform. 30 (6), 1132.

Owen, A.M., Hampshire, A., Grahn, J.A., Stenton, R., Dajani, S., Burns, A.S., Ballard, C.G., 2010. Putting brain training to the test. Nature 465 (7299), 775.

Pashler, H., 1988. Familiarity and visual change detection. Percept. Psychophys. 44 (4), 369–378.

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

Rensink, R.A., O'Regan, J.K., Clark, J.J., 1997. To see or not to see: the need for attention to perceive changes in scenes. Psychol. Sci. 8 (5), 368–373.

Sekuler, R.W., 1965. Spatial and temporal determinants of visual backward masking. J. Exp. Psychol. 70 (4), 401.

Solomon, J.A., 2000. Channel selection with non-white-noise masks. JOSA A 17 (6), 986–993.

Sperling, G., 1960. The information available in brief visual presentations. Psychol. Monogr. Gen. Appl. 74 (11), 1–29.

Sperling, G., Melchner, M.J., 1978. The attention operating characteristic: examples from visual search. Science 202 (4365), 315–318.

Stromeyer, C., Julesz, B., 1972. Spatial-frequency masking in vision: critical bands and spread of masking. J. Opt. Soc. Am. 62, 1221–1232.

Truppa, V., De Simone, D.A., Mortari, E.P., De Lillo, C., 2014. Effects of brief time delays on matching-to-sample abilities in capuchin monkeys (Sapajus spp.). Behav. Brain Res. 271, 240–248.

Vogel, E.K., Woodman, G.F., Luck, S.J., 2006. The time course of consolidation in visual working memory. J. Exp. Psychol. Hum. Percept. Perform. 32 (6), 1436.

Wheeler, M.E., Treisman, A.M., 2002. Binding in short-term visual memory. J. Exp. Psychol. Gen. 131 (1), 48.

White, K.G., 1974. Temporal integration in the pigeon. Br. J. Psychol. 65 (3), 437–444.

Wilson, H.R., McFarlane, D.K., Phillips, G.C., 1983. Spatial frequency tuning of orientation selective units estimated by oblique masking. Vision Res. 23 (9), 873–882.

Wolfe, J.M., Oliva, A., Butcher, S.J., Arsenio, H.C., 2002. An unbinding problem? The disintegration of visible, previously attended objects does not attract attention. J. Vis. 2 (3), 5.

Wright, A.A., 2013. Functional relationships for investigating cognitive processes. Behav. Processes 93, 4–24.

Wright, A.A., Katz, J.S., 2006. Mechanisms of same/different concept learning in primates and avians. Behav. Processes 72 (3), 234–254.

Wright, A.A., Katz, J.S., Magnotti, J., Elmore, L.C., Babb, S., 2010. Testing pigeon memory in a change detection task. Psychon. Bull. Rev. 17 (2), 243–249.

Wright, A.A., Santiago, H.C., Sands, S.F., Kendrick, D.F., Cook, R.G., 1985. Memory processing of serial lists by pigeons, monkeys, and people. Science 229 (4710), 287–289.