

UC Davis

UC Davis Previously Published Works

Title

A note on the identification of change detection task models to measure storage capacity and attention in visual working memory

Permalink

<https://escholarship.org/uc/item/0642557g>

Journal

Behavior Research Methods, 51(3)

ISSN

1554-351X

Authors

Feuerstahler, Leah M
Luck, Steven J
MacDonald, Angus
et al.

Publication Date

2019-06-01

DOI

10.3758/s13428-018-1082-z

Peer reviewed



Published in final edited form as:

Behav Res Methods. 2019 June ; 51(3): 1360–1370. doi:10.3758/s13428-018-1082-z.

A Note on the Identification of Change Detection Task Models to Measure Storage Capacity and Attention in Visual Working Memory

Leah M. Feuerstahler, Steven J. Luck, Angus MacDonald III, and Niels G. Waller

LMF, University of California-Berkeley; NGW and AMIII, University of Minnesota, Twin-Cities; SJL, Department of Psychology, University of California-Davis

Abstract

The change detection task is a common method for assessing the storage capacity of working memory, but estimates of memory capacity from this task can be distorted by lapses of attention. When combined with appropriate mathematical models, some versions of the change detection task make it possible to separately estimate working memory and the probability of attentional lapses. In principle, these models should allow researchers to isolate the effects of experimental manipulations, group differences, and individual differences on working memory capacity and on the rate of attentional lapses. However, the present research found that two variants of a widely accepted model of the change detection task are not mathematically identified.

Working memory plays a key role in many large-scale theories of cognition (Anderson, Matessa, & Lebiere, 1997; Meyer & Kieras, 1997), and individual differences in the storage capacity of working memory are widely thought to play a major role in determining individual and group differences in cognitive ability (Baddeley, 2003; Cowan, 2012; Engle, Tuholski, Laughlin, & Conway, 1999). The visual change detection paradigm has become a popular approach for estimating working memory capacity (Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Cusack, Lehmann, Veldsman, & Mitchell, 2009; Luck & Vogel, 2013). In the canonical version of this paradigm (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001), each trial consists of a sample array of visually presented objects followed by a short delay and then a test array (see Figure 1A). The sample and test arrays are identical except that, on some percentage of trials (typically 50%), one item differs between the sample and test arrays. On each trial, participants make an unspeeded two-alternative forced-choice response to indicate whether a change was present or absent. Working from this general paradigm, Pashler (1988, p. 370) developed a simple equation for estimating the number of items stored in memory. The logic behind this equation is that, if the participant has a working memory capacity of K items and the array contains S items, there will be a K/S probability that the changed item is one of the items being held in working memory and that the change will therefore be detected. This equation has become widely used to assess how memory storage capacity, K , varies across experimental conditions, across groups of participants, and across individuals within a group (Alvarez & Cavanagh, 2004; Cowan et

al., 2006; Cusack et al., 2009; Johnson, McMahon, Robinson, Harvey, Hahn, Leonard, Luck, & Gold, 2013).

One advantage of the change detection task is that—when used in conjunction with appropriate mathematical models (Rouder, Morey, Cowan, Zwillling, Morey, & Pratte, 2008)—it can be used to measure attentional engagement concurrently with working memory storage capacity. By *attentional engagement*, we mean that the participant is actively encoding the stimuli and following the rules of the task. When attention is not engaged, a lapse of attention occurs. Following previous work (Rouder et al., 2008; Rouder, Morey, Morey, & Cowan, 2011), we assume that a participant is either “on task” or “off task” on a given trial. When participants are on task, they (a) perceive the sample array, (b) encode as much of that information as possible, (c) maintain as much of this information as possible over the delay, (d) compare the memory of the sample array to the test array to arrive at a decision, and (e) make the appropriate response for this decision. Alternatively, when they are off task, they fail at one or more of the aforementioned task components.

In recent years, many variants of the change detection task have been described in the literature, and quantitative methods have been developed to estimate working memory capacity and attentional engagement in at least three of these variants (Gibson, Wasserman, & Luck, 2011; Rouder et al., 2008; Rouder et al., 2011). These models have not been widely used in the empirical literature, in which researchers often simply estimate capacity without considering lapses of attention (Adam, Mance, Fukuda, & Vogel, 2015; Fukuda, Vogel, Mayr, & Awh, 2010; Luria & Vogel, 2011). However, there is growing evidence that mind wandering and other forms of attentional disengagement play a significant role in both typical and disordered cognition (Erickson, Albrecht, Robinson, Luck, & Gold, 2017; Mrazek, Smallwood, Franklin, Chin, Baird, & Schooler, 2012), and future research on individual and group differences in working memory would benefit greatly from models that can separately estimate capacity and attentional engagement.

The present note demonstrates that attentional engagement is not necessarily estimated accurately in models of two of three variants of the change detection task. In these variants, the underlying mathematical model is not identified, meaning that different parameter values will predict the same pattern of behavioral performance. Stated differently, in these models, widely different levels of attentional engagement may be consistent with a given pattern of observed data. This occurs because the attentional engagement parameter depends mathematically on the two guessing parameters. In the remainder of this paper, we prove that these models are not identified and we demonstrate the consequences of this inconvenient mathematical truth for model estimation and interpretation. We then describe simple model modifications that ensure parameter identification. We end by describing the overall strengths and weaknesses of the three variants of the change detection task with respect to psychological factors that may impact their validity.

The models considered in this study are mathematical instantiations of a class of theories of visual working memory that are often called “slot” or “item-limit” theories (reviewed by Luck & Vogel, 2013). These theories assume that there is a limit on the number of items that can be stored in working memory; when the sensory input exceeds this limit, a subset of

items are stored with some reasonable precision, and no information is retained about the other items. In contrast, “resource” theories assume that all items are stored in working memory, but that the quality of the individual item representations declines as the number of items in memory increases because of resource limitations or noise (Bays, 2014; van den Berg, Awh, & Ma, 2014). These two classes of theories have been hotly debated, and the field has not reached a consensus. Importantly, researchers in a variety of domains have a pressing need to be able to estimate working memory capacity and attentional engagement, and this research cannot wait decades for memory experts to reach unanimity. Researchers should certainly not use parameter estimates from models that are not mathematically identified. Thus, the present note is timely even if the debate between item-limit theories and resource theories has not been completely resolved.

Three Variants of the Change Detection Task

Figure 1 shows the three variants of the change detection task that we consider in this report. The first variant (Luck & Vogel, 1997; Vogel et al., 2001) is illustrated in Figure 1A. We call this model the *canonical change detection* task because it was the first variant to become widely used (and it is still the most commonly used of the three variants). The second variant limits the decision to a single item in the test array (Figure 1B), which is typically accomplished either by presenting a postcue around the one item in the test array that might have changed or by presenting only one item in the test array (see Experiments 4–6 in Vogel et al., 2001). We call this the *partial report change detection* task because it is analogous to the partial report paradigm made famous in Sperling’s classic iconic memory experiments (Sperling, 1960). The third variant, called the *multiple change detection* task, is similar to the canonical change detection task, except that multiple items can change on a given trial (Figure 1C). Rather than varying the set size, this task fixes the number of items in the display and varies the number of items that change (Gibson et al., 2011). However, the task remains the same as in the other variants; participants report whether a change was present (irrespective of the number of changes) or absent.

Figure 2 illustrates the pattern of data expected from the model of the canonical change detection task. In this model, the probability of reporting a change is a function of S (the set size), whether a change actually occurred, and four person parameters: the individual’s level of storage capacity (K), attentional engagement (A), and two guessing parameters (G and U). This figure shows trial response probabilities for five hypothetical individuals with K values that range from 1 to 5, $A = .85$, $G = .15$, and $U = .35$. In general, the probability of detecting a change on change-present trials declines systematically as the set size (S) increases beyond a given individual’s storage capacity (K). At this point in the manuscript, it is not important to understand the other parameters. Rather, it is sufficient to recognize that under the canonical change detection paradigm, trial response probabilities can be modeled as a function of set size, whether or not a true change occurred, and the previously described four person parameters. Notice in Figure 2 that the probability of a change response on change-absent trials does not depend on set size or storage capacity.

The general logic for estimating attentional engagement (Gibson et al., 2011; Rouder et al., 2008; Rouder et al., 2011) is that participants will attend to the task on a certain proportion

of trials (denoted A), and they will have lapses of attention for the remaining proportion of trials ($1-A$). When participants are attending (i.e., they are on task), they are expected to exhibit perfect performance if the set size is less than or equal to their storage capacity (i.e., when $S \leq K$). Accuracy is predicted to fall off at higher set sizes. When participants lapse (i.e., they are off task), they presumably guess randomly and are predicted to produce a constant error rate at all set sizes. Pashler's (1988, p. 370) original equation for estimating working memory storage capacity (K) assumed that the participant was consistently on task and would respond correctly on a change-present trial if a change was present and if the changed item was in working memory. However, because a change might occur in an item that was not stored in memory, it would be reasonable for participants to occasionally guess that a change was present even if none of the items in memory changed. In Pashler's model, this guessing rate is denoted G . Pashler assumed that G could be estimated by the probability of a change-present response on trials without a change. Cowan (2001) created a modification of this equation for the partial report variant of the task.

The Pashler (1988) and Cowan (2001) models did not take into account the possibility of lapses of attention. Rouder et al. (2011) extended these models to include the possibility of lapses, assuming that participants guess when they are off task (see also Rouder et al., 2008). Rouder et al. (2011) used the letter U to denote the probability of a change-present guess when a participant is off task (i.e., when a participant is not attending to the task). In this scheme, U refers to *uninformed* guessing and G refers to *informed* guessing. When a participant is on task, guesses are informed by factors such as the number of items in the array and the fact that no changes were detected in the K items that were present in working memory. The guessing rate when a participant is on task (G) may be quite different from the guessing rate when a participant is experiencing attentional lapses (U). Importantly, lapses in attentional engagement may occur for many reasons such as an ill-timed blink that prevents perception of the sample array, mind wandering that prevents encoding of the sample array in memory, or response-mapping confusions at the time of report. In the partial report change detection model, all guessing is considered uninformed guessing (Rouder et al., 2011). In other words, the partial report change detection model includes U but not G . In this model, if the probed item is not in working memory due to capacity limits, then the subject presumably guesses in the same manner as when the item is not in working memory due to an attentional lapse.

As will be proven below, in the models of the canonical and multiple change detection tasks, the attention (A), and two guessing (G , U) parameters are not mathematically identified without additional constraints. In less formal language, this implies that different values of these parameters can produce the same pattern of responses on a change detection task (and a given set of observed data is therefore consistent with multiple combinations of parameter values). In the next section we demonstrate this point mathematically.

Model Underidentification in the Canonical and Multiple Change Detection Tasks

Both the canonical and multiple change detection tasks include four individual-level parameters. In the following mathematical equations, no subscripts are used for the individual or trial to avoid notational clutter. As described earlier, the four person parameters of these models are:

- K : capacity, the number of objects held in visual working memory,
- A : attention, the probability that attention is engaged on a trial ($0 \leq A \leq 1$),
- G : informed guessing, the probability of correctly guessing that an object changed, on a given trial, if that object is not in working memory when the participant is on task ($0 \leq G \leq 1$), and
- U : uninformed guessing, the probability of indicating that a change has occurred, on a given trial, when the participant has no information to constrain the guessing ($0 \leq U \leq 1$).

In the following, we assume that on a given trial S objects are shown on a computer display. For the test array, C objects change. In the canonical change detection task $C \in \{0,1\}$ and S varies. In the multiple change detection task $C \in \{0,1,\dots,S\}$ for a constant S . Let $D_{C,S}$ denote the probability that an object has changed and exists in working memory for a trial with set size S in which C objects have changed. In Gibson et al. (2011) $D_{C,S}$ is modeled as a function of K , C , and S :

$$\text{if } C = 0, D_{0,S} = 0 \quad (1a)$$

$$\text{if } C > 0, D_{C,S} = \begin{cases} \min\left[1, 1 - \prod_{i=S-C+1}^S \frac{i-K}{i}\right] & \text{if } C \leq S-K \\ 1 & \text{otherwise.} \end{cases} \quad (1b)$$

Note that this model is a generalization of the equations provided by Rouder et al. (2011) for modeling change detection performance in the canonical task when lapses of attention are possible. That is, when only a single change is possible, Equations (1a) and (1b) reduce to the equations provided by Rouder et al. (2011; see p. 328). The K , A , G , and U parameters are estimated from the observed responses to each experimental trial type (where trial types are defined by C and S combinations). For a trial with C changes, we write $P_{C,S}$ to denote the probability of a “change” response. These probabilities (i.e., change proportions) are modeled as functions of the four person-level parameters described above: A , G , U , and (indirectly via $D_{C,S}$) K ,

$$P_{C,S} = A[D_{C,S} + (1 - D_{C,S})G] + (1 - A)U. \quad (2)$$

We now demonstrate that in the multiple change and the canonical change detection tasks, three of the four person-level parameters (A , G , U) cannot be uniquely estimated without making additional assumptions. Recall that Figure 2 shows the hit rates and false alarm rates as a function of S for five K values in the canonical change detection task. Notably, for any K value, the values in Figure 2 could arise from sets of parameters other than the figure-generating values of $A = .85$, $G = .15$, and $U = .35$. For example, the data shown in Figure 2 could be generated by (A, G, U) values of $(.903, .199, .000)$, $(.785, .080, .546)$ or $(.722, .000, .649)$. Moreover, this pattern of data is consistent with values of A ranging from .722 to .903, and it is impossible to determine the actual value of A that generated the data (even with an infinite number of trials). This is a very broad range of possible A values, and it is unlikely that researchers who employ these tasks could draw meaningful conclusions about individual or group differences in attentional engagement given this degree of indeterminacy. To clarify the underlying cause of this underidentification, we introduce a mathematical coefficient to these models, B , that makes it possible to generate all sets of fungible (A, G, U) triples that can produce a given set of data.¹ The coefficient B is related to A , G , and U as follows:

$$\tilde{A} = A/B, \quad (3)$$

$$\tilde{G} = 1 - B(1 - G), \quad (4)$$

and

$$\tilde{U} = \frac{A - \frac{A}{B} + (1 - A)U}{1 - \frac{A}{B}}. \quad (5)$$

As is formally proved in the Appendix, with a suitable choice of B , any triple of $(\tilde{A}, \tilde{G}, \tilde{U})$ that is generated from Equations (3)–(5) will yield the same values of $P_{C,S}$ (for all values of C and S) for given values of (A, G, U) . This fact justifies our claim that the multiple change and the canonical change detection models are not identified. Further illustration of the implications of model non-identification are given in Online Appendix A. Note, however, that the choice of B is subject to the constraints that $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ must each lie between 0 and 1. As shown in Online Appendix B, these constraints will be satisfied when-ever

$$\frac{A}{A + (1 - A)U} \leq B \leq \frac{1}{1 - G}. \quad (6)$$

¹The coefficient B serves a similar function as the rotation matrix in factor analysis or principal components analysis. In factor analysis or principal components analysis, orthogonal and oblique rotations predict the same patterns of observed data and do not influence model-data fit. Likewise, permissible values of B do not influence predicted trial responses and thus also do not influence model-data fit.

In the Online Appendix B, we also derive the bounds on the $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ parameters that are implied by these bounds on B .

Identified Visual Short Term Working Memory Models

Solution 1: Equal Guessing Probabilities

One way to eliminate the indeterminacy of the change detection models is by adding constraints on the relationship between the two guessing rate parameters, G and U . Specifically, when $G = U$ there are 3 identified person-level parameters when there are at least 3 trial types. This identification constraint is justified whenever the probability of a “change” response due to inattention equals the probability of a “change” response when a changed object is not in working memory. In this updated model, the probability P of a “change” response to trial type C equals

$$P_{C,S} = A D_{C,S} (1 - G) + G. \quad (7)$$

Solution 2: An Ideal Guesser

Another way to resolve the non-identification of the canonical and multiple change detection models follows a suggestion by Rouder et al. (2011) about the normative relationship between the G and U parameters. In an ideal guesser, the informed guessing probability G will depend on set size, capacity, and the estimated base rate of changes. Specifically, Rouder et al. (2011, p. 327) propose that, for an ideal guesser,

$$G_{C,S} = \frac{(1 - D_{C,S})U}{(1 - D_{C,S})U + (1 - U)}, \quad (8)$$

and therefore

$$P_{C,S} = A [D_{C,S} + (1 - D_{C,S})G_{C,S}] + (1 - A)U. \quad (9)$$

The identification constraint given in Equation 8 is psychologically valid only if participants guess in a manner that is ideal. However, to our knowledge, there is no empirical evidence that this ideal relationship holds in change detection tasks, and it is well known that humans and other species often deviate from normative behavior across a wide range of probabilistic decision-making scenarios (Estes, 1962; Kahneman & Tversky, 2000). As a consequence, this solution is not necessarily on stronger psychological footing than Solution 1.

Simulations

To evaluate the performance of each of these solutions, we conducted two small simulation studies in which we varied the extent to which the assumptions of the solutions were violated. These simulations were designed to evaluate whether the suggested model identification restrictions lead to accurate working memory capacity estimates even if the

model identification assumptions do not hold exactly. In both studies, trial responses were simulated for the canonical change detection task with set sizes $S = \{1, 2, 3, 4, 5, 6, 7\}$. At each set size, we simulated responses to 30 change-present and 30 change-absent trials for a total of 420 trial responses. In the first simulation, we generated a trial response grid of K , A , G , and U values such that $K \in \{1, 2, 3, 4, 5\}$, $A \in \{.6, .7, .8, .9, 1\}$, $G \in \{0, .1, .2, .3, .4, .5\}$, and $U \in \{0, .1, .2, .3, .4, .5\}$. That is, in the first simulation, we assumed that informed guessing did not depend on set size and we systematically examined violations of the $G = U$ assumption of Solution 1. All levels of K , A , G , and U were fully crossed, leading to 900 unique conditions. Capacity values K were then estimated using model identification Solution 1.

In the second simulation, we generated data according to the same grid of K , A , and U values. We also assumed that informed guessing rates depended on set size and K according to Equation 8, but multiplied the resulting $G_{C,S}$ values by a factor j , $j \in \{0.6, 0.8, 1.0, 1.2, 1.4\}$, to systematically examine the effects of violating Solution 2's assumption of the relationship between G and U . For $j = 1$, the ideal observer assumption holds, and for $j \neq 1$, the informed guessing rate is systematically higher or lower than the ideal rate. All levels of K , A , U , and j were fully crossed, leading to 750 unique conditions. In both simulations, response data for 100 subjects were generated for each unique condition (i.e., set of parameters $\{K, A, G, U\}$ or $\{K, A, U, j\}$). Capacity values K were then estimated using model identification Solution 2.

For each subject in both simulations, maximum likelihood parameter estimates were obtained using the Nelder-Mead optimization routine as implemented in the *dfoptim* package (Varadhan & Borchers, 2016) in R (R Core Team, 2016)². To avoid local maxima, for each simulated response vector we obtained 10 sets of parameter estimates from 10 different sets of start values. We retained the set of parameter estimates that produced the largest log likelihood value. This design makes it possible to investigate how well the two model identification solutions perform as their model identification restrictions are or are not met. Note that we simulated a relatively large number of subjects (100 per condition) to obtain reliable estimates of bias and variance in the parameter estimates, not because this would ordinarily be required in an empirical study that was designed to estimate the parameters themselves.

Results of the first simulation are shown in Figure 3, and results for the second simulation are shown in Figure 4. Both figures show box plots of the distribution of bias ($\hat{K} - K$), and within each box, the black horizontal line represents the median bias value and the whiskers extend to 1.5 times the interquartile range of bias values. Figure 3 shows the distribution of bias for all combinations of data-generating K , A , and the absolute difference between G and U ($|G - U|$). In this figure, the bottom row of panels represents data simulated under the condition that $G = U$, and thus the boxes in the bottom row represent recovery for when the

²One issue that must be addressed when fitting any visual change detection model is that $D_{C,S}$ is a discontinuous function of K . Specifically, as K increases, the probability that a changed item is in working memory ($D_{C,S}$) "jumps" at $K = S - C$ ($C > 0$) (see Equation 1b). Because $D_{C,S}$ is a discontinuous function of K , the likelihood function for this model is also discontinuous. One implication of this result is that the derivatives of the likelihood function are not defined at any $K = S - C$ ($C > 0$), and therefore gradient-based methods of finding the maximum likelihood estimates for this model should not be used. Maximum likelihood parameter estimates can instead be computed using derivative-free methods such as the Nelder-Mead algorithm (Nelder & Mead, 1965). The *dfoptim* package was used rather than the Nelder-Mead algorithm included in the *optim* function in the base library for R because *dfoptim* includes Nelder-Mead optimization with box constraints, whereas *optim* does not at this time.

$G = U$ assumption is met. As Figure 3 shows, median $\hat{K}$ values are very close to their data-generating values across all conditions, even if the $G = U$ assumption is not met. In other words, if data are generated under the assumption that G is independent of the other parameters, median K values can be estimated without significant bias by assuming that $G = U$. Average (mean) estimates, however, are somewhat biased, and bias tends to be negative. Interestingly, K values estimated under Solution 1 tend to be negatively biased whereas K values estimated under Solution 2 tend to be positively biased. Although median $\hat{K}$ is nearly unbiased at all combinations of K , A , G , and U , there are some differences in the variability of $\hat{K}$ across conditions, as shown by the sizes of the conditional box plots. Most prominently, $\hat{K}$ variability increases substantially as the data-generating A decreases from 1.0 to 0.6. In other words, smaller attention values (A) are associated with greater variability in $\hat{K}$ than larger attention values.

Figure 4 shows the distribution of bias for all combinations of data-generating K , A , and j , where j is multiplied by the ideal informed guessing rate $G_{C,S}$ (as described above, if $j = 1$, then the ideal observer of Solution 2 holds). Consequently, the middle row of panels represent recovery when the ideal observer assumption of Solution 2 holds. Overall, the results displayed in Figure 4 are similar to those displayed in Figure 3. Namely, median K estimates are nearly unbiased in most conditions, especially at higher values of A . Moreover, even if the assumption of Solution 2 does not hold exactly, capacity estimates are largely unaffected by this misspecification. Perhaps the most prominent conclusion from both of these simulations is that when a subject does not pay attention on some proportion of trials, it becomes increasingly difficult to reliably estimate working memory capacity because many of the trial responses are the result of inattentiveness. This suggests that researchers should design experiments that maximize person attentiveness when trying to obtain the most accurate estimates of working memory capacity. In other words, although an appropriate model can avoid the problem of lapses of attention leading to an underestimate of the true working memory storage capacity, low attentiveness remains a problem in terms of reliability.

Discussion

The ability to separately assess working memory capacity and attentional engagement is both practically and theoretically important, and change detection tasks can make this possible when combined with appropriate mathematical models. However, the present findings show that the canonical and the multiple change detection tasks are not mathematically identified if these models take attentional capacity into account. Stated plainly, in this model of working memory capacity, widely different estimates of attentional engagement can lead to identical data sets. We have suggested two model identification solutions: the first assumes that the two guessing parameters, G and U , are equal to each other, and the second assumes that informed guessing is a normatively ideal function of uninformed guessing, working memory capacity, and set size. However, neither of these solutions has been supported by empirical evidence, and further research is needed to determine what model identification method is most psychologically plausible.

The typical model for the partial report variant of the change detection task is identified because all guessing in this variant is uninformed, eliminating the need to have separate G and U parameters. This variant uses a postcue to indicate which item should be reported. If the cued object is not present in memory, the participant has no information about whether or not that item might have changed, just as a participant has no information about the presence or absence of a change if attention has lapsed. Assuming that this logic is justified, the partial report change detection task is identified and should therefore be used when the goal of the research is to estimate both working memory storage capacity and attentional engagement.

Each of the three variants discussed here—the canonical variant, the multiple change variant, and the partial report variant—has advantages and disadvantages that are not captured in the typical mathematical models used to describe task performance. An advantage of the partial report variant is that only one item from the sample array must be compared with the test array, which minimizes contributions from “decision noise” (Vogel et al., 2001). However, the partial report version assumes that participants can completely ignore the uncued items. This assumption may not be true in some groups of participants. In addition, some experiments implement the partial report task by means of a test display that contains only the one item to be compared with memory, and this may lead to difficulty in knowing exactly which item from the sample array corresponds to the one item in the test display (Levillain & Flombaum, 2012). It should also be noted that location binding errors, in which the features are remembered but their locations are lost (Bays, Catalao, & Husain, 2009), will contribute to capacity estimates in the partial report task but should have little or no impact on performance in the multiple change detection task. The role of these bindings is less clear in the canonical change detection task.

An advantage of the multiple change detection version is that every trial appears to be equally easy to the participant (because the set size is held constant, and the participant has no way of knowing in advance how many items will change). This avoids the possibility that some participants will “give up” on trials with large set sizes, a scenario that would violate the assumption of the existing models that attentional engagement is independent of the set size. Such violations will lead to poor model fit and an underestimation of actual storage capacity (see Rouder et al., 2008). Moreover, participants with low storage capacity should be able to detect changes quite well when most or all of the items change between the sample and test arrays, avoiding the feeling of frequent failure that these participants experience at large set sizes in the canonical and partial report versions of the task. Consequently, the multiple change detection task may yield more accurate estimates of storage capacity in participants who are not highly motivated to maximize their performance, whereas capacity estimates with the other tasks may be distorted by individual differences in effort.

Thus, although the mathematical model of the partial report variant makes it possible to separately estimate working memory capacity and attentional engagement, it may involve assumptions about task performance (such as equivalent attentional engagement across different set sizes or an ideal observer) that may be inappropriate in some studies. As a result, if the main goal of a study is to estimate working memory capacity in a manner that is

unbiased by the degree of attentional engagement, and not to estimate attentional engagement *per se*, the canonical variant or the multiple change variant might be more appropriate than the partial report variant.

An alternative to the model identification solutions proposed in this paper is to use tasks that are not variants on change detection. For example, considerable research now uses delayed estimation tasks, in which the participant provides a continuous estimate of the remembered feature value (e.g., by clicking on the corresponding location on a color wheel; see Wilken & Ma, 2004; Zhang & Luck, 2008). These tasks make it easy to create models that capture the precision of the memory representations as well as storage capacity, and models are also available that do not assume a fixed number of discrete representations (van den Berg, et al., 2014). However, the existing models of delayed estimation do not take lapses of attention into account, and this would be a useful avenue for future model development.

Concluding Remarks

Presently, there is no perfect means of concurrently assessing working memory capacity and attentional engagement, although considerable progress has been made over the past decade. The issue of lapses of attention was largely ignored prior to the study of Rouder et al. (2008), and several different methods are now available to separately estimate working memory capacity and attentional engagement. Our work has demonstrated that the canonical change detection task and the multiple change detection task cannot be used to estimate attentional engagement. It has also demonstrated that when best practices are used, researchers can use working memory tasks to make reliable inferences about working memory capacity and attentional engagement.

Acknowledgments

This work was supported through the National Institute of Mental Health grants R01MH084861 (LF, AM, NW) and R01MH065034 (SL) and R01MH076226 (SL). Correspondence concerning this article should be addressed to Niels G. Waller, Department of Psychology, N218 Elliott Hall, University of Minnesota, 75 East River Road, Minneapolis, Minnesota, 55455. EMAIL: nwaller@umn.edu

Appendix

In this appendix we provide a formal proof that the multiple change and the canonical change detection models are not identified. Using the notation introduced in the main paper, rewrite Equation (2) as

$$P_{C,S} = AD_{C,S} + AG - AD_{C,S}G + (1 - A)U, \quad (A1)$$

$$= AD_{C,S}(1 - G) + P_{0,S}. \quad (A2)$$

Next, rewrite Equations (3), (4), and (5) as

$$A = \tilde{A}B, \quad (\text{A3})$$

$$G = \frac{B-1+\tilde{G}}{B}, \quad (\text{A4})$$

and

$$U = \frac{(1-\tilde{A})\tilde{U} - \tilde{A}B + \tilde{A}}{1-\tilde{A}B}. \quad (\text{A5})$$

Substituting Equations (A3), (A4), and (A5) into Equation (A1) with $C = 0$ yields,

$$\begin{aligned} P_{0,S} &= AG + (1-A)U \\ &= \tilde{A}B \left(\frac{B-1+\tilde{G}}{B} \right) + (1-\tilde{A}B) \frac{(1-\tilde{A})\tilde{U} - \tilde{A}B + \tilde{A}}{1-\tilde{A}B} \quad (\text{A6}) \\ &= \tilde{A}(B-1+\tilde{G}) + (1-\tilde{A})\tilde{U} - \tilde{A}B + \tilde{A} \end{aligned}$$

$$= \tilde{A}\tilde{G} + (1-\tilde{A})\tilde{U}. \quad (\text{A7})$$

Note that Equations (A6) and (A7) have the same mathematical form. This result proves our conjecture that this model is not identified for $P_{0,S}$. To show that our conjecture also holds for $P_{C,S}$ using the result from Equation (A7), let

$$\begin{aligned} P_{C,S} &= AD_{C,S}(1-G) + P_{0,S} \\ &= \tilde{A}BD_{C,S} \left(1 - \frac{B-1+\tilde{G}}{B} \right) + P_{0,S} \quad (\text{A8}) \\ &= \tilde{A}BD_{C,S} \left(\frac{1-\tilde{G}}{B} \right) + P_{0,S} \end{aligned}$$

$$= \tilde{A}D_{C,S}(1-\tilde{G}) + P_{0,S}, \quad (\text{A9})$$

As before, note that Equations (A8) and (A9) have the same mathematical form. Because $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ do not depend on C , S , or K , this result holds generally for the multiple change and canonical change detection models, regardless of the type of trials or number of replications. This completes our proof: the multiple change and the canonical change detection models are not identified.

Online Appendix A

To illustrate the implications of model non-identification for model estimation, suppose a canonical change detection task for which there are 30 change-absent and 30 change-present trials at each of seven set sizes $S = \{1, 2, 3, 4, 5, 6, 7\}$. Further imagine that for the change-absent trials, a subject provides a “change” response to 7, 5, 8, 7, 8, 6, and 8 trials at set sizes 1 through 7 (out of 30 trials for each condition), and for the change-present trials, a subject provides a “change” response to 30, 28, 24, 20, 25, 18, and 13 trials at set sizes 1 through 7 (out of 30 trials for each condition). For these data, we estimated 100 sets of maximum likelihood $\hat{K}$, $\hat{A}$, $\hat{G}$, and $\hat{U}$ using the Nelder-Mead derivative-free optimization routine (Nelder & Mead, 1965) in R (R Core Team, 2016) as implemented in the *dfoptim* package (Varadhan & Borchers, 2016). This algorithm requires the user to specify starting values for optimization. To search for local maxima we initiated the algorithm with 100 unique sets of starting values. Note that we did not estimate K using the limited information approaches that have been described in the literature (see Cowan, 2001; Pashler, 1988; Rouder et al., 2011).

For this simulation, the optimization routine produced 100 unique sets of estimates, each of which was associated with the same maximized model negative log likelihood value. Recall that in a mathematically identified model only a single set of estimates should maximize the log likelihood. Five sets of estimates from our analyses are reported in Table A1. Note in this table that although $\hat{K}$ and the negative log likelihood values are the same across replications, the values of $\hat{A}$, $\hat{G}$, and $\hat{U}$ differ. The final column in Table A1, labeled B , denotes the scaling factor with reference to the first row of parameter estimates. Because the negative log likelihood is the same for each solution, each set of parameter estimates describes (i.e., fits) the data equally well. In other words, each set of parameter estimates is associated with the same global maximum of the (log) likelihood function. Without external information, there is no way to know which set of parameter estimates most closely corresponds to the individual’s response process.

Table A1:

Five sets of maximum likelihood estimated parameters ($\hat{K}$, $\hat{A}$, $\hat{G}$, and $\hat{U}$) that yield identical model predictions. For set sizes $\{1, 2, 3, 4, 5, 6, 7\}$, the observed proportion of “change” responses equalled $\{7/30, 5/30, 8/30, 7/30, 8/30, 6/30, 8/30\}$ for the change-absent trials and $\{30/30, 28/30, 24/30, 20/30, 25/30, 18/30, 13/30\}$ for the change-present trials.

	$\hat{K}$	$\hat{A}$	$\hat{G}$	$\hat{U}$	-log likelihood	B
1	3.0000	.8187	.1715	.5227	31.8188	1.0000
2	3.0000	.6864	.0118	.7241	31.8188	1.1927
3	3.0000	.7626	.1106	.6354	31.8188	1.0735
4	3.0000	.8887	.2368	.2226	31.8188	0.9212
5	3.0000	.9133	.2574	.0014	31.8188	0.8964

The scatterplots of the resulting 100 sets of maximum likelihood parameter estimates from these simulations are shown in the upper triangular panels (i.e., the panels above the main

diagonal) of Figure A1. The diagonal panels of Figure A1 display the histograms of the estimated values and the lower triangular panels (i.e., those below the main diagonal) display the theoretical values that we computed from Equations (3)–(5). To generate these theoretical values we first computed hypothetical bounds on $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ using equations that are described in Appendix B. For this example, we found that $.68 \leq \tilde{A} \leq .91$, $0 \leq \tilde{G} \leq .26$, and $0 \leq \tilde{U} \leq .73$. Next, using the range of B implied by (6), we generated 1000 sets of exchangeable parameter estimates. An important result of these analyses is that the relationships shown in the upper triangular panels of Figure A1 are identical (but transposed) to those shown in the lower triangular panels of the figure. These results demonstrate that the multiple sets of parameter estimates in Table A1 can be derived, without error, from (3), (4), and (5) thus providing an empirical demonstration that the multiple change and the canonical change detection models are not identified.

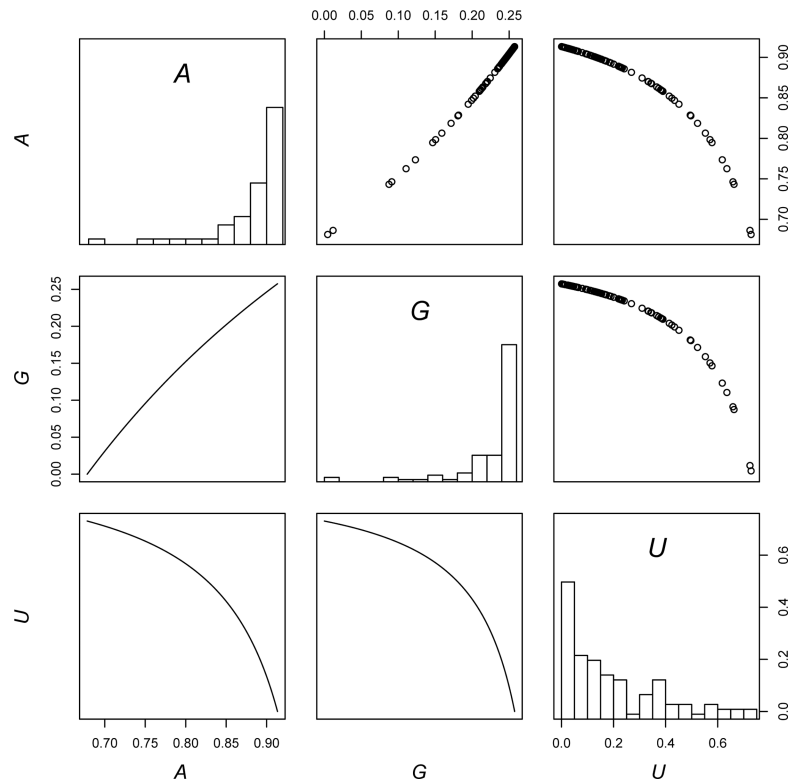


Figure A1:

Example of non-identification of the canonical change detection model. In this example, we created a set of known data from a single participant in an imaginary experiment with 30 change and 30 no-change trials at each of seven set sizes {1,2,3,4,5,6,7}. The number of change-present responses made by the participant was {30/30, 28/30, 24/30, 20/30, 25/30, 18/30, 13/30} for the change-present trials and {7/30, 5/30, 8/30, 7/30, 8/30, 6/30, 8/30} for the change-absent trials. The goal of the analysis was to demonstrate that very different estimated A, G, and U parameters could fit this single set data set. The panels along the diagonal labeled A, G, and U display histograms of the (maximum likelihood) estimated A, G, and U parameters across 100 estimation runs (with different random starting values). These histograms show that a broad range of parameter values could fit these data equally well. The upper triangular panels (i.e., panels above and to the right of the histogram panels) show scatter plots of the estimated parameters. The lower triangular panels (i.e., panels below and to the left of the histogram panels) show the theoretical relationships between the parameter estimates, as given by Equations (3)–(5). The rows and columns of the figure display estimated A, G, and U values. For instance, Panel₁₂ (row 1, column 2) displays the scatter plot between the maximum likelihood estimates of A and G. Panel₂₁ (row 2, column 1) displays 1000 pairs of exchangeable G and A estimates from Equations (3)–(5).

Online Appendix B

In this section we demonstrate that the values of B are limited by constraints on $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$. Further, we derive the range of possible values for B , $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ for given values of A , G , and U . Recall that A , G , and U must be bounded between 0 and 1. For a value of B to be

admissible, it must imply $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ values (see Equations (3)–(5)) that are also bounded between 0 and 1. We now derive the bounds on B from the bounds on $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$. Starting with $\tilde{A}$,

$$\begin{aligned} 0 &\leq \tilde{A} \leq 1 \\ 0 &\leq A/B \leq 1 \\ 0 &\leq A \leq B. \end{aligned}$$

Next, adding the constraint on $\tilde{G}$,

$$\begin{aligned} 0 &\leq \tilde{G} \leq 1 \\ 0 &\leq 1 - B(1 - G) \leq 1 \\ -1 &\leq -B(1 - G) \leq 0 \\ 0 &\leq B(1 - G) \leq 1 \\ 0 &\leq B \leq \frac{1}{1 - G}. \end{aligned}$$

Finally, accommodating the constraint on $\tilde{U}$,

$$\begin{aligned} 0 &\leq \tilde{U} \leq 1 \\ 0 &\leq \frac{AB - A + (1 - A)UB}{B - A} \leq 1 \\ 0 &\leq AB - A + (1 - A)UB \leq B - A \\ A &\leq B(A + (1 - A)U) \leq B \\ A/B &\leq A + (1 - A)U \leq 1 \\ \frac{A}{A + (1 - A)U} &\leq B \leq \frac{B}{A + (1 - A)U}. \end{aligned}$$

Note that because $(A + (1 - A)U) \leq 1$, the last upper bound for B is trivial. Combining the above results, and recognizing that a finite upper bound for B occurs only when $B \leq \frac{1}{1 - G}$ and $G < 1$, it can be shown that

$$A \leq B,$$

and

$$\frac{A}{A + (1 - A)U} \leq B.$$

Moreover, because $A + (1 - A)U \leq 1$, the range of admissible B values is

$$\frac{A}{A + (1 - A)U} \leq B \leq \frac{1}{1 - G}. \quad (\text{A10})$$

Given these bounds on B , we can also compute bounds for $\tilde{A}$, $\tilde{G}$, and $\tilde{U}$ given A , G , and U . Specifically, because $B = A/\tilde{A}$, for $\tilde{A}$,

$$\begin{aligned}
\frac{A}{A + (1 - A)U} &\leq A/\tilde{A} \leq \frac{1}{1 - G} \\
1 - G &\leq \tilde{A}/A \leq \frac{A + (1 - A)U}{A} \quad (\text{A11}) \\
A(1 - G) &\leq \tilde{A} \leq A + (1 - A)U.
\end{aligned}$$

Next, recognizing that Equation (4) implies $B = \frac{1 - \tilde{G}}{1 - G}$, for $\tilde{G}$,

$$\begin{aligned}
\frac{A}{A + (1 - A)U} &\leq \frac{1 - \tilde{G}}{1 - G} \leq \frac{1}{1 - G} \\
\frac{A(1 - G)}{A + (1 - A)U} &\leq 1 - \tilde{G} \leq 1 \\
\frac{A(1 - G)}{A + (1 - A)U} - \frac{A + (1 - A)U}{A + (1 - A)U} &\leq -\tilde{G} \leq 0 \quad (\text{A12}) \\
\frac{-AG - (1 - A)U}{A + (1 - A)U} &\leq -\tilde{G} \leq 0 \\
0 &\leq \tilde{G} \leq \frac{AG + (1 - A)U}{A + (1 - A)U}.
\end{aligned}$$

In terms of $\tilde{U}$, after substituting Equation (3) into Equation (5) and rearranging terms, we find that

$$B = \frac{A - \tilde{U}A}{(1 - A)U - \tilde{U} + A}. \quad (\text{A13})$$

Note that because $B \geq 0$, $0 \leq A \leq 1$, and $0 \leq \tilde{U} \leq 1$, both the numerator and denominator of Equation (A13) must be positive. Thus,

$$\begin{aligned}
\frac{A}{A + (1 - A)U} &\leq \frac{A - \tilde{U}A}{(1 - A)U - \tilde{U} + A} \leq \frac{1}{1 - G} \\
\frac{A[(1 - A)U - \tilde{U} + A]}{A + (1 - A)U} &\leq A - \tilde{U}A \leq \frac{(1 - A)U - \tilde{U} + A}{1 - G} \quad (\text{A14}) \\
\frac{A[(1 - A)U - \tilde{U} + A]}{A + (1 - A)U} + \tilde{U}A &\leq A \leq \frac{(1 - A)U - \tilde{U} + A}{1 - G} + \frac{\tilde{U}A(1 - G)}{1 - G}.
\end{aligned}$$

Isolating the right-hand side of Equation (A14):

$$\begin{aligned}
\frac{A[(1-A)U - \tilde{U} + A]}{A + (1-A)U} + \tilde{U}A &\leq A \\
\frac{(1-A)U - \tilde{U} + A}{A + (1-A)U} + \tilde{U} &\leq 1 \\
(1-A)U - \tilde{U} + A + \tilde{U}(A + (1-A)U) &\leq A + (1-A)U \\
\tilde{U}(A + (1-A)U - 1) &\leq A + (1-A)U - (1-A)U - A \\
\tilde{U}(A + (1-A)U - 1) &\leq 0 \\
0 &\leq \tilde{U}.
\end{aligned} \tag{A15}$$

Note that the final inequality in Equation (A15) occurs because $A + (1-A)U \leq 1$. Finally, from the above it can be proved that,

$$\begin{aligned}
A &\leq \frac{(1-A)U - \tilde{U} + A}{1-G} + \frac{\tilde{U}A(1-G)}{1-G} \\
A(1-G) &\leq (1-A)U - \tilde{U} + A + \tilde{U}A(1-G) \\
A(1-G) - (1-A)U - A &\leq \tilde{U}[A(1-G) - 1] \\
\tilde{U} &\leq \frac{A(U-G) - U}{A(1-G) - 1}.
\end{aligned}$$

Thus,

$$0 \leq \tilde{U} \leq \frac{A(U-G) - U}{A(1-G) - 1}$$

gives the bounds on $\tilde{U}$ for given A , G , and U values.

References

- Adam KC, Mance I, Fukuda K, & Vogel EK (2015). The contribution of attentional lapses to individual differences in visual working memory capacity. *Journal of Cognitive Neuroscience*, 27, 1601–1616. [PubMed: 25811710]
- Alvarez GA, & Cavanagh P (2004). The capacity of visual short-term memory is set both by information load and by number of objects. *Psychological Science*, 15, 106–111. [PubMed: 14738517]
- Anderson JR, Matessa M, & Lebiere C (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction*, 12, 439–462.
- Baddeley A (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829–839. [PubMed: 14523382]
- Bays PM (2014). Noise in neural populations accounts for errors in working memory. *Journal of Neuroscience*, 34, 3632–3645. [PubMed: 24599462]
- Bays PM, Catalao RF, & Husain M (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9, 1–11.
- Cowan N (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185. [PubMed: 11515286]
- Cowan N (2012). Working memory capacity: Psychology press.
- Cowan N, Naveh-Benjamin M, Kilb A, & Saults JS (2006). Life-span development of visual working memory: when is feature binding difficult? *Developmental Psychology*, 42, 1089. [PubMed: 17087544]

- Cusack R, Lehmann M, Veldsman M, & Mitchell DJ (2009). Encoding strategy and not visual working memory capacity correlates with intelligence. *Psychonomic Bulletin & Review*, 16, 641–647. [PubMed: 19648446]
- Engle RW, Tuholski SW, Laughlin JE, & Conway AR (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309–331. [PubMed: 10513398]
- Erickson MA, Albrecht MA, Robinson BM, Luck SJ, & Gold JM (2017). Impaired suppression of delay-period alpha and beta is associated with impaired working memory in schizophrenia. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 2, 272–279.
- Estes WK (1962). Learning theory. *Annual Review of Psychology*, 13, 107–144.
- Fukuda K, Vogel E, Mayr U, & Awh E (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, 17, 673–679. [PubMed: 21037165]
- Gibson B, Wasserman E, & Luck SJ (2011). Qualitative similarities in the visual short-term memory of pigeons and people. *Psychonomic Bulletin & Review*, 18, 979–984. [PubMed: 21748417]
- Johnson MK, McMahon RP, Robinson BM, Harvey AN, Hahn B, Leonard CJ, Luck SJ, & Gold JM (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology*, 27, 220–229. [PubMed: 23527650]
- Kahneman D, & Tversky A (2000). *Choices, values, and frames*. Cambridge University Press.
- Levillain F, & Flombaum JI (2012). Correspondence problems cause repositioning costs in visual working memory. *Visual Cognition*, 20, 669–695.
- Luck SJ, & Vogel EK (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281. [PubMed: 9384378]
- Luck SJ, & Vogel EK (2013). Visual Working Memory Capacity: From Psychophysics and Neurobiology to Individual Differences. *Trends in Cognitive Sciences*, 17, 391–400. [PubMed: 23850263]
- Luria R, & Vogel EK (2011). Visual search demands dictate reliance on working memory storage. *Journal of Neuroscience*, 31, 6199–6207. [PubMed: 21508243]
- Meyer DE, & Kieras DE (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, 104, 3–65. [PubMed: 9009880]
- Mrazek MD, Smallwood J, Franklin MS, Chin JM, Baird B, & Schooler JW (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology: General*, 141, 788–798. [PubMed: 22468669]
- Pashler H (1988). Familiarity and change detection. *Perception & Psychophysics*, 44, 369–378. [PubMed: 3226885]
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.Rproject.org/>
- Rouder JN, Morey RD, Cowan N, Zwilling CE, Morey CC, & Pratte MS (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, 105, 5975–5979.
- Rouder JN, Morey RD, Morey CC, & Cowan N (2011). How to measure working memory capacity in the change detection paradigm. Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm *Psychonomic Bulletin & Review*, 18, 324–330. [PubMed: 21331668]
- Sperling G (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74 (11, Whole No. 498).
- van den Berg R, Awh E, & Ma WJ (2014). Factorial comparison of working memory models. *Psychological Review*, 121, 124. [PubMed: 24490791]
- Varadhan R, & Borchers HW (2016). *dfoptim: Derivative-free optimization*. R package version 2016.7–1 <https://CRAN.R-project.org/package=dfoptim>

- Vogel EK, Woodman GF, & Luck SJ (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 92–114. [PubMed: 11248943]
- Wilken P, & Ma WJ (2004). A detection theory account of change detection. *Journal of Vision*, 4, 1120–1135. [PubMed: 15669916]
- Zhang W, & Luck SJ (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233–235. [PubMed: 18385672]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

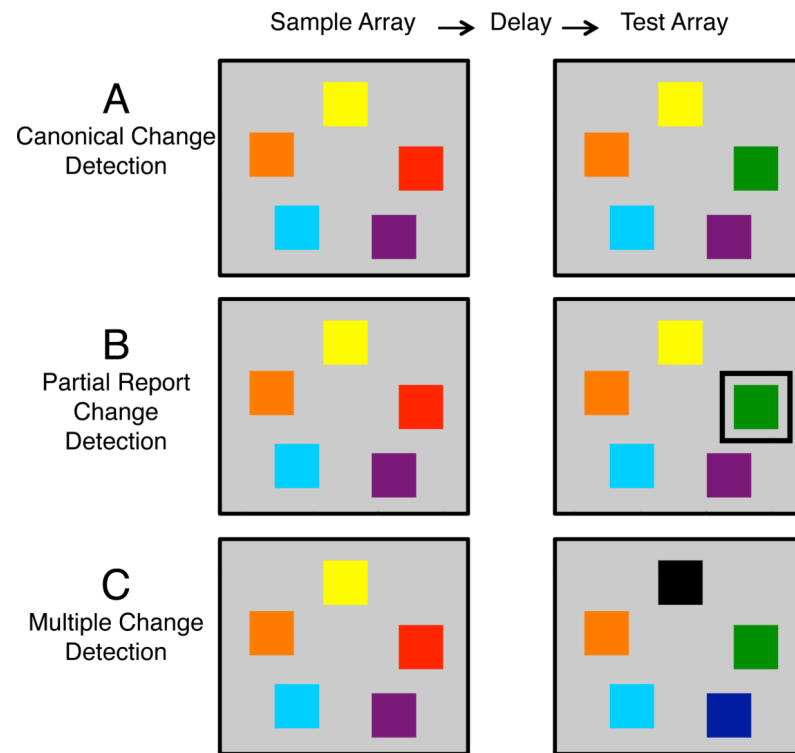


Figure 1:
Illustration of three visual change detection task variants.

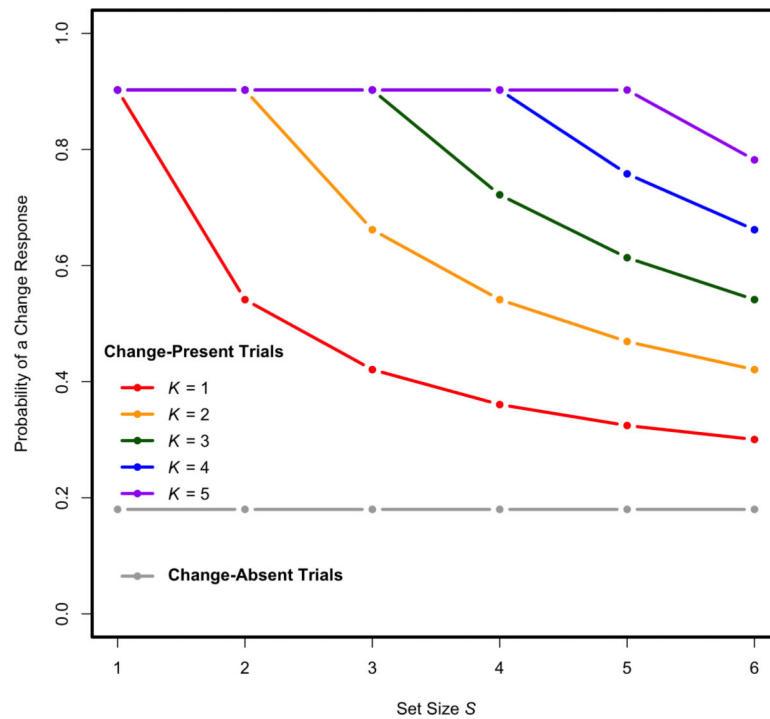
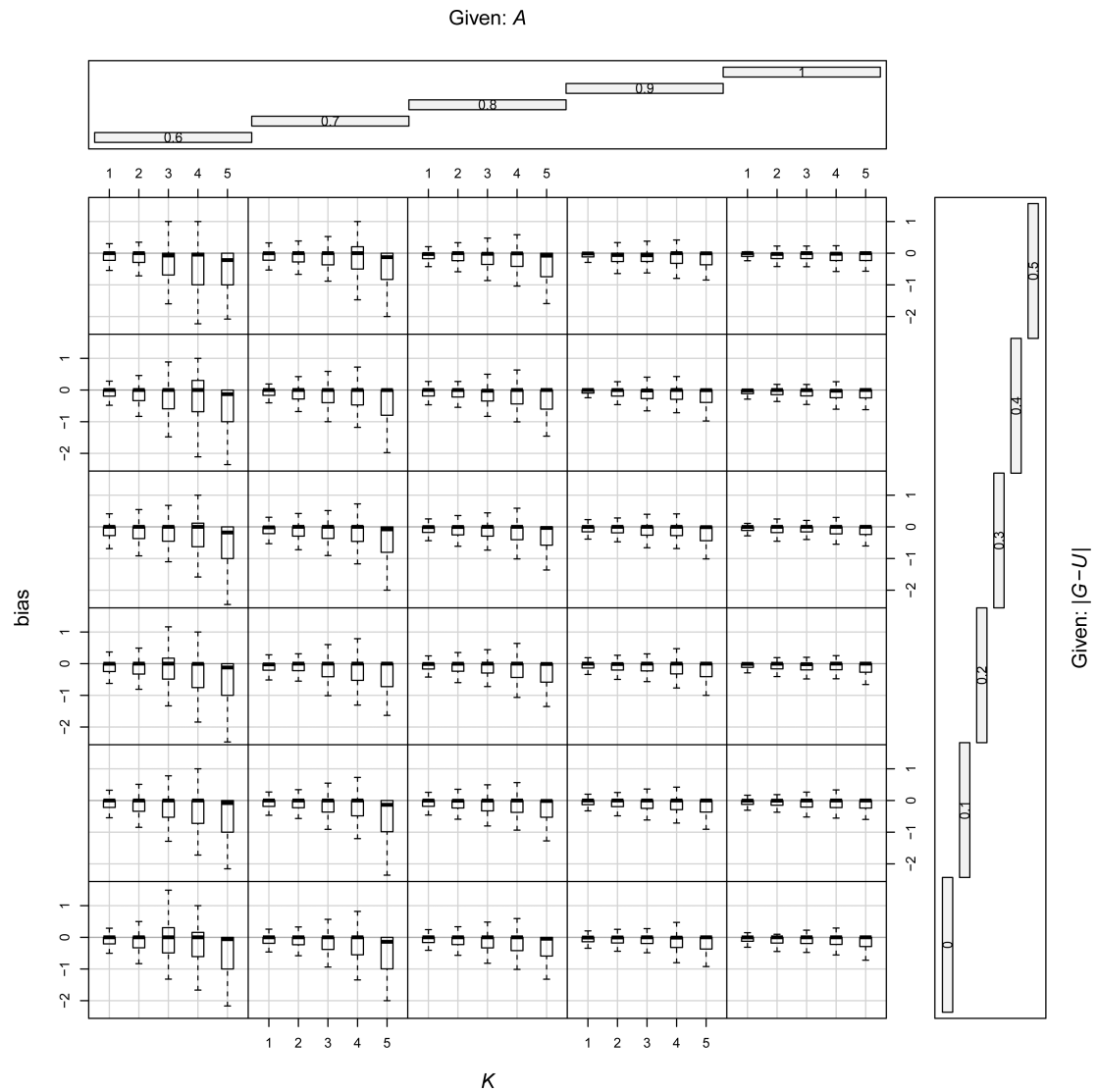
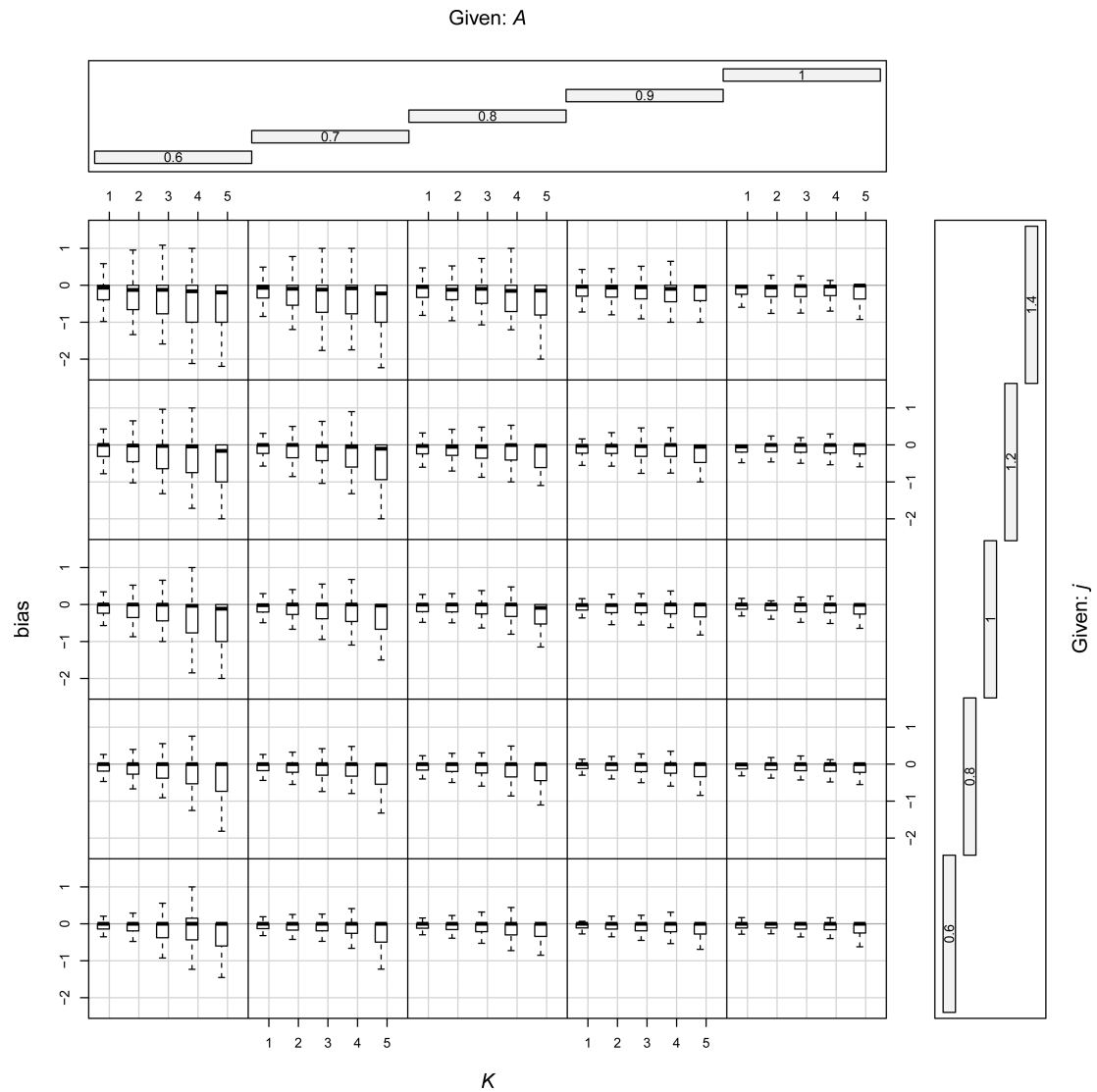


Figure 2:

Probability of change detection in the canonical change detection model. In this example, $S \in \{1,2,3,4,5,6\}$, $K \in \{1,2,3,4,5\}$, $A = .85$, $G = .15$, and $U = .35$. Importantly, the same picture arises from different (A, G, U) triples. For example, (A, G, U) triples of $(.903, .199, .000)$, $(.785, .080, .546)$, and $(.722, .000, .649)$ make identical predictions as $(.85, .15, .35)$.

**Figure 3:**

Conditioning plot of $\hat{K}$ bias ($\hat{K} - K$) for various combinations of data-generating K , A , and the absolute difference between G and U (i.e., $|G - U|$) values for the canonical change detection task. Each cell represents five levels of K . Data were generated according to the unidentified canonical change detection model. Parameters were estimated under the Solution 1 assumption that $G = U$, and the bottom row of panels represent simulations for which the Solution 1 assumption holds. The furthest right column of cells displays results for perfect attentional engagement ($A = 1$), and columns of cells further to the left correspond to lower attentional engagement.

**Figure 4:**

Conditioning plot of $\hat{K}$ bias ($\hat{K} - K$) for various combinations of data-generating K , A , and j values for the canonical change detection task, where j is the factor by which the ideal observer guessing rates $G_{C,S}$ are multiplied (if $j = 1$), then the Solution 2 assumption holds). Each cell represents five levels of K . Parameters were estimated under the ideal relationship between set size, capacity, and informed guessing defined by Solution 2. The furthest right column of cells displays results for perfect attentional engagement ($A = 1$), and columns of cells further to the left correspond to lower attentional engagement.