Published in final edited form as:

J Exp Psychol Anim Learn Cogn. 2015 January ; 41(1): 32–38. doi:10.1037/xan0000050.

Monkey Visual Short-Term Memory Directly Compared to Humans

L. Caitlin Elmore and Anthony A. Wright

University of Texas Health Science Center Medical School at Houston

Abstract

Two adult rhesus monkeys were trained to detect which item in an array of memory items had changed using the same stimuli, viewing times, and delays as used with humans. Although the monkeys were extensively trained, they were less accurate than humans with the same array sizes (2, 4, & 6 items), with both stimulus types (colored squares, clip art), and showed calculated memory capacities of about one item (or less). Nevertheless, the memory results from both monkeys and humans for both stimulus types were well characterized by the inverse power-law of display size. This characterization provides a simple and straightforward summary of a fundamental process of visual short-term memory (how VSTM declines with memory load) that emphasizes species similarities based upon similar functional relationships. By more closely matching of monkey testing parameters to those of humans, the similar functional relationships strengthen the evidence suggesting similar processes underlying monkey and human VSTM.

Keywords

change detection; visual short-term memory; monkey; visual working memory

Visual short-term memory (VSTM) refers to the ability to transiently store visual information over brief time intervals of a few seconds or less. VSTM underlies numerous cognitive and motor functions including: detecting changes in the environment, planning and executing goal directed movements, and combining information across eye movements (e.g., Brouwer & Knill, 2007; Irwin, 1991; Henderson, 2008). For 25 years, the task of choice for investigating human VSTM has been *change detection* where participants view an array of visual stimuli and after a short delay report which stimulus changed or whether there was a change (Alvarez & Cavanagh, 2004; Cowan et al., 2001; Eng, et al., 2005; Luck & Vogel (1997); Pashler, 1988; Rensink, 2002). More recently, the emphasis has been on identifying the nature of VSTM limitations in terms of capacity or accuracy (Anderson et al., 2011; Bays & Husain, 2008; Donkin et al., 2013; Elmore et al., 2011; Gorgoraptis et al., 2011; Keshvari et al, 2013; Pashler, 1988; Rouder et al., 2008; Sims et al., 2012; van den Berg et al., 2012; Wilken & Ma, 2004; Zhang & Luck, 2008; Zhang & Luck 2011).

Correspondence concerning this article should be addressed to L. Caitlin Elmore, Department of Neuroscience, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. lelmore@cns.bcm.edu.

L. Caitlin Elmore, Department of Neuroscience, Baylor College of Medicine; Anthony A. Wright, Department of Neurobiology and Anatomy, University of Texas Health Science Center.

Despite extensive behavioral characterization of human VSTM, similar studies of nonhuman animal VSTM have only been conducted recently (Buschman et al., 2011; Elmore et al., 2011; Elmore et al., 2012; Gibson et al., 2011; Heyselaar, et al., 2011; Lara & Wallis, 2012; Wright et al., 2010). This lag in testing animal VSTM is, on one hand, surprising because of the importance in understanding how VSTM works in other species, discovering evolutionary continuity of VSTM in a nonhuman animal species, and establishing a model system for invasive (brain) studies of VSTM. On the other hand, meaningful behavioral characterizations of VSTM requires high task accuracy that is often difficult to achieve with animals—even for the rhesus monkey (*Macaca mulata*), the standard medical model.

Nevertheless, by developing training procedures, we (and others) have been able to achieve reasonably accurate performance from rhesus monkeys in VSTM tasks similar to some of those used to test humans (Buschman et al., 2011; Elmore et al., 2011; Elmore et al., 2012; Heyselaar, et al., 2011; Lara & Wallis, 2012). Although rhesus monkeys are typically not as accurate as humans in these similar tasks, they have shown some of the same signature trends of human VSTM, for example, progressive and systematic declines in accuracy as the number of to-be-remembered items is increased (e.g., Elmore et al., 2011; Heyselaar et al., 2011). For these species comparisons, we tested humans in the same basic task that we used to test rhesus monkeys. Nevertheless, some aspects of the task (e.g., stimuli, viewing times, delays) had to be made somewhat easier for monkeys to maintain accurate performance, and had to be somewhat more difficult to avoid ceiling effects with humans (Elmore et al., 2011). Despite an easier task, the monkeys were still less accurate than humans, revealing a VSTM capacity slightly less but close to one item, whereas humans showed a memory capacity of 3 items, a value similar to that shown by some other researchers (e.g., Eng, Chen, & Jiang, 2005; Cowan, 2001, Luck & Vogel, 1997). Despite these calculated capacity differences, performance across the different set sizes (number of items to be remembered) was well fit by the same theoretical function (based on signal detection theory) for both species. These similar functional relationships suggested similar VSTM processing, whereas accuracy level and capacity differences emphasized species' differences. A closer matching of monkey testing parameters to those used by humans would strengthen these comparisons.

Similar to our experience in training rhesus monkeys in other challenging memory tasks (e.g., visual and auditory list-memory tasks), continued training in the change-detection task helped to stabilize performance, allowing parameters (intervals, display sizes, etc.) to be gradually made equivalent to those used to test humans while still maintaining a reasonable performance level (e.g., ~70% correct). These matched parameters included the same visual angle, 1-s viewing times, 1-s delay times, and stimulus sets of 6 colored squares and 976 clip art images. Moreover, the monkeys were tested with five intermixed sample display sizes (instead of only three display sizes as in the Elmore et al., 2011 study), providing better constraints to the model fits and thereby strengthening monkey and human VSTM comparisons. The purpose of the present study was to better compare monkey and human VSTM using better matched change-detection tasks to better assess accuracy and capacity differences and VSTM functional relationship similarities.

Methods

Subjects

Two adult male rhesus monkeys, Cisco and Captain, participated in the experiment. They were 10 and 14 years old respectively at the start of the experiment. Both monkeys had extensive prior experience in the change detection task (see Elmore et al. 2011; Elmore et al. 2012 for details). The monkeys were housed in individual cages in a primate colony room with several other monkeys. They were tested 5–6 days per week in periods that did not exceed two hours in duration. The monkeys received their daily ration of primate chow and water following testing. Supplemental fruits and vegetables were provided on non-testing days. During the task, the monkeys received reinforcement in the form of banana pellets (Bio-Serv, 300-mg, Frenchtown, NJ) and Cherry Koolaid. The two types of reinforcers were delivered following correct responses. Reinforcer type was allocated pseudorandomly according to the individual monkey's preference. Cisco received 70% Koolaid and 30% pellets and Captain received 60% Koolaid and 40% pellets. All animal procedures conformed to National Institutes of Health guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Texas Health Science Center at Houston.

Apparatus

Chambers—Custom-made aluminum test chambers were used to test the monkeys. The chambers were 47.5 cm wide $\times$ 53.13 cm deep $\times$ 66.25 cm high. The monkeys were unrestrained and free to move about within the confines of the test chamber. A sound machine (Homedics, Commerce Township, MI) located outside of the chamber was used to produce white noise to mask extraneous noise. A 17" computer monitor (EIZO) equipped with an infrared touchscreen (Unitouch, ELO, Round Rock, TX) was fitted in the back wall of the chamber 30 cm above the chamber floor on which the monkeys sat. The touchscreen was used to detect touch responses to the computer monitor. On the left side of the back wall of the chamber, 14 cm below the touchscreen, was a pellet cup (5.6 cm in diameter, 2.5 cm deep) which received delivery of banana pellets from a pellet dispenser (Gerbrands, G5-120, Arlington, MA) located outside of the chamber. Cherry-Koolaid was dispensed via plastic tubing to a metal spout located 8 cm below the touch screen on the right side of the chamber's back wall.

Stimuli & Displays

The stimuli were either six 4-cm wide colored squares (Colors with RGB 24 bit values: aqua – 0, 255, 255, blue – 0, 0, 255, green – 0, 255, 0, magenta – 255, 0, 255, red – 255, 0, 0, yellow – 255, 255, 0), or 976 different clip art images (example images are shown in Figure 1). Each of the clip art images fit within a 4 $\times$ 4 cm square. The stimuli were presented in random locations on an invisible 4 $\times$ 4 matrix (26 $\times$ 22 cm). This matrix was aligned to a clear Plexiglas response template which had 16 circular cutouts, each of which was 4 cm in diameter. Based on an estimate of the monkeys' average distance from the screen, the stimuli subtended a visual angle of 5.75 degrees.

Test Procedures

The task progression is illustrated in Figure 1. Each session began with the presentation of stimuli on the computer monitor (sample display). The critical parameter that was varied was display size (number of stimuli to remember per trial). Within each session, display sizes of two, three, four, five, and six items were intermixed. The monkeys were tested with 20 alternating 96-trial blocks of either colored squares or clip art. Each 96-trial block contained just one of the two stimulus types. The stimuli were viewed for a variable viewing times ranging from 1000–5000ms (in 500 ms increments) to encourage vigilance during the session. Trials with a viewing time of 1000-ms constituted the majority (56%) of trials tested in across sessions and were the only trials included in the analyses presented here, the other 44% trials were split between the other 8 viewing times (1500-ms, 2000-ms, 2500-ms, 3000-ms, 3500-ms, 4000-ms, 4500-ms, and 5000-ms). Blocks were not identical in the number of trials at each viewing time, as 96 trials do not divide evenly among 9 viewing times, but over the course of the 20 blocks, the monkeys completed 1080 trials with 1000-ms viewing times and 840 trials split between the other 8 viewing times (105 each). Next, the stimuli disappeared and the chamber was darkened for a 1000-ms delay interval.

Following the delay, two stimuli appeared on the screen, one of which matched (in identity and location) one of the stimuli from the original sample display, and the other was presented in a location previously occupied by a stimulus during the sample display, but had changed in identity. The monkey's task was to touch the stimulus that had changed. After the monkeys' response, if they responded correctly, a tone was presented followed by reinforcement. Pellets were dispensed immediately, but there was a brief delay before juice reward was dispensed to allow time for the monkeys to access the juice spout. If an incorrect response was made, a click noise was heard and no reinforcement was provided. Next, the chamber was illuminated by a green 25-watt bulb located outside the chamber for a 15-s intertrial interval. The light entered the chamber through a small gap between the touchscreen and monitor. At the end of the 15-s intertrial interval, the chamber was darkened and the next trial began.

Results and Discussion

As shown in the upper panel of Figure 2, percent correct performance declined as a function of display size. Separate repeated measures ANOVAs of display size $\times$ stimulus type showed a significant effect of display size for both monkeys [M1: $F(4, 36) = 11.09, p < 0.001$, partial $\eta^2 = 0.33$, 95% CIs = 0.15, 0.44; M2: $F(4, 36) = 7.59, p < 0.001$, partial $\eta^2 = 0.25$, 95% CIs = 0.09, 0.36], as well as a significant effect of stimulus type [M1: $F(1, 9) = 20.67, p < 0.001$, partial $\eta^2 = 0.19$, 95% CIs = 0.06, 0.32; M2: $F(1, 9) = 12.48, p < 0.001$, partial $\eta^2 = 0.13$, 95% CIs = 0.02, 0.25]. The monkeys performed more accurately with clip art than with colored squares (mean difference in percent correct was 10.57%).

Testing Models of VSTM

Fixed-Capacity Model—VSTM capacity was calculated using Equation 1, originally developed by Eng et al. (2005).

$$A = \left[\frac{N-C}{N} \right]^2 \times 50\% + \left\{ 1 - \left[\frac{N-C}{N} \right]^2 \right\} \times 100\% \quad \text{Equation 1}$$

In this equation, A is the empirical accuracy, N is the display size tested, and C is VSTM capacity. The likelihood that a single test item was not among the C items remembered is $(N-C)/N$, and the likelihood that both test items were not among the C items remembered is $[(N-C)/N]^2$. If both test items were among the C items remembered, then accuracy is 100% whereas if neither were among the C items remembered, then accuracy is 50%, because the participant would be guessing. Thus, the equation above can be used to compute capacity, given that $N > C$. For each monkey, each display size, and both stimulus types, capacity was computed by solving for C . Mean capacity estimates for both stimulus types and each display size are shown in Figure 2 – middle panel. By averaging across display sizes and monkeys, the mean capacities for the two stimulus types were computed. Overall, mean capacity for colored squares was 0.33 ± 0.10 (S.E.M.) and mean capacity for clip art was 0.84 ± 0.08 . These capacity estimates are somewhat smaller than those in our previous experiment (Elmore et al., 2011), where capacity was estimated to be 0.71 ± 0.24 and 1.02 ± 0.19 for colored squares and clip art respectively. The present data indicate that according to the fixed-capacity model of VSTM, monkeys were on the average maintaining substantially less than one colored-square stimulus (and more or less one clip-art stimulus) in memory during the delay interval.

However, it is worth noting that Equation 1's capacity formula assumes that if subjects do not remember one of two sample display items being tested, but remember the other one was the same as the sample item, then they can *infer* that the not-remembered item must be the one that changed. However, it is unknown whether or not monkeys make this inference. If they do not make this inference, then it could be assumed that the correct response is made if the changed item was remembered (probability of C/N), or by a correct guess of 0.5 (we are indebted to Nelson Cowan for this suggestion). In this case, Equation 1 would become:

$$A = \left[\frac{C}{N} + \left(1 - \frac{C}{N} \right) * 0.5 \right] * 100 \quad \text{Equation 2}$$

Equation 2 actually produces larger capacity estimates, since correct answers are based on VSTM for changes plus guessing (but not inference) when the VSTM does not detect a change. With this equation, in the colored square condition, the monkeys' mean capacity would be 0.61 ± 0.27 (instead of 0.33 ± 0.10), and in the clip art condition mean capacity would be 1.44 ± 0.13 (instead of 0.84 ± 0.08), thus producing a modest rise in the maximum capacity estimates for monkeys. In either case, if roughly one item is the VSTM capacity limit of monkeys, then survival by avoiding predators, finding and remembering food sources and successful troop interactions (mates and dominance hierarchy) might be in jeopardy. But of course rhesus monkeys adapt and survive very well. The resolution to this capacity issue does not likely lie with a lack of ecological validity of the stimuli used in this experiment. Performance of 85% correct for the two clip-art condition is very good accuracy in these kinds of tasks, by any nonhuman species. Perhaps the monkeys remembered two items perfectly on some trials, but at other times remembered no items (inattention),

resulting in a mean capacity that vacillates around 1 item (we are indebted to an anonymous reviewer for this suggestion). And yet another possibility, considered in the next section, is that all memory items are encoded and remembered imperfectly—for example imperfect memory might be distributed across many if not all items to-be-remembered. Furthermore, distribution of a limited memory resource would result in less perfect memory as the number of to-be-remembered items increased.

Continuous-Resource Model—The continuous-resource model employs d' values from signal detection theory as a measure of memory sensitivity (Wilken & Ma, 2004; Green & Swets, 1966; MacMillan & Creelman, 2005, Elmore et al., 2011). d' was computed using Equation 3.

$$d' = \frac{[z(H) - z(FA)]}{\sqrt{2}} \quad \text{Equation 3}$$

The difference of the z scores of the hits and false alarms are divided by the square root of 2. The division by square root of 2 is necessary because the task is a two-alternative forced-choice task (2AFC) and there are two ways to make a correct response: by remembering that one is the same as the sample display (and choosing the other), or by noticing the object that has changed and choosing it (MacMillan & Creelman, 2005). Hits and false alarms were defined based on stimulus location in the test display. Locations were numbered from 1 to 16 as the locations went from left to right and then down into the row below, such that the bottom right corner was location number 16. A hit was defined as a correct response to the lower numbered location in the test display. So if test stimuli were displayed in locations 2 and 9 and the stimulus in 2 was the changed object, a correct response to location 2 would constitute a *hit*. A *false alarm* was defined as a response to the lower numbered location when that location did not contain the changed item. The definitions of a “hit” and a “false alarm” are arbitrary but equivalent to the obverse. d' values for each stimulus type and display size are plotted in Figure 2 – lowest panel. There was little or no response bias for colored squares and clip art, in keeping with findings from humans in two-alternative forced-choice as opposed to yes-no procedures (Green & Swets, 1966, p. 408).

The d' values for each stimulus type were fit with power law functions. As anticipated, power law functions were found to be good fits to each individual monkey's d' values as well as to the group mean, for both stimulus types. Power law functions should provide good fits according to the continuous-resource model, because memory sensitivity (d') should be proportional to $1/N$, where N is the number of items in the display. The general form of the power law function is shown in Equation 4, where Y is a constant, N is the number of items in the sample display, and x is the exponent.

$$d' = Y * N^{-x} \quad \text{Equation 4}$$

Individual power law fits revealed r^2 values of 0.74 and 0.66 for clip art, and 0.72 and 0.84 for colored squares for Captain and Cisco, respectively. The group power law fits were statistically significant, [Colored Squares: $F(1,8) = 81.80, p < 0.0001$, Cohen's $f^2 = 10.13$,

95% CIs = 6.0, 26.0; Clip Art: $F(1,8) = 43.57$, $p < 0.001$, Cohen's $f^2 = 5.45$, 95% CIs = 3.1, 13.9]. The mean power law functions provided good fits to the group data with r^2 values of 0.86 and 0.94 for clip art and colored squares, respectively. These fits are comparable to the fits in our prior experiment (Elmore et al., 2011) where r^2 values were 0.99 and 0.98 for clip art and colored squares, respectively.

Comparisons to Humans

For comparison, the results for humans from our prior study are shown in Figure 3 (Elmore et al., 2011). Much like the monkeys from the study presented here (Figure 2), performance declined as a function of display size. On average, the humans outperformed the monkeys by 19.2% for clip-art trials and 27.3% for colored-square trials. A repeated measures ANOVA of display size (2, 4, and 6 only) $\times$ stimulus type $\times$ species revealed a main effect of display size [$F(1,6) = 8.81$, $p = 0.01$, partial $\eta^2 = 0.72$, 95% CIs 0.59, 0.86] and species [$F(2,12) = 12.15$, $p = 0.01$, partial $\eta^2 = 0.63$, 95% CIs 0.53, 0.75].

Capacity estimates also differed between the two species. The humans' capacity estimates were larger with means of 2.8 ± 0.4 for clip art and 2.5 ± 0.4 for colored squares. Thus, the humans could remember approximately 2 more items than the monkeys according to a fixed capacity model of VSTM. The human capacity estimates are somewhat lower than what is often reported (magic number 4 ± 1 – Cowan, 2001), but they are within the range typically found for human subjects (e.g., Vogel & Machizawa, 2004). Thus, the fixed-capacity model provides a reasonable framework for understanding human VSTM, but appears to be somewhat less meaningful for monkeys.

Monkey and human performances were also compared using d' values according to the continuous-resource model. The d' values for both species are well characterized by power law functions, and the exponents of these functions are in a similar range. The exponents were -0.70 and -1.02 for clip art, and -0.94 and -0.86 for colored squares, for monkeys and humans, respectively. An unpaired t -test showed that there were no significant differences in exponent value across species [$t(14) = 1.54$, $p = 0.15$, Hedge's $g = 0.84$, 95% CIs = 0.42, 2.01]. Not surprisingly, however, the coefficients of the power law functions were significantly greater for humans [unpaired t -test, $t(14) = 2.26$, $p = 0.04$, Hedges' $g = 1.23$, 95% CIs = 0.09, 2.43], indicating a greater overall memory sensitivity (d') for humans than monkeys.

Conclusions

We have shown that rhesus monkeys can be trained and tested in change-detection tasks with parameters closely matched to those used to test humans for more direct VSTM comparisons. Humans, of course, come to the task with a lifetime of game-playing and test-taking experience and therefore require little or no training, except instructions, to reach asymptotic accuracy. Monkeys, on the other hand, learn the 'rules of the task' through the contingencies of reinforcement. With continued training, the rhesus monkeys in this study learned a generalized rule of change (Elmore et al., 2012). They maintain accurate performance similar to that shown in a previous study even though the task was made progressively more difficult by gradually changing parameters (visual angle, 1-s viewing

times, 1-s delay times, and stimulus sets of 6 colored squares and 976 clip art images) to match those used to test humans (Elmore et al., 2011). In the Elmore et al. (2011) study, these monkeys were tested with 50-ms delays, whereas in the present study they were tested with 1000-ms delays. We observed no major differences in performance between these two delays and found similar accuracies for groups of intermediate delay values. With too short of a delay (less than a few hundred milliseconds) there has been in the past (e.g., Pashler, 1988) a concern of attentional capture governing performance. This is why we chose to use the 1000-ms delay; to exclude attentional capture as opposed to change detection based on VSTM.

Findings and conclusions from the better matched experiment of this article support and strengthen many of the conclusions from our previous experiment (Elmore et al., 2011). Among these conclusions is that both monkeys and humans have very limited VSTM. They are only able to accurately remember a small amount of information over the course of a brief delay, and as such, VSTM accuracy declines with display size. Said otherwise, the more items that the participants must remember, the lower is their accuracy. The experiment presented in this article also highlights and supports monkey and human VSTM differences shown in these experiments. The monkeys were less accurate than humans at similar set-sizes (2, 4, & 6 items) and showed calculated capacities of less than one item with one of the item types (colored squares), similar to what they had previously been shown for colored circles (Elmore et al., 2011).

Implications of such findings were more straightforward when the earlier experiment (Elmore et al., 2011) was published, and the results were summarized as supporting a continuously distributed resource account of VSTM as opposed to all-or-nothing fixed capacity account. Since that publication, other theories and models have been proposed that blend these accounts, plus other factors (e.g., attention lapses) that have been added—all of which make model distinctions more difficult, particularly with the procedures we used in this experiment (e.g., Donkin et al., 2013; Gorgoraptis et al., 2011; Keshvari et al, 2013; Sims et al., 2012; van den Berg et al., 2012; van den Berg, Awh, & Ma, 2014; Zhang & Luck, 2011).

Despite differences in accuracy and capacity that emphasize species differences, a continuous-resource account provides a simple and straightforward explanation based upon similar functional relationships which emphasizes species similarities. Memory sensitivity (d') is specified to decline precisely as an inverse power law function of N (display size). The d' values from both species and both stimulus types were well fit by these power law functions. Even with the additional constraints of five display sizes for monkeys (similar to humans), the continuous-resource model and resulting functional relationships still account for 85% of the variance. By closer matching of monkey testing parameters to those of humans, conclusions based upon the similar functional relationships shown here strengthen the evidence for similar VSTM processing between monkeys and humans.

Acknowledgments

Support for this research was provided by NIMH Grants R01MH-072616 and R01MH091038 (A. A. Wright). We thank Dr. Wei Ji Ma for his contributions to this research. This research was conducted following the relevant ethics guidelines for research with animals and was approved by UTHSC's institutional IACUC.

References

- Alvarez GA, Cavanagh P. The capacity of visual short-term memory is set by both visual information load and by number of objects. *Psychological Science*. 2004; 15:106–111. [PubMed: 14738517]
- Anderson DE, Vogel EK, Awh E. Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal of Neuroscience*. 2011; 31:1128–1138. [PubMed: 21248137]
- Bays PM, Husain M. Dynamic shifts of limited working memory resources in human vision. *Science*. 2008; 321:851–854. [PubMed: 18687968]
- Brouwer AM, Knill DC. The role of memory in visual guided reaching. *Journal of Vision*. 2007; 7:1–12. [PubMed: 18217846]
- Buschman TJ, Siegel M, Roy JE, Miller EK. Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Science*. 2011; 108:11252–11255.10.1073/pnas.1104666108
- Cowan N. The magic number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Science*. 2001; 24:87–114.
- Donkin C, Nosofsky RM, Gold JM, Shiffrin RM. Discrete slot models of visual working-memory response times. *Psychological Review*. 2013; 4:873–902. [PubMed: 24015956]
- Elmore LC, Ma WJ, Magnotti JF, Leising KJ, Passaro AD, Katz JS, Wright AA. Visual Short-Term Memory Compared in Rhesus Monkeys and Humans. *Current Biology*. 2011; 21:975–979.10.1016/j.cub.2011.04.031 [PubMed: 21596568]
- Elmore LC, Magnotti JF, Katz JS, Wright AA. Change detection by rhesus monkeys (*Macaca mulatta*) and pigeons (*Columba livia*). *Journal of Comparative Psychology*. 2012; 126:203–212.10.1037/a0026356 [PubMed: 22428982]
- Eng HY, Chen D, Jiang Y. Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin and Review*. 2005; 12:1127–1133. [PubMed: 16615339]
- Gibson B, Wasserman E, Luck SJ. Qualitative similarities in the visual short-term memory of pigeons and people. *Psychonomic Bulletin & Review*. 2011; 18:979–984.10.3758/s13423-011-0132-7 [PubMed: 21748417]
- Gorgoraptis N, Catalao RG, Bays PM, Husain M. Dynamic updating of working memory resources for visual objects. *The Journal of Neuroscience*. 2011; 31:8502–8511. [PubMed: 21653854]
- Green, DM.; Swets, JA. *Signal Detection Theory and Psychophysics*. New York: Wiley; 1966.
- Henderson, JM. *Visual Memory*. Oxford: Oxford University Press; 2008. Eye movements and visual memory; p. 87-121.
- Heyselaar E, Johnston K, Pare M. A change detection approach to study visual working memory of the macaque monkey. *Journal of Vision*. 2011; 11:1–10.10.1167/11.3.11
- Irwin DE. Information integration across saccadic eye movements. *Cognitive Psychology*. 1991; 23:420–456. [PubMed: 1884598]
- Keshvari S, van den Berg R, Ma WJ. No evidence for an item limit in change detection. *PLoS Computational Biology*. 2013; 10.1371/journal.pcbi.1002927
- Lara AH, Wallis JD. Capacity and precision of an animal model of visual short-term memory. *Journal of Vision*. 2012; 12:1–12.10.1167/12.3.13
- Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature*. 1997; 390:279–281. [PubMed: 9384378]
- Macmillan, NA.; Creelman, CD. *Detection Theory*. USA: Lawrence Erlbaum Associates, Inc; 2005.
- Pashler H. Familiarity and visual change detection. *Perception and Psychophysics*. 1988; 44:369–378. [PubMed: 3226885]

- Rensink RA. Change detection. *Annual Review of Psychology*. 2002; 53:245–277.
- Rouder JN, Morey RD, Cowan N, Zwilling CE, Morey CC, Pratte MS. An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*. 2008; 105:5975–5979.
- Sims CR, Jacobs RA, Knill DC. An ideal observer analysis of visual working memory. *Psychological Review*. 2012; 119:807–830. [PubMed: 22946744]
- Van den Berg R, Awh E, Ma WJ. Factorial comparison of working memory models. *Psychological Review*. 2014; 121:124–149. [PubMed: 24490791]
- Van den Berg R, Shin H, Chou WC, George R, Ma WJ. Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*. 2012; 109:8780–8785.
- Vogel EK, Machizawa MG. Neural activity predicts individual differences in visual working memory capacity. *Nature*. 2004; 428:748–751. [PubMed: 15085132]
- Wilken P, Ma WJ. A detection theory account of change detection. *Journal of Vision*. 2004; 4:1120–1135. [PubMed: 15669916]
- Wright AA, Katz JS, Magnotti J, Elmore LC, Babb S, Alwin S. Testing pigeon memory in a change detection task. *Psychonomic Bulletin and Review*. 2010; 17:243–249.10.3758/PBR.17.2.243 [PubMed: 20382927]
- Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. *Nature*. 2008; 453:233–235. [PubMed: 18385672]
- Zhang W, Luck SJ. The number and quality of representations in working memory. *Psychological Science*. 2011; 22:1434–1441. [PubMed: 21987693]

Figure 1.
Progression of events in the change detection task.

Figure 2.

Monkey change-detection performance and models. Top panel: Percent correct in the change detection task with Clip Art and Colored Squares. Middle panel: Capacity estimates calculated based on change detection performance. Lower panel: Power law fits for d' values calculated based on change detection performance. Error bars represent standard error of the mean.

Figure 3.

Human change detection performance and models. Top panel: Percent correct in the change detection task with Clip Art and Colored Squares. Middle panel: Capacity estimates calculated based on change detection performance. Lower panel: Power law fits for d' values calculated based on change detection performance. Error bars represent standard error of the mean.