Change detection in visual short-term memory: The relative impact of pairwise switches and identity substitutions

Raju P. Sapkota[1*], Shahina Pardhan[1] & Ian van der Linde[1,2]

[1]Vision & Eye Research Unit (VERU), Postgraduate Medical Institute,

Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK.

[2]Department of Computing & Technology, Anglia Ruskin University, East Road,

Cambridge CB1 1PT, UK.

Author Note

[*]Raju P. Sapkota, Vision & Eye Research Unit (VERU), Anglia Ruskin University.

Correspondence concerning this article should be addressed to Raju P. Sapkota, Vision & Eye Research Unit (VERU), Postgraduate Medical Institute, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK. E-mail: raju.sapkota@anglia.ac.uk

Phone No.: +44 845 1962642

Abstract

Numerous kinds of visual events challenge our ability to keep track of objects that populate our visual environment from moment to moment. These include blinks, occlusion, shifting visual attention and changes to object's visual and spatial properties over time. These visual events may lead to objects falling out of our visual awareness but can also lead to unnoticed changes, such as undetected object replacements and positional exchanges. Current visual memory models do not predict which visual changes are likely to be the most difficult to detect. We examined the accuracy with which *switches* (where two objects exchange locations) and *substitutions* (where one or two objects are replaced) are detected. Inferior performance for one-object substitutions vs. two-objects switches, along with superior performance for two-object substitutions vs. two-object switches was found. Our results are interpreted in terms of object file theory, trade-offs between diffused and localized attention, and net visual modification.

Keywords: visual short-term memory, switch, substitution, change detection.

Change detection in visual short-term memory: The relative impact of pairwise switches and identity substitutions

Several seminal studies have used the change detection paradigm to study human visual short-term memory (VSTM) low-capacity transient memory system proposed to hold visual information online to support on-going cognitive tasks (Luck & Vogel, 1997; Phillips, 1974; Pashler, 1988; Wheeler & Treisman, 2002; Alvarez & Cavanagh, 2004). In the change detection paradigm, typically, two multi-item visual displays are presented one after another, separated by a brief interlude (of the order of seconds). Observers are required to indicate if the second display has changed visually from the first (Luck & Vogel, 1997; Phillips, 1974; Wilken & Ma, 2004). Changes often entail the exchange of two object locations, referred to as an object *swap* (Hollingworth, 2006) or *switch* (Simons, 1996), or the replacement of one or more objects by new objects, referred to as an object *substitution* (Luck & Vogel, 1997; Pashler, 1988) [1].

VSTM creates a temporal bridge that serves to link perceptual experiences across interruptions, such as the display intervals used in change detection studies. It follows then that limitations in VSTM, stemming from its low capacity and proclivity for decay and inter-stimulus confusion, can produce perceptual experiences that appear stable and unchanging despite genuine modifications, producing *change blindness,* a phenomenon wherein ostensibly obvious visual changes go unnoticed across a transient period, such as a brief interlude (Rensink, 1996, 2000). Change detection performance will be influenced by the quality of the memory representations created for items presented in the first display (Wilken & Ma, 2004), *viz*., encoding fidelity, along with

[1]Change detection experiments can also entail the addition of new objects, object deletion, or the updating of one or more objects in some visuo-spatial dimension (such as size, location, or color).

memory maintenance and subsequent retrieval costs, which collectively facilitate a comparison with items presented in the second display (Marois & Ivanoff, 2005).

Object file theory (OFT) proposes that information is stored in VSTM in the form of spatially indexed object files (Kahneman, Treisman, & Gibbs, 1992). Object files are hypothetical memory structures, within which successive states of objects are said to be linked and integrated, allowing us to establish perceptual continuity between objects viewed across multiple fixations, blinks or saccades (Irwin 1991; Irwin & Andrews, 1996; Kahneman et al., 1992). An extension of OFT called ‘trans-saccadic object file theory’ incorporates eye movements, proposing that 3-4 object files may be accurately stored in VSTM across changes in gaze position (Irwin 1991; Irwin & Andrews, 1996; Irwin & Zelinsky, 2002).

Our proficiency at detecting changes to scenes produced by object switches and substitutions may differ, since each requires its own set of (putative) OFT operations. However, neither the VSTM nor OFT literature comprehensively addresses this question, as experiments that compare change detection performance for switches and substitutions are scarce and indirect. Simons (1996) addresses this question to some degree by comparing the change detection performance for a two-object location switch vs. a single-item identity substitution (wherein a new object replaces a previously seen object). In experiments using displays comprising 5 novel geometric shapes, object switch detection was found to be superior to object substitution detection, but only for short inter stimulus intervals, ISIs (250 ms, rather than 4.3 s). However, to detect an object switch, observers must update the location information of two extant object files; to detect an object substitution, a new object file must be created in VSTM and potentially, an old object file overwritten/deleted. In Simon’s (1996) study, the number of spatial locations modified as a result of the experimental manipulations was not equated (i.e., in the object switch condition, two of the five items presented underwent a change, whereas in the object substitution condition, only a single item underwent a change). It is therefore difficult to know whether observers were simply noticing the net visual modification between the two display intervals; i.e., that the switch vs. substitution experimental parameter may have been secondary and incidental to spatially diffuse low-level change detection that encompassed the entire display but was not reliant upon memory for *individual* objects.

Although one might predict that a two-object location switch will produce a greater change detection performance relative to a one-object substitution due to the greater net visual modification produced by updating two spatial locations rather than one, an alternative hypothesis might predict that the persistence of the same object identities between displays in the switch condition itself will be a weaker indicator of a change than the introduction of a novel item, which may alter the summary statistics (or overall scene *gist*) and produce a corresponding ‘pop-out’ effect (Jonides & Yantis, 1988). Indeed, a greater performance for one-item substitutions over two-item switches was shown in experiments 1, 2 and 5 of Simons (1996) Fig. 2. Furthermore, when viewing objects in motion, positional exchanges are entirely feasible visual events that nevertheless might not be as attention-grabbing as the introduction of a new object, despite entailing the modification of a greater proportion of the visual field.

We present results from two experiments that measure performance for detecting two-object location switch *vs*. one or two-object substitution changes. Our results have important implications in understanding change blindness by quantifying the relative cognitive burden of the putative *create* and *update* operations of OFT (i.e., in terms of their impact on VSTM performance). Our results also address the question of whether the higher detectability of switch *vs*. substitution changes reported in Simons (1996) can be explained purely by the number of objects modified between study and test displays (net visual modification).

Experiment 1

This experiment compares VSTM change detection performance for a two-object *switch* *vs*. a one-object *substitution*. Although a partial replication of Simons (1996), experiment 1 is valuable and necessary as a basis of comparison since Simons (1996) study reports variable results (higher but non-significant performance for one-object *substitution* vs. two-object *switch* in experiments 1-2 and 5, and higher but non-significant performance for two-object *switch* vs.one-object *substitution* in experiment 3. Only experiment 4 of Simons (1996) reports significantly higher performance for two-object *switch* vs. one-object *substitution*, but using non-verbal stimuli displayed very briefly. Performance variability between experiments for detecting a two-object *switch* and a one-object *substitution* in Simons' (1996) study, is possibly the consequence of the variability in the experimental procedures used (stimulus class, display intervals, use or not of a verbal suppression task).

Method

Participants

Fifteen normal/corrected-to-normal subjects with Mini-Mental State Examination (MMSE) scores ≥ 27 (i.e., without impaired memory function) participated in the study (M_{age} = 26.5 years, SD = 5.30) of which 9 were male. All participants were naïve to the purpose of the study, and were paid for their time. Participants were treated in accordance with appropriate ethical guidelines; the tenets of the Helsinki Declaration were observed.

Stimuli

Stimuli comprised a set of 170 Snodgrass line drawings of real-world objects (Snodgrass & Vanderwart, 1980), each centered within an invisible square subtending 2° of visual angle, at a testing distance of 57 cm. Object line drawings from each of 14 different conceptual categories were used (e.g., four-footed animals, articles of furniture, parts of the human body). Example stimuli are shown in Fig. 1. Stimulus presentation was controlled by MATLAB (Mathworks, Natick, MA) with the PsychToolbox/VideoToolbox extensions (Brainard 1997; Pelli, 1997). Stimulus background was set to white.

<< Figure 1 About Here >>

Apparatus

Stimuli were displayed on a LCD monitor set at a spatial resolution of 1024 × 768 pixels and a refresh rate of 75 Hz. The monitor was positioned at 57 cm from participants (such that the spatial extent of the display was approx. 34° × 27°). A chin/forehead rest was used to stabilize viewing position and distance. Ambient light was held constant across trials and between participants.

Procedure

Experimental procedures are shown diagrammatically in Fig. 2. Each trial began with the presentation of a two-digit number (for 400 ms) within an invisible square (2°) at the display screen center. This was followed by the presentation of a study display, in which four to-be-remembered stimuli were shown simultaneously (for 400 ms) at the four vertices of an imaginary square (8° × 8°) positioned about the display screen center. Observers spoke aloud the two-digit number simultaneously during the examination of study stimuli in order to discourage verbal encoding of visual stimuli (i.e., a *verbal suppression task*, Baddeley, 1986; Todd & Marois, 2004), and to minimize in-trial transfer of episodic representations into more general semantic representations. This was followed by a blank 1000 ms ISI and the presentation of a test display, in which a second set of four stimuli were shown (using the same screen coordinates as the study display).

<< *Figure 2 About Here* >>

In 50% of trials, study and test displays matched exactly, i.e., study stimuli were re-displayed at their original locations (no change). In the remaining trials, the test display changed from the study display in one of the following two ways: 1. a pair of study objects (selected at random) exchanged locations (object switch); 2. one new object substituted for one randomly selected study object (object substitution). The number of switch and substitution (change) trials within each block was equal. Change and no-change trials were randomly interleaved (but counterbalanced across trial type). The pairing of objects to locations was randomized across trials such that no two study displays were exactly the same. Participants were not required to explicitly report the type of change that had been applied, merely whether a change *had* or *had not* occurred. Participants responded ‘change’ or ‘no change’ for each trial by pressing specific keys on a response keypad. 'No change' responses submitted for trials in which either a switch or substitution change had been applied were taken to be indicative of change blindness. Per-trial auditory feedback (correct/incorrect) was provided. Furthermore, participants were briefed prior to data capture that change and no change trials would occur equally often, with an aim to discourage strategic response biases. The next trial started immediately after a response was submitted.

Four study stimuli were used to envelop the commonly cited 3-4 item capacity of VSTM (Irwin & Zelinsky, 2002; Luck & Vogel, 1997; Pashler, 1988). Stimuli were chosen from different semantic categories to avoid processing competition that may have arisen if stimuli from the same category had been used (owing to a greater number of shared properties, Bright, Moss, Stamatakis, & Tyler, 2005). The 8° × 8° display area ensured that study stimuli were sufficiently close to be encompassed by the macular area on central viewing, but sufficiently far apart to minimize spatial crowding effects (Flom, Weymouth & Kahneman, 1963; Polat & Sagi, 1993). The importance of response accuracy (rather than speed) was emphasized to participants.

Each participant completed a practice block of 20 trials (using stimuli not featuring in the main experiment), followed by the main experiment, in which five blocks of 80 trials (i.e., 400 trials in total, 200 no change, 100 switch, 100 substitution) were run. Approximately 30 minutes of data capture per participant were required. Rest breaks were permitted between blocks; data were captured within a single session for all participants.

The number of correct/incorrect responses were used to calculate performance (measured as % correct detection) for no change, switch, and substitution trials, and analyzed using one-way repeated measures ANOVA. A corresponding analysis of reaction time (RT) data was made. Overall sensitivity (*d*') and response bias (c) were calculated using formulae given in Stanislaw & Todorov (1999). Bonferroni correction was used to control type I error rate. Where the assumption of sphericity was violated (identified using Mauchly's test), degrees of freedom were adjusted using the

Greenhouse-Geisser procedure. The number of items stored in memory (*K*) was calculated using formulae given in Cowan (2001). A one-sample *t*-test was used to compare performance against chance level.

Results

Fig. 3 shows how mean % correct detection (pooled across participants) varied by trial type. Mean % correct detection for no change, switch, and substitution trials were found to be 88.89 (SD = 5.60), 75.25 (SD = 8.75), and 53.32 (SD = 12.07), respectively. A one-sample *t*-test showed performance in one-object substitution trials to average just above chance level (although significantly so $t(14) = 17.10$, $p < 0.01$), similarly to Simons (1996). These data led to an overall sensitivity (*d*') of 1.69 (SD = 0.45). A response bias (c) of 0.46 (SD = 0.27) was found, suggesting that participants were slightly inclined towards responding ‘no change’. A one-way repeated measures ANOVA with trial type (3 levels) as a within-subjects factor showed a significant main effect of trial type on % correct detection, $F(1.22, 17.15) = 68.69$, $p < 0.01$, $\eta^2 = 0.83$. A pairwise comparison (Bonferroni-corrected) between individual trial types revealed significantly greater detection for no change trials compared to either switch (mean difference = 13.64%, $p < 0.01$) or substitution (mean difference = 35.57%, $p < 0.01$) trials, highlighting greater VSTM performance where objects and positions are repeated from study to test display without modification. However, more germane to the purpose of the present study, is that significantly greater performance for switch *vs*. substitution trials was found on pairwise comparison (mean difference = 21.93%, $p < 0.01$), demonstrating superior change detection performance where two study object locations are exchanged (switched), relative to where one new object replaces an old object. This finding was mirrored in the corresponding RT data, wherein significantly shorter reaction times ($p < 0.01$, paired *t*-test) for detecting object switch (340 ms, SD = 40) *vs*. substitution (400 ms, SD = 80) were found, reflecting a less effortful decision.

<< Figure 3 About Here >>

One could argue that the results of this experiment, and others like it (such as Simons, 1996; although he only found an equivalent result in his experiment 4, and the opposite, and/or non-significant results in experiment 1-3 and 5), superior performance in the object switch condition may be attributable to the fact that there are two visuo-spatial events that could highlight a change (i.e., two objects exchanged locations in the test display, relative to the study display), whereas in the object substitution condition only a single visuo-spatial event occurs (i.e., one study object is substituted for a new object in the test display). In other words, greater net visual modification between the study and the test display occurs in the two-object switch condition (requiring less effort to detect such change) relative to the one-object substitution condition, which may be responsible for its superior performance independently of any underlying memory operations that relate to single objects. To test this possibility, a second experiment was conducted.

Experiment 2

Like experiment 1, experiment 2 examines change detection performance for object switch and substitution, but it equates the number of spatial locations undergoing a change. In object switch trials, a pair of objects exchange locations from study to test display (i.e., identical to experiment 1); in object substitution trials two new objects replace two original objects from the study display. This equates net modification to the visual display between conditions, and enables us to test the hypothesis that Simons (1996) finding that identity updates (substitutions) are more difficult to detect than positional updates (switches) is due to net visual modification. Furthermore, experiment 2 also enables us to establish whether effect upon scene gist of a two-object substitution will yield greater detection performance than a two-object switch (i.e., whether participants employ object identity in a significant manner in change detection tasks of this kind, or whether the net area updated, visual modification, is indeed the primary factor, as experiment 1 suggests).

Method

Methods are similar to experiment 1, except for the following; fifteen new participants with MMSE scores ≥ 27 were recruited (M_{age} = 26.5 years, SD = 5.97) of which 9 were male. In change trials in which object substitution occurred, a pair of new objects replaced two original objects from the study display, selected at random.

Results

Fig. 4 shows how mean % correct detection (pooled across participants) varied by trial type. Mean % correct detection rates for no change, switch, and substitution trials were found to be 91.72 (SD = 4.81), 78.38 (SD = 8.04), and 83.07 (SD = 6.37), respectively. These data led to an overall sensitivity (*d'*) of 2.33 (SD = 0.46). A response bias (c) of 0.30 (SD = 0.24) was found, suggesting that, like in experiment 1, participants were inclined towards responding ‘no change’. A one-way repeated measures ANOVA with trial type (3 levels) as a within-subjects factor shows a significant main effect of trial type on % correct detection, $F(2, 28) = 21.60$, $p < 0.01$, $\eta^2 = 0.61$. A pairwise comparison between trial types shows significantly greater performance for no change relative to either switch (mean difference = 13.33, $p < 0.01$) or substitution (mean difference = 8.65, $p < 0.01$) trials, confirming the observation made in experiment 1 that performance is greater where objects and positions are repeated from study to test display. However, in contrast to experiment 1, significantly lower performance for object switch relative to object substitution trials on pairwise comparison was found (mean difference = -4.68, $p = 0.02$). This yielded a lower number of items stored (K) in VSTM in experiment 1 (mean = 2.13, SD = 0.43) than in experiment 2 (mean = 2.86, SD = 0.35) , $z = -5.10$, $p < 0.01$, demonstrating superior change detection performance where two new objects replace two old objects, relative to where two old objects merely exchange locations. Shorter mean RT was found for object substitution (390 ms, SD = 40) *vs*. switch (410 ms, SD = 40) change detection, although the RT difference was not statistically significant, possibly because, unlike experiment 1, the number of spatial locations undergoing change *was* equated between conditions, making the relative difficulty of these two conditions more similar.

<< Figure 4 About Here >>

General Discussion

The novelty of this study stems from the fact that change detection performance for object switch and substitution trials has been compared directly by equating the number of spatial positions undergoing a change. This manipulation is relevant to understanding the mechanisms that underlie change blindness since it compares the relative impact of different kinds of change in terms of visual and spatial continuity. Our results are also interpretable in terms of OFT, enabling a comparison of the relative impact of creating *vs*. updating object files on VSTM maintenance and retrieval.

Three main observations emerged from our data (although more germane to the purpose of the study are the second and the third observations): 1. participants were better at detecting ‘no change’ vs. ‘change’ trials, regardless of the change type, and whether one or two objects were involved in the change; 2. participants were better at detecting a change produced by switching a pair of objects than by substituting one object for a new object; 3. In contrast to the second observation, participants were better at detecting a change produced by substituting a pair of objects than by switching a pair. The performance disparity in detecting the two different change types was also reflected in the RT data, although this did not reach statistical significance for experiment 2.

First, we will explain our findings in the context of object file theory (OFT). With observation 1, since in ‘no change’ trials the same set of objects were seen in the same locations in both study and test displays, only the refreshing of extant object files was required during memory retrieval. This is assumed to be minimally disruptive to object memory relative to switch (wherein two object files must be updated) and substitution (wherein a new object file must be *created*, and potentially another *deleted* or *overwritten*) conditions. The finding of significantly shorter RT for ‘no change’ trials compared to both types of change trial supports the interpretation that superior performance in this condition is due to the ‘same position advantage’ rather than merely a consequence of positive response bias (e.g., Hollingworth, 2006, 2007; Sapkota, Pardhan, & van der Linde, 2011).

Our second observation goes beyond basic OFT, and suggests that *updating* object file information for *two* objects produces a change that is easier to detect than the *creation* of *one* new object file (and potentially the deletion of another), despite introducing no novel visual information. A comparison between the object substitution conditions of experiment 1 and 2 reveals that the performance for change detection where *two* objects are substituted (experiment 2) relative to where *one* object is substituted (experiment 1) was as high as 57% (z = 8.44, p < 0.01), which may be due to an increased 'pop out' effects (see below).

Our third observation shows that the change detection advantage observed where two object files are spatially *updated* (i.e., object switch), relative to where new object files are *created* (i.e., object substitution), is limited *only* to the situation in which *one* new object file replaces *one* original object file (Experiment 1). As the number of new object files created during the test display increases beyond one, a reversal in change detection performance occurs; VSTM supports superior change detection performance in the object substitution condition relative to the pairwise object switch condition. This suggests that a change in which two object files are updated is more difficult to detect than a change in which two new object files are created.

One plausible explanation for this is that the abrupt change in the summary statistics between the visual displays (global gist) where two new objects are added may have caused the second display to sufficiently 'pop out' from the first display (Jonides & Yantis, 1988), despite the net screen area undergoing change being equal in both conditions (see below). Alternatively, it could be that the display disruption accompanying a two-object switch attracts more attention than a one-object substitution, facilitating superior change detection performance, but when the number of locations undergoing a change is equated, greater attention is attracted where the objects themselves alter identity rather than merely relocate.

One might propose that, of the two items that exchanged locations between display intervals in switch trials, only one item may have been encoded in memory, and thus any detected change could in practice have been perceived as a single new item appearing at a monitored location (an object substitution), rather than as two items exchanging locations (a pairwise switch). If this had been the case, a correct response could still be emitted if a single monitored item was among the two items that were switched. This proposition, however, does not adequately account for the findings in experiment 2, in which change detection performance for a pairwise object switch trials was inferior to a two-object substitution trials, despite equating the number spatial locations undergoing a change.

One might argue that participants may have examined only one of the four items displayed (i.e., localized their attention), and hence had only a 25% chance of having witnessed a visual change in object substitution trials vs. a 50% chance of having witnessed a visual change in object switch trials. However, this hypothesis does not adequately explain the lower performance in one-object substitution trials relative to two-object switch trials, since, if *only* restricted spatial vigilance were an overriding strategy, one would expect similar performance in two-object substitution and object switch trials, since the chance of having attended a changed item is equal in both cases. This was not found. It is however, possible, that in both experiments, participants may have monitored a restricted number of spatial locations (e.g., one or two) and, in parallel, monitored changes in the summary statistics of the whole display, giving rise to an alternative data interpretation to the OFT. Any task benefit due to changes in the summary statistics between the visual displays may be evident only where at least two new items substitute two original items. These data support Wolfe, Reinecke and Brawn (2006) claim that selective and non-selective pathways affect processes underlying change detection.

In both experiments, the number of stimuli presented was set to four, corresponding to the commonly cited capacity of VSTM (Luck & Vogel, 1997; Pashler, 1988, Cowan, 2001). However, the number of items stored, according to Cowan's formula [Number of items stored, K = (Hit Rate + Correct Rejection Rate − 1) × Number of items presented], was < 3, in agreement with several earlier studies using similarly complex stimuli (Alvarez & Cavanagh, 2004; Olsson & Poom, 2005).

Had we used fewer than four stimuli, our analyses would have been complicated by ceiling effects in VSTM performance, which could mask genuine differences in performance between object switch and substitution trials (in the present study, average performance was already 78% for two-object switch and 83% for two -object substitution in experiment 2). Furthermore, if set size were reduced to 2 (the observed VSTM capacity produced by Cowan's formula), the probability that participants could detect a change for two-object switch trials would be theoretically 100%, even if only a single spatial location were monitored. Using more than four stimuli is likewise inappropriate, since it not only exceeds the commonly cited 3-4 item capacity (Luck & Vogel, 1997), but would also increase the disparity between the number of stimuli displayed and VSTM capacity, thereby increasing guess rate, and/or may result in limited memory resources being divided across a greater number of items than can be stored with a sufficient fidelity to support subsequent recognition (Wilken & Ma, 2004). Increasing the number of stimuli changed (switched/substituted) beyond two, whilst keeping the number of stimuli presented at four, would also have increased performance to ceiling levels as a result of the substantial visual modification produced by updating the majority of the display. For these reasons, experimental permutations using fewer or more than four stimuli, or entailing the manipulation of more than two stimuli, were not run.

A similar investigation to our experiment 1 was conducted in experiments 3 and 4 of Simons (1996), wherein five novel shapes stimuli were presented that could change between two display intervals in one of three ways: (i) a one-object substitution; (ii) a two-object switch; (iii) a configuration change (wherein one item was moved to a new location). ISIs of either 250 ms or 4.3 s were used. It was found that object switches produced superior performance to object substitutions at 250 ms ISI (at 4.3 s, no significant difference in performance between switch and substitution was found, although configuration changes produced far superior performance in both experiments). Our study, which also showed greater performance for two-item switch *vs*. one-item substitution, using a 1000 ms ISI, extends Simons' (1996) study by comparing change detection performance when the effective screen area changed is equated between change types, enabling us to report a reversal in this effect for a two-object substitution *vs*. a two-object switch. This demonstrates that the underlying cause of the difference in performance between switch and substitution trials is not purely attributable to net visual modification, a potential confound that was not considered in Simons (1996) study. Experiment 1 of the present study also partly replicates Simons (1996) experiment 4 for confirmation of the reliability of the reported effect, since of the 5 experiments in that study, only experiment 4 showed a significantly greater performance for two-item switch over one-item substitution. Other experiments (1, 2 and 5) showed superior performance for one-item substitution over two-item switch, but for which the comparison is complicated by issues including whether or not a verbal suppression task was used, the differences in the ISI, and type of stimuli used.

Change blindness has been shown to exist across a variety of situations, ranging from sparse displays wherein small numbers of items are presented against an empty background (e.g., Simons, 1996), to displays rich in visual information, such as photographs of indoor scenes (e.g., Hollingworth, 2003; Hollingworth & Henderson, 2002). Furthermore, change blindness may occur for passively viewed scenes, where items remains static between displays (such as those used in the present study), and for real-world interactive scenes, such as those used by Levin et al. (2002) in which a conversational partner is surreptitiously switched mid-discourse. The results of this study indicate that, for a sparse display, change blindness may be more severe where a new object replaces a previously viewed object, relative to where two previously viewed objects exchange their locations (at least where the objects are of roughly equal size and visual saliency). In contrast, change blindness may be more pronounced where two previously viewed objects exchange locations, relative to where two previously viewed objects are replaced by two new objects, presumably as a consequence of their attention-grabbing impact upon scene gist.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, **15**, 106-111. doi: 10.1111/j.0963-7214.2004.01502006.x.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436. doi: 10.1163/156856897X00357.

Bright, P., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2005). The anatomy of object processing: The role of antero-medial temporal cortex. *The Quarterly Journal of Experimental Psychology*, **58(B)**, 361-377. doi: 10.1080/02724990544000013.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, **24**, 87-114. doi: 10.1017/S0140525X01003922.

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and contour interaction. *Journal of the Optical Society of America*, **53**, 1026-1032. doi: 10.1364/JOSA.53.001026.

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, **29**, 388-403. doi: 10.1037/0096-1523.29.2.388.

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **32**, 58-69. doi: 10.1037/0278-7393.32.1.58.

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object Arrays. *Journal of Experimental Psychology: Human Perception and Performance*, **33**, 31-47. doi: 10.1037/0096-1523.33.1.31.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 113-136. doi: 10.1037/0096-1523.28.1.113.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, **23**, 420-456. doi: 10.1016/0010-0285(91)90015-G.

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements, in Attention and Performance XVI: Information Integration in Perception and Communication Eds T Inui, J L McClelland (Cambridge, MA: MIT Press) pp 125-155.

Irwin, D. E., & Zelinsky, G. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, **64**, 882–895. doi: 10.3758/BF03196793.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, **43**, 346-354. doi: 10.3758/BF03208805.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object specific integration of information. *Cognitive Psychology*, **24**, 175 -219. doi: 10.1016/0010-0285(92)90007-O.

Levin, D. T., Simons, D. J., Angelone, B. L., & Chabris, C. F. (2002). Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology*, **93**, 289-302. doi: 10.1348/000712602760146224.

Luck, S. J., & Vogel, E. K. (1997). Capacity of visual working memory for features and conjuctions. *Nature*, **390**, 279-281. doi: 10.1038/36846.

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, **6**, 296-305. doi: 10.1016/j.tics.2005.04.010.

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences*, **102**, 8776-8780. doi: 10.1073/pnas.0500810102.

Pashler, H. (1988). Familarity and visual change detection. *Perception & Psychophysics*, **44**, 369-378. doi: 10.3758/BF03210419.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442. doi: 10.1163/156856897X00366.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, **16**, 283-290. doi: 10.3758/BF03203943.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, **33**, 993-999. doi: 10.1016/0042-6989(93)90081-7.

Rensink, R. A. (1996). Change detection. *Annual Review of Psychology*, **53**, 245-277. doi: 10.1146/annurev.psych.53.100901.135125.

Rensink, R. A. (2000). Seeing, sensing and scrutinizing. *Vision Research*, **40**, 1469-1487. doi: 10.1016/S0042-6989(00)00003-1.

Sapkota, R., Pardhan, S., & van der Linde, I. (2011). Object-position binding in visual short-term memory for sequentially presented unfamiliar stimuli. *Perception,* **40**, 538-548. doi: 10.1068/p6899.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science,* **7**, 301-305. doi: 10.1111/j.1467-9280.1996.tb00378.x.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardised set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 174-215. doi: 10.1037/0278-7393.6.2.174.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, **31**, 137-149. doi: http://dx.doi.org/10.3758/BF03207704.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature,* **428**, 751-754. doi: 10.1038/nature02466.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology. General,* **131**, 48-64. doi: 10.1037/0096-3445.131.1.48.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, **4**, 1120-1135. doi: 10.1167/4.12.11.

Wolfe, J. M., Reinecke, A., & Brawn, P. (2006). Why don't we see changes? The role of attentional bottlenecks and limited visual memory. *Visual Cognition*, **14**, 749-780. doi: http://dx.doi.org/10.1080/13506280500195292.

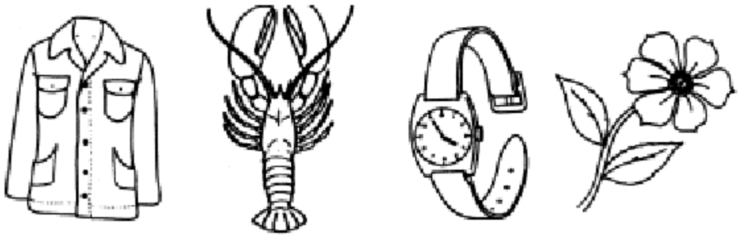

Figure 1. Example stimuli.

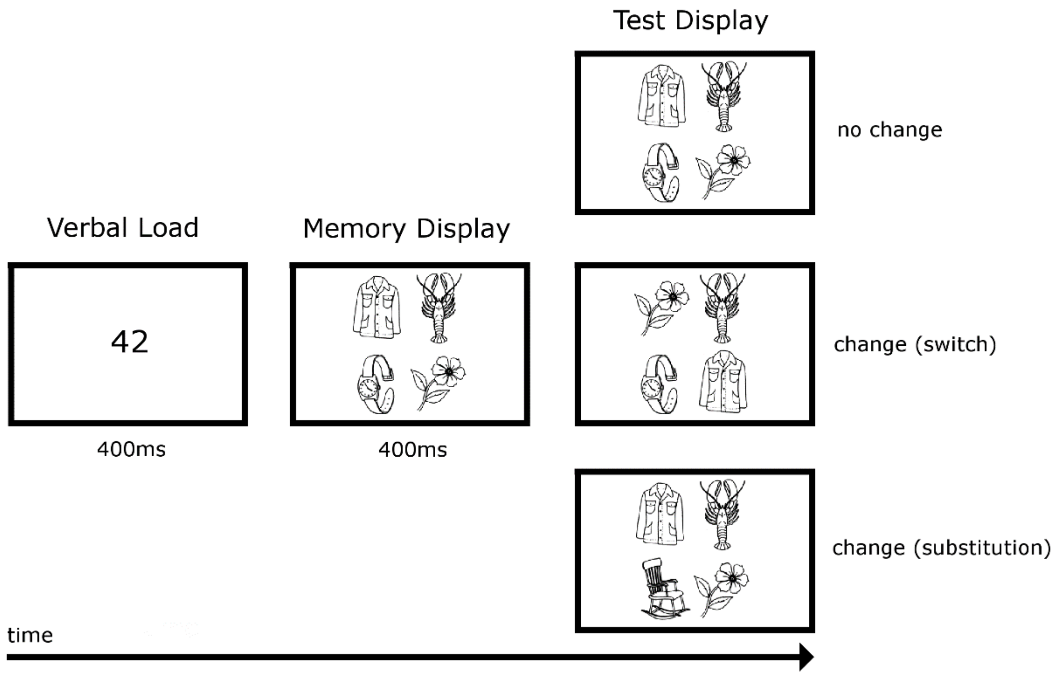

Figure 2. Schematic representation of the procedure for no change, object switch and object substitution trials in experiment 1. In any given trial, only one change type (i.e., no change, object switch, or object substitution) occurred. An identical procedure was used in experiment 2, except that, in object substitution trials, two memory objects (instead of one) were substituted by two new objects in the test display.

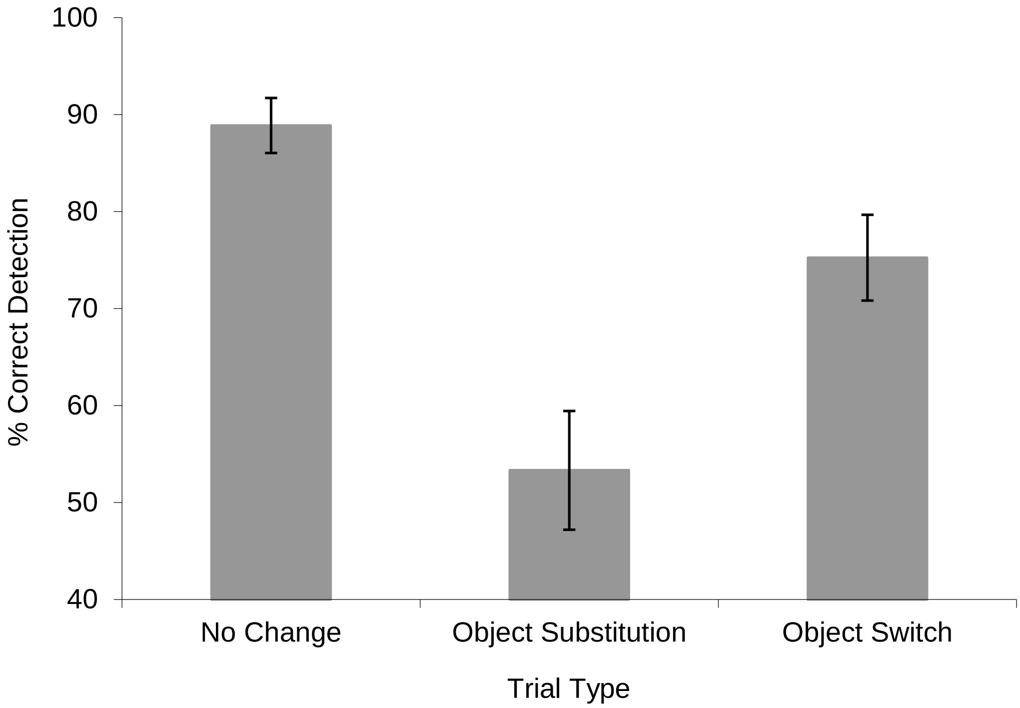

Figure 3. Percent correct detection across trial type in Experiment 1. Error bars represent ±1.96SE.

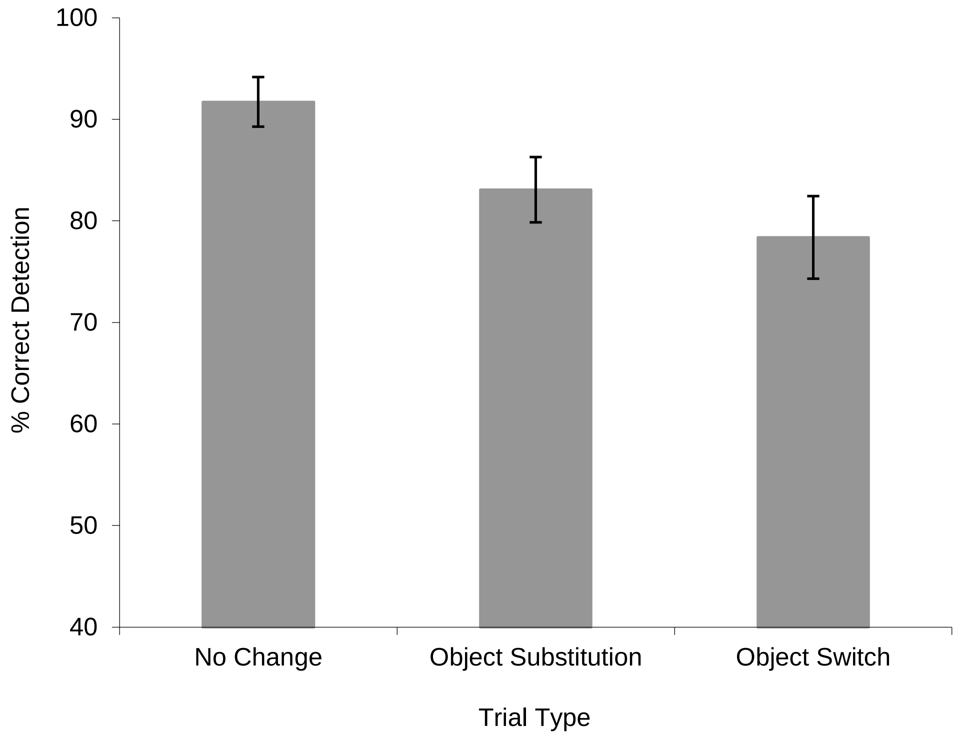

Figure 4. Percent correct detection across trial type in Experiment 2. Error bars represent ±1.96SE.