Perception, 2011, volume 40, pages 299 – 316

doi:10.1068/p6892

Flexible resource allocation for the detection of changing visual features

Alex Burmester, Guy Wallis
Perception and Motor Control Laboratory, Department of Human Movement Studies, University of Queensland, St Lucia, Brisbane 4072, Australia; e-mail: alexburm@gmail.com; website: http//hms.health.uq.edu.au/vislab
Received 16 December 2010, in revised form 19 February 2011; published online 18 April 2011

Abstract. Failure to detect change under circumstances where visual input is interrupted or attention is distracted is indicative of the capacity limits of visual short-term memory. The current study attempts to probe the nature of these limits. In experiment 1, the appearance of single Gabor patches was altered across colour, size, or speed, and set size was manipulated by means of a visual cue. In experiment 2, performance for detecting single and multiple changes to Gabor patches was compared under the constraint that the inherent detectability of each individual change was the same. Experiment 1 yielded a particular set size (4) and a particular level of change magnitude at which performance was equivalent across change type. On the basis of these parameter values, experiment 2 revealed that the detectability of two features changing within one object was the same as the detectability of a single feature changing across two objects, and that this level of detectability could be predicted by a simple model of probability summation. Together, these results suggest that performance is determined by the magnitude of featural changes independently of the way they are distributed across objects. We suggest they are adequately explained by a flexible-resource-allocation model rather than a slot-allocation model.

1 Introduction

We consider here the question of how information is represented in visual short-term memory (vSTM). Two competing classes of models that both seek to answer this question can be referred to as the flexible-resources and fixed-slots models respectively (Bays and Husain 2008). Flexible-resources models suggest that a dynamic resource is shared out amongst items that are to be retained in vSTM, while fixed-slots models suggest that each item is allocated one or more dedicated slots for processing [where each slot has a fixed (Luck and Vogel 1997) or a variable (Luck 2008) resolution]. There are several member models for each class (eg the 'slots plus averaging' model of Zhang and Luck 2008).

In the current study, we seek only to distinguish between the two broader classes of models (ie flexible resources versus fixed slots). To this end, we present evidence that supports a flexible-resource account of vSTM by showing that two changes (of the same or different features) taking place across two objects is as easily detectable as two changes taking place within one object and, furthermore, that this level of detectability can be predicted by a simple model of probability summation for independent channels. This shows that performance for detecting change is critically dependent on the degree of featural change and the number of features changing, independently of the number of objects. This supports the notion that it is the complexity of the objects changing that dictates detection performance (see Alvarez and Cavanagh 2004), rather than the number of objects changing independently of their complexity (as in Luck and Vogel 1997).

Like many of the studies looking at how information is represented in vSTM, the current study uses a change-detection paradigm in which an observer must search through a small number of objects and decide whether or not a change occurs across a blanking of the screen. The following sections deal with the change-detection and visual-search paradigms and how they can be used to look at how information is represented in vSTM.

1.1 *Change detection and change blindness*

Evidence suggests that we do not have the ability to process all parts of the visual scene simultaneously (Rensink 2002). An important cue to orient attention and cognitive resources is change. For this reason, vision scientists have been studying change detection for several decades. Typically, change detection is studied by adjusting the salience of a change over space and/or time (eg by introducing visual distractors at the moment of change, or by making the change across a longer period of time). The study of change detection has led to the discovery of the change-blindness phenomenon.

Change blindness refers to the difficulty observers have in detecting (otherwise obvious) changes to a scene when these changes are obscured, made gradually, or made simultaneously with the onset of distracting information in the scene (Simons and Levin 1997). Change blindness has been shown to occur when changes are masked by a blank screen (Rensink et al 1997); or occur simultaneously with mudsplats (a large number of blobs randomly placed across the display) (O'Regan et al 1999), eye-blinks (Grimes 1996), or saccades (McConkie and Currie 1996). The effect has also been shown when the changes are made gradually over time (eg when the change involves adding something to the scene, it can be made by 'fading-in' or increasing the contrast of the changing region gradually over time) (Simons 2000).

All of the above techniques are thought to impair the localisation of the motion signal that accompanies a change (Simons and Rensink 2005). For instance, a blank screen, an eye-blink and a saccade all interrupt the input of visual information at the time of the change. In contrast, a mudsplat floods the visual system with transient motion signals, impairing selection of any individual signal. It is the single, unambiguous motion transient accompanying a change under normal conditions that is thought to attract focused attention to the location of a change. Given that the conditions that induce change blindness necessarily involve adding noise to this signal, and consequently misdirecting attention (or at least preventing its direction) the phenomenon is often cited as support for the idea that focused attention is critical for detecting changes in a scene (Rensink 2000, 2002). Rensink (2002) suggests that this property of the phenomenon can be used to map out the nature of visual attention. However, early change-detection studies were originally used to map out the nature of vSTM. It should be noted here that the difference between visual attention and vSTM may be arbitrary, or dependent on perspective. For instance, vSTM is a construct that allows people to focus on the more static aspects of change detection (eg maintenance) while visual attention is a construct that allows people to focus on the more dynamic aspects (eg loading or comparison) (Rensink 2010).

1.2 *Change blindness and vSTM*

Change-blindness experiments typically follow a sequence $A - B - A'$, where A is the scene without the change, B is the transformation (eg blank, mudsplat) during which the change occurs, and A' is the scene with the change. This sequential-presentation paradigm was first used in experiments trying to determine the functional properties of visual memory. For instance, Phillips (1974) presented observers with an array of black and white squares and changed one of the squares across the delay period. For short-delay periods (up to 600 ms), accuracy in detecting the change was found to be constant and insensitive to the complexity of the pattern. For long delay periods (greater than 600 ms), accuracy decreased with time and became sensitive to the complexity of the pattern. This dissociation was taken to reflect the operation of two distinct systems: iconic (or sensory) memory at shorter intervals and a more abstract vSTM at longer intervals.

Phillips introduced the notion that vSTM was severely capacity-limited. This has been verified by tasks in many different paradigms (Baddeley 2003). Therefore, the presence of a capacity limit is not debated. The nature of this limit is, however, still disputed.

The dominant model of vSTM (see Luck and Vogel 1997) suggests that only a limited number of items can be simultaneously represented in vSTM. A competing model suggests that it is not the number of items that is fixed, but the amount of a 'resource' that can be shared between items, biased by the selective attention of the observer (see Bays and Husain 2008).

In a study that provided support for the 'fixed item limit' model of vSTM, Luck and Vogel (1997) used a sequential comparison procedure in which observers were presented with several elements and instructed to monitor all of them for a change in a single feature (colour, orientation, presence of a gap, size) across a delay interval. A common pattern of results was found for all change types (see figure 1)—detection accuracy was relatively constant for set sizes of 4 and below, but decreased linearly for set sizes above 4. This same pattern was found when observers were required to monitor the objects for changes in either of two features simultaneously. Given that performance appeared to be determined by the number of objects monitored, rather than the number of features, Luck and Vogel (1997) suggested that vSTM codes at the level of complete, integrated visual objects rather than at the level of visual features. However, both Wilken (2001) and Wheeler and Treisman (2002) replicated the critical conditions of Luck and Vogel (1997) and found support for the idea that it was the number of features changing, rather than the number of objects changing that dictated performance.

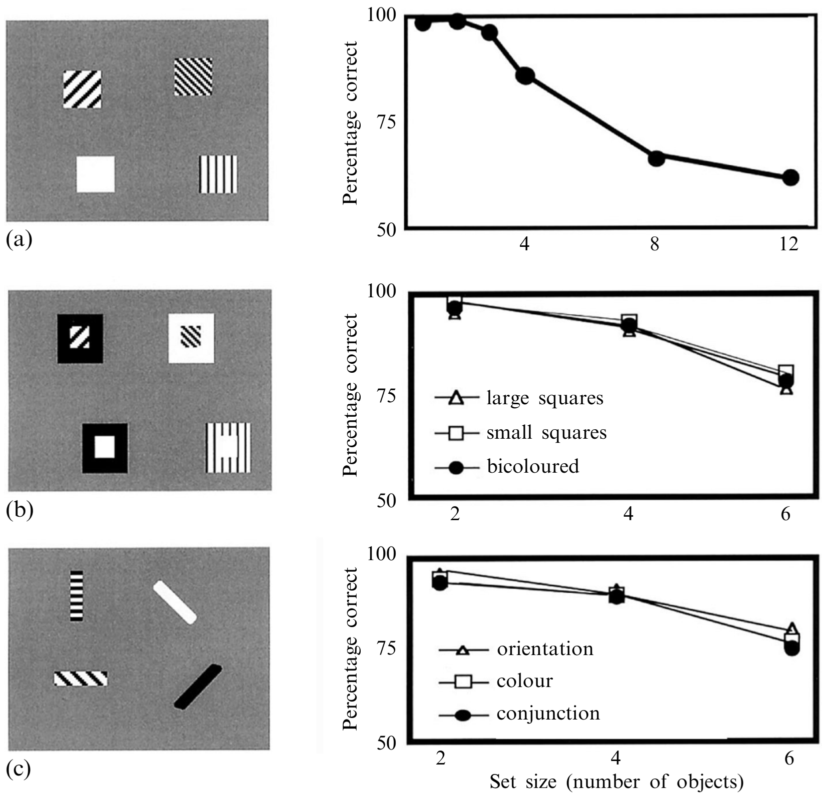

Figure 1. An example of stimuli (left) and data (right) from one of the experiments in the Luck and Vogel (1997) study.

The question of whether vSTM codes information at the level of features or objects has also been examined by Xu (2002). She used a conjunction display in which features were grouped into objects, and a disjunction display where they were separated into different objects. She found an object-based advantage only when the object contained features from different dimensions (eg colour and orientation) and not when the objects contained features from the same dimension (eg two overlapping oriented bars) (see also Wing and Allport 1972 and Duncan 1993). Olson and Jiang (2002) also found evidence to support this idea, by showing that orientation – colour conjunctions were remembered as accurately as the single features, but that colour – colour conjunctions were not. Olson and Jiang (2002) suggest their results support the 'weak-object' hypothesis of vSTM, which suggests vSTM is limited by the number of features to be remembered, but that this limitation can be alleviated by the grouping of features into 'chunks' or conjunctions. Olson and Jiang (2002) distinguished this hypothesis from several others, including the 'strong-object' hypothesis and the 'multiple-resources' hypothesis. The strong-object hypothesis suggests that vSTM is limited only by the number of objects to be remembered (see Luck and Vogel 1997 and Vogel et al 2001). The multiple-resources hypothesis suggests that vSTM is limited by the number of features to be remembered, regardless of whether they are separate or conjoined into objects. It suggests there are separate pools of resources for each feature type. This notion has been supported in the visual-search paradigm, in which observers must search for a target defined by a difference (in one or more features) to a background of distractors (see Wolfe et al 1990).

1.3 *Distinguishing between hypotheses regarding the nature of the vSTM representation*

The studies discussed above involved manipulating the number of features and the number of objects that changed over a blank interval. In the current study we also aim to do this, but do so using magnitudes of change that have been equated in terms of their detectability when presented alone. This is an important control that has not been employed by other studies looking at the detectability of multiple changes (eg Luck and Vogel 1997; Xu 2002). Experiment 1 is designed primarily to establish a magnitude of change for the three change types (colour, speed, and size) that is equivalent. Experiment 2 then uses this 'baseline' magnitude of change to compare changes in one or two features across one or two objects. This is done because the different hypotheses regarding the vSTM representation make different predictions for detection performance in these conditions.

A strong-object hypothesis would propose that detection performance will decrease with the number of objects changing independently of the number of features changing. For instance, performance should be the same for one object changing one feature as for one object changing two features. In contrast, a multiple-resource account would suggest that one object changing on one feature will have a lower detectability than one object changing on two features. However, the weak-object hypothesis would also predict that one object changing on two features will be more detectable than one object changing on one feature, as this hypothesis suggests that "vSTM may be limited in the number of features to be remembered" (Olson and Jiang 2002, page 1056). For single objects changing then, the weak-object and multiple-resources hypotheses cannot be distinguished. This distinction can be made, however, in the case of two objects changing.

The multiple-resources hypothesis suggests that two objects each changing on a single feature should be as detectable as a single object changing on two features. The weak-object hypothesis predicts "superior performance in the conjunction condition relative to the feature condition, but the benefit may be less than doubling vSTM capacity for features" (Olson and Jiang 2002, page 1056). This statement suggest that two features changing with one object will produce greater performance than two features changing across two objects, as the conjunction condition (two features within one object) should yield superior performance compared to the feature condition (ie where the same number of features change, but are not conjoined in a single object).

Comparing two features changing across two objects with two features changing within one object should allow us to distinguish between the multiple-resources and the weak-object hypotheses.

1.4 *The current study*

The current study had three primary aims:

(i) To investigate the relationship between the detectability of changes, target – distractor discriminability, and set size;

(ii) To determine a level of change magnitude and set size where performance is equivalent across different types of change;

(iii) To investigate the detectability of multiple changes, and compare these for cases where multiple changes take place within and between objects, where the magnitude of each featural change is equivalent.

The current study uses a display paradigm designed to manipulate the attentional load independent of the sensory load (see Palmer et al 2000). This is done to ensure that inferences about vSTM (ie capacity limits and the nature of its code) aren't contaminated by low-level variation in the visual field, but are instead only dependent upon top – down task constraints.

Two experiments were conducted and the basic stimulus configuration (6 Gabors surrounding fixation) and response type (yes/no) were kept constant. In experiment 1, participants were required to detect a change in a single feature (either colour, speed, or size) for different cued set sizes. In experiment 2, the change could occur on either of two features (across one or two objects) on any given trial.

2 General method

In many visual-search experiments, the target is different from all distractors on some featural dimension such as colour, size, orientation, or spatial frequency. In both experiments in the current study, however, targets were elements that underwent a change along some featural dimension across a blank interval (ie before the blank interval they were unchanged, and after it they were changed). This paradigm has been called 'visual search for change', as it involves searching for a target that is defined by a change (Rensink 2000).

2.1 *Equipment and stimuli*

The experiment was run on an SGI Onyx 300 machine using custom software. The experiments were displayed on a Sony Trinitron 24 inch monitor and the display size was 43 cm × 29 cm, with participants sitting with their eyes 100 cm from the surface of the display. The display therefore subtended 23.72 deg × 15.94 deg of visual angle.

On each trial 6 circular Gabors were presented in a ring around fixation. This ring had a radius of 5.0 deg, meaning the centre of each Gabor was 5.0 deg from the centre of the fixation cross, and 5.0 deg from the centre of either of the two adjacent Gabors.

In addition, a cue was presented at fixation at the beginning of each trial. The cue was a miniature representation of the locations that the observer needed to attend to. It contained a series of black blobs, one for each to-be-attended element. The centre of each black blob was 0.33 deg from the centre point, and the diameter of each blob was 0.22 deg. Each blob lay along an imaginary line from the centre of the screen to the centre of the element that blob represented.

All experiments involved variation along one or more featural dimensions in each trial. The program controlling stimulus presentation explicitly coded size, orientation, colour on the red – green axis, speed of motion, direction of motion, and spatial frequency. In real-world visual stimuli, many featural dimensions covary (eg colour and luminance). To ensure that variation along featural dimensions could be controlled, in the

current study we employed the following measures in order to minimise covariation across features:
(i) To control covariation of colour and luminance, all elements were made isoluminant (illuminance of 22 lux).
(ii) Given that the motion component orthogonal to the orientation of an object is its most salient, Gabors could only move in the two directions (left and right) orthogonal to their orientation (vertical).
(iii) The Gabors were always defined with a window of 1.8 deg or more, and at least two bars were visible. This was to ensure sufficient motion information was available in all different forms in which the Gabor appeared (ie to ensure changes in spatial frequency and/or size did vary perceived speed for a given actual speed).

2.2 *Determining colour points*

Because the colours used had to be isoluminant, there was no on-line modification of colour values in any experiment—all colour values were preset. To do this, the red and green values of the *RGB* computer colour space were varied so that a series of points were created that were progressively less red and more green, but all of the same lightness value (as read by a Minolta CL-100 colorimeter operating in CIE 1976 $L*a*b$ colour space, where L is lightness). The use of this procedure means that the space between consecutive colour points is not uniform on a metric scale—the placement of consecutive points was instead made on the basis of colour appearance. In other words, points were created so that each subsequent point appeared (to the experimenter) to be as distinct from the previous as other pairs on the scale and all points were made isoluminant. Each colour point was measured in the CIE $L*a*b$ colour space[(1)] where L is lightness, and a and b are colour coordinates on each of the two colour-opponent dimensions (a is red–green and b is blue–yellow). These coordinates were converted as shown in equation (3) to give values in the $L*C*h$ colour space (C and h represent chromaticity and hue) so that each colour point could be described as a single point—its h value.

$$L \text{ (lightness)} = L \quad (1)$$

$$C \text{ (chromaticity)} = \sqrt{a^2 + b^2} \quad (2)$$

$$h \text{ (hue)} = \tan^{-1}(b/a) \quad (3)$$

2.2.1 *Training.* Before taking part in the main experiment, during which data were collected, each participant took part in a short training run. The trials in the training run were the same as the trials in the experiment (albeit in a different order), but only a maximum of 30 trials was run (the number of trials run depended on how quickly the participant became comfortable with the task). Also, the to-be-attended element in the training phase had boxes around them, with a red box identifying the target and white boxes identifying non-targets.The purpose of this was to show clearly to participants what the changes looked like, and how the cue at the start of each trial was identifying to-be-attended elements.

3 Experiment 1. Change to stationary Gabor stimuli using cues to manipulate set size

3.1 *Introduction*

In this experiment, the set size was manipulated by having the same number present on each trial, but cuing a different subset of those elements at the beginning of each trial. This is a cued-set size or relevant-set size (see Palmer et al 1993) manipulation.

[(1)] This colour space is also referred as to '$L*u*v$' and 'CIELAB'. It is derived from the CIE 1931 *XYZ* colour space and was created to be more perceptually relevant and uniform than the 1931 colour space.

It keeps the amount of sensory information relatively constant on each trial, and has an advantage over reducing stimulus numbers in that the set-size manipulation can be considered more as an attentional manipulation (ie manipulating attentional load), rather than a combined attentional and sensory manipulation. Furthermore, manipulating set size independently of the number of elements displayed keeps statistical decision noise constant and so means results are not contaminated by this effect.

3.2 *Aims*

(i) To determine the effect of changes in set size on the detection of a single element changing on a single dimension.
(ii) To determine a single set of conditions suitable for comparing single feature changes with multiple feature changes occurring simultaneously across several objects and within the same object.

3.3 *Predictions*

(i) It was predicted that set-size effects would be similar for colour, speed, and size, in agreement with Palmer (1994).
(ii) It was predicted that increases in change magnitude would produce corresponding increases in performance (following Scott-Brown and Orbach 1998) and, further, that the slope of the change magnitude – performance relationship would be similar for the different change types, owing to the magnitudes of change being roughly equated in terms of detectability through pilot testing.

3.4 *Method*

3.4.1 *Participants.* Nine participants took part in this experiment, seven males and two females. All were volunteers recruited from the department of Human Movement Studies at the University of Queensland. All participants were naive to the purpose of the study and were not experienced psychological observers. However, all participants were trained in the experimental task before data were collected. Details of the training phase can be found in section 2. The participants' mean age was 29.44 years (range 21 – 42 years).

3.4.2 *Procedure and design.* The progression of a trial in this experiment can be schematised as:

$$\text{cue} \rightarrow \mathrm{A} \rightarrow \mathrm{B} \rightarrow \mathrm{A}' \rightarrow \mathrm{B} \rightarrow \mathrm{A} \rightarrow \mathrm{B} \rightarrow \mathrm{A}' ,$$

where A is the pre-change presentation, A′ is the post-change presentation, and B is a blank screen. On trials in which no change occurred, the elements in A′ are the same as the elements in A. Fixation lasted 1500 ms, pre- and post-change presentations lasted 1500 ms. Blank screens were presented for 120 ms. Figure 2 shows an example trial.

Independent variables in this experiment were the presence of change (present or absent), the type of change (colour, speed, size), the cued set-size (1, 2, 4, or 6), and the magnitude of change (1 – 4 steps). Therefore, there were 48 unique trials in which a change occurred and 48 in which no change occurred. These 96 trials were repeated 4 times in 4 separate blocks and the trials were run in a random order in each block.

3.4.3 *Stimulus configuration.* The stimulus points are given below.
(i) Colour (hue): 0.98, 1.01, 1.04, 1.08, 1.11.
(ii) Size (deg): 1.80, 1.98, 2.18, 2.40, 2.64.
(iii) Speed (deg s^{-1}): 0.30, 0.42, 0.59, 0.82, 1.52.

The stimulus dimensions that did not undergo change—orientation, spatial frequency, and luminance—were kept constant throughout (orientation: vertical; spatial frequency: 1.5 cycles deg^{-1}). The values for size and speed lie on a logarithmic scale and follow Weber's law, as it was found in pilot testing that equally discriminable steps in size and speed followed this relationship.

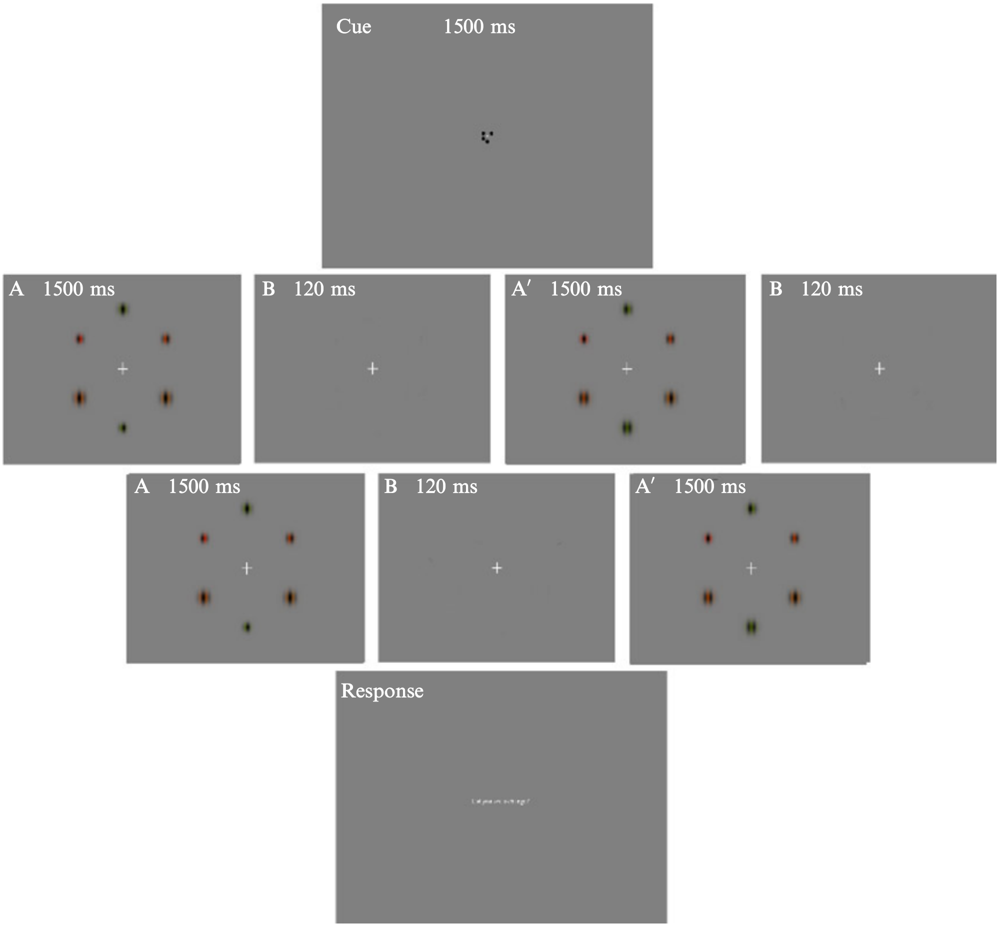

Figure 2. [In colour online, see http://dx.doi.org/10.1068/p6892] Example of a size-change trial. The changing element is the one directly below fixation point. It changes from a smaller-size value in A to a larger-size value in A′ (increase in size).

3.5 *Results*

3.5.1 *Hit rate versus proportion correct.* There are two obvious dependent variables (DVs) that can be used to analyse these data: hit rate and proportion correct. Hit rate is the proportion of trials in which a change is correctly identified (eg change trials in which the participant indicates there was a change). Proportion correct is the hit rate added to the rate of correct rejections. A correct rejection is when a participant correctly indicates on a no-change trial that there was no change. In addition, a false alarm is when a participant indicates there was a change on a no-change trial. Equations (4), (5), and (6) show how hit rate, proportion correct, and the number of correct rejections, respectively, are calculated.

$$\text{hit rate} = \text{hits}/\text{change trials}\,, \tag{4}$$

$$\text{proportion correct} = (\text{correct rejections} + \text{hits})/\text{trials}\,, \tag{5}$$

$$\text{correct rejections} = (\text{trials}/2) - \text{false alarms}\,. \tag{6}$$

Trials in which a correct rejection occurs are not unique in the way that hit trials are—they cannot be uniquely attributed to a category or 'cell' defined by a combination of independent variables (IVs) (change type, magnitude of change) as there is no change occurring in a correct-rejection trial. Therefore, any difference in performance when using proportion correct as a DV rather than hit rate will be random. This is because the difference is due only to the correct-rejection trials included in the analysis. These trials cannot be included on the basis of the combination of conditions occurring in that trial (as the trials are exactly the same regardless of what condition the program/experiment may attribute them to), but on a purely random basis instead. Therefore, hit rate is the more meaningful DV to use for the current experiment.

3.5.2 *Data.* Figures 3, 4, and 5 show how performance (ie hit rate) varies with change magnitude and set size for colour, size, and speed, respectively. These graphs show effects of set size and change magnitude for each change type.

Figure 6 shows how hit rate changes according to set size for each change type. The lines of best fit (from robust regression) and their standard errors are given in table 1. A set-size effect is present for all feature dimensions. The size of the effect is similar for size and speed and is smaller for colour.

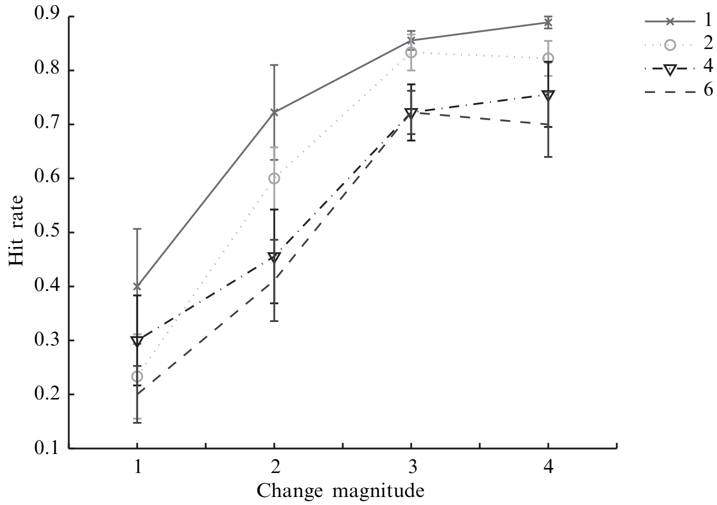

Figure 3. How performance changes with change magnitude and set size for colour changes. Error bars represent ±1 SE.

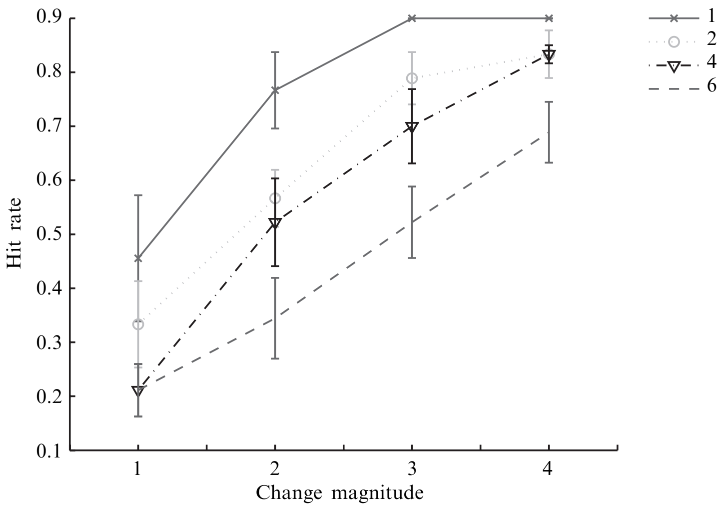

Figure 4. How performance changes with change magnitude and set size for size changes. Error bars represent ±1 SE.

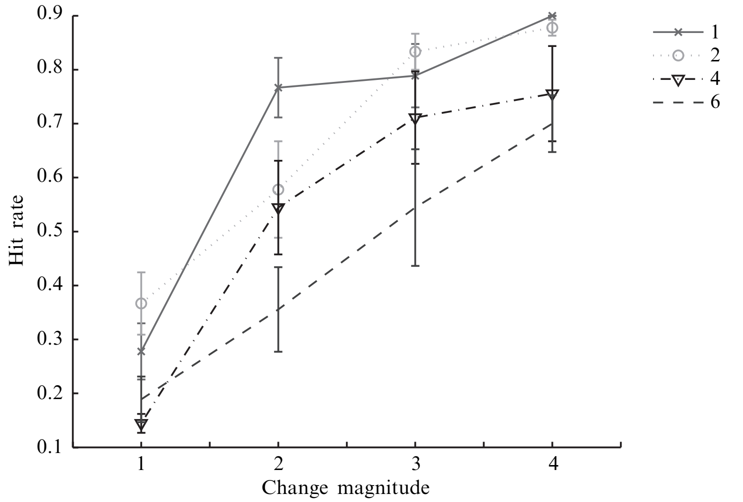

Figure 5. How performance changes with change magnitude and set size for speed changes. Error bars represent ±1 SE.

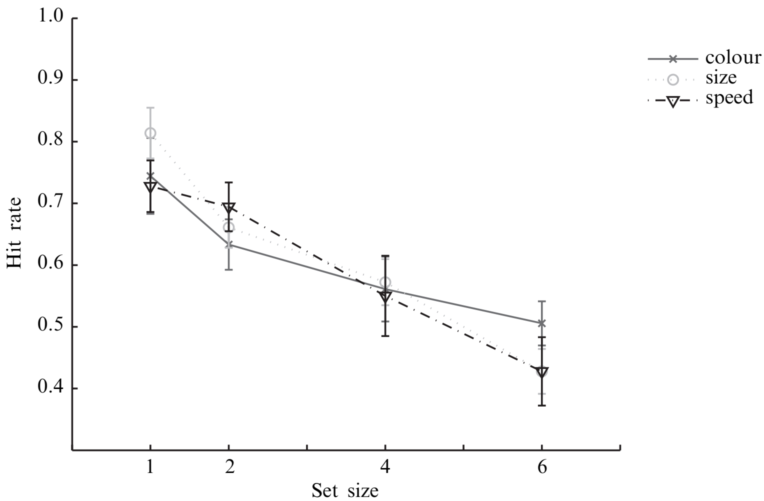

Figure 6. Set-size effects for the different change types. Error bars represent ±1 SE.

Table 1. Experiment 2. Set-size effect, slopes, and errors.

Feature	Slope	Error
Colour	−0.0443	0.0103
Size	−0.0713	0.0113
Speed	−0.0621	0.0040

It is apparent from the data that performance was substantially below ceiling for set sizes of 1. This perhaps indicates that the cuing was not totally effective and/or the constant presence of other elements in addition to those being attended to created noise in the data (there were always 6 elements present).

From figure 6, it is clear that performance is most similar across change type for a set size of 4. Figure 7 includes data only from trials with a set size of 4 and shows how performance changes with change magnitude for the three different change types. The variability across different change types is similar for magnitudes of 2, 3, and 4.

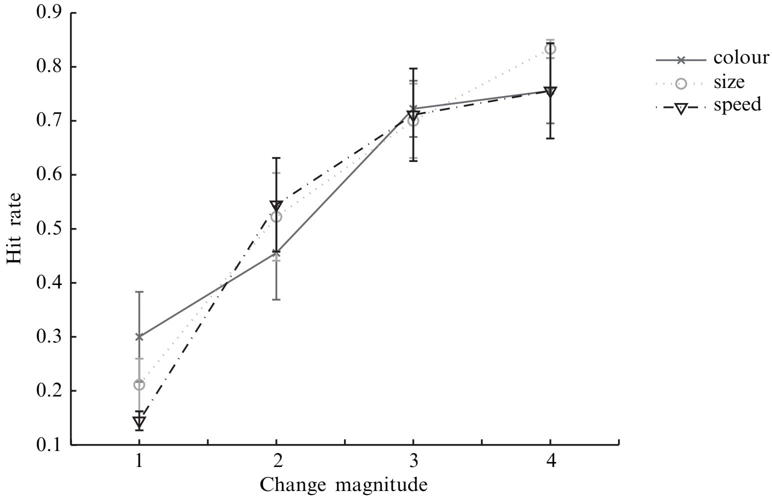

Figure 7. How hit rate varies with change magnitude for different change types when the set size is 4. Error bars represent ±1 SE.

The predictions of similar set-size effects across change type was supported. Effects of change magnitude were also similar for colour, size, and speed—this, however, was not quantified by slope calculation in the current experiment. The different set-size effects require further investigation. The results for set sizes of 1 were substantially below ceiling. This indicates that the constant presence of distractors created noise, be it sensory or decision noise, that reduced performance for what should be an easily detectable change. In Palmer's 1998 experiments, it was found that both the display set-size (ie no cue—the number of elements present is the same as the set size) and relevant set-size manipulations produced ceiling effects at a set size of 1 and, furthermore, yielded very similar set-size effects. The cue used in Palmer's studies, however, was different from that used in the current experiment. Palmer presented cues in the position of the elements, rather than at fixation (as in the current experiment). It is possible, therefore, that the cue presented at fixation is not as efficient as cues presented in element positions in terms of directing attention to those positions. This could create extra noise at both a sensory and a decision-based level in the visual system, as more noisy inputs are required to be integrated because they are not filtered out at a low enough level. Furthermore, Palmer measured set-size effects for individual observers and then averaged these, whereas in the current experiments set-size effects were generated from the average performance of all observers. The idea that the set-size manipulation may have been more effective with a better cue is important, especially if we aim to make detailed set-size effect measurements (eg using psychometric functions and psychometric slopes). This, however, was not the main aim of the current experiment.

The main aim of this experiment was to determine a set of conditions that produced equivalent levels of detectability across the three change types. The set size for which detectability across change type was most similar was 4. The equivalent levels of detectability were established so that a comparison can be made between single and multiple changes in the next experiment, where each stimulus increment of change is equivalent.

When multiple changes are made, it is possible that the overall detectability of this composite change will be additive. Therefore, it is important to choose a level of change magnitude that, as well as producing similar detectability across change type, produces a level of performance that is not so high that it could produce ceiling effects in the multiple change situation (if the effects are indeed additive). Also, the 'equivalence' level of detectability should be substantially above chance (50%) so that a reliable baseline of detectability can be established in the single-change case.

The combination of equivalence and being substantially below ceiling and above floor levels is fulfilled best by changes of magnitude 3 with a set size of 4 (see figure 7). Variability of hit rate for these changes is very small and the mean hit rate is around 0.75, which is ideal as it is halfway between floor (chance) and ceiling.

4 Experiment 2. Multiple changes within and between objects

4.1 *Aim*

The aim of experiment 2 was to look at interactions between different visual features (colour, size, and speed), where the detectability of each individual change was relatively similar. In order to do this, a restricted set of conditions from experiment 1 was employed (a constant set-size and a constant magnitude of change on each dimension).

4.2 *Predictions*

It was predicted that:

(i) Following Wilken (2001), performance would be determined by the number of features changing, rather than the number of objects changing (as in Luck and Vogel 1997). Accordingly, it was expected that two features changing within an object would produce the same level of performance as two features changing across two objects.

(ii) Following Luck and Vogel (1997), it was expected that different features changing across two objects would produce the same level of performance as the same feature changing across two objects.

4.3 *Participants*

Ten participants took part in this experiment, four males and six females. All were undergraduate students at the University of Queensland and were paid Australian $10 for their participation. Their mean age was 20.9 years (range 18 – 24 years). All participants were naive to the purposes of the study and were not experienced psychophysical observers. However, all participants were trained in the experimental task before data were collected. Details of the training phase can be found in section 2.

4.4 *Procedure and design*

Each trial followed the presentation pattern used in experiment 1. At the beginning of each trial, stimulus variables were configured in the same way as for experiment 1. However, in this experiment, more than one element and one feature could change, but the magnitude of that change was fixed at the levels obtained from experiment 1. The independent variables for this experiment were: number of objects changed (1 – 2), number of features changed (1 – 2), and feature type(s) changed (size, colour, speed). The experiment therefore followed a 2 (absence/presence of change) × 2 (number of objects changed) × 2 (number of features changed) × 3 (feature change type) design. When two objects changed, only one feature within each object would change. Therefore, when two objects changed, they could change on either the same feature (single-feature change) or each change on a different feature (double-feature change). The possible double-feature change combinations were: colour and speed, speed and size, and colour and size. Figure 8 shows an example of a trial in which one element changes size and one changes colour.

4.5 *Results*

Figure 9 shows the proportion of hits for the one- and two-object conditions where one and two features are involved. These data also show a level of detectability that would be expected from probability summation alone—detectability will increase with two targets present simply because more suprathreshold targets are present. In other words, observers are more likely to spot a target (ie produce a “yes” response) when

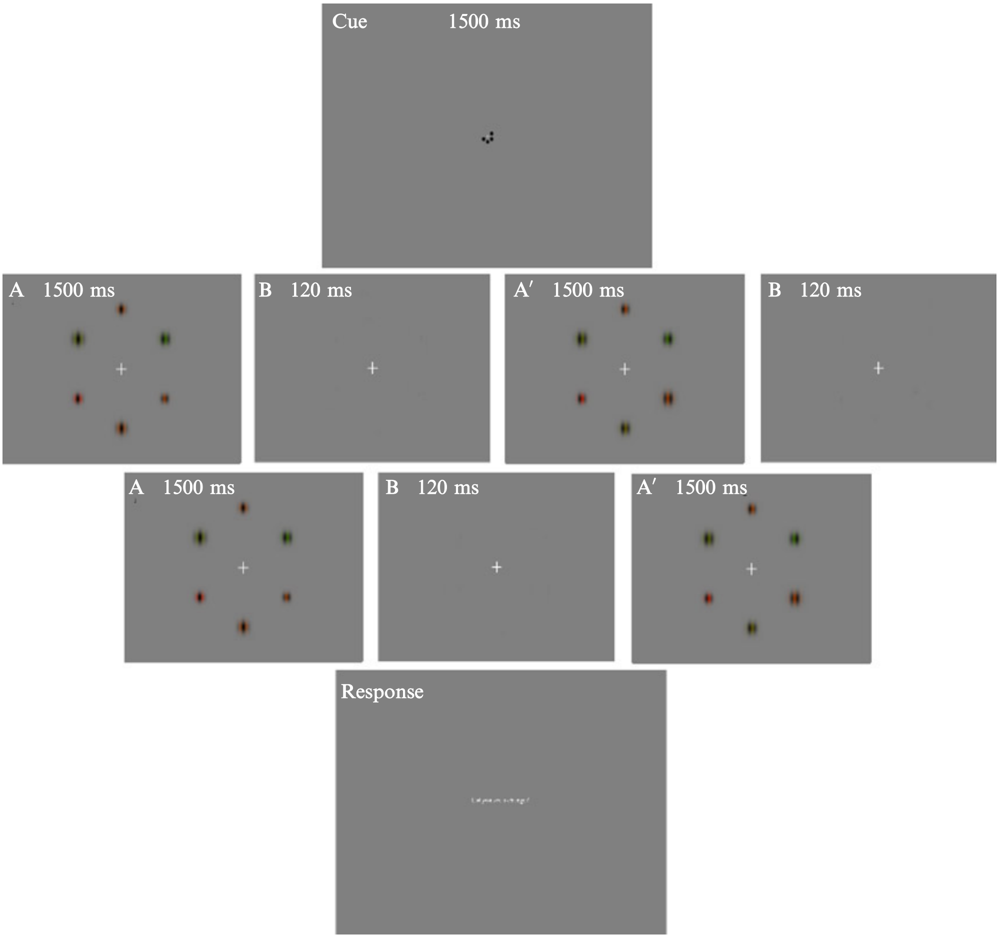

Figure 8. [In colour online] Example of a trial where one element changes size and one changes colour. One changing element is directly below fixation (colour) and one is diagonally down and to the right (size).

more targets are present in the search field. The calculation of probability summation is given in equation (7). This formula has been used in psychophysics detection studies to calculate the expected increase in detectability of a stimulus increment when two stimulus variables (eg contrast and orientation) are changing simultaneously (eg see Reisbeck and Gegenfurtner 1998). It is also the formula used to calculate statistical decision noise from added distractors in visual search. The formula effectively calculates the probability of either one of two independent channels detecting a signal, where the probability of one detecting a signal is p, the probability of it not detecting that signal is $1-p$, the probability of both not detecting a signal is $(1-p)^2$, and, therefore, the probability of either of them detecting a signal is $1-(1-p)^2$.

$$p_{\text{summ}} = 1-(1-p)^n \ . \qquad (7)$$

Two important results are shown in figure 9. First, there is no significant difference between the two double-feature conditions (ie when two features change within an object as compared to when they change across two different objects). Second, there is no significant difference between the probability summation level and all of the double object/feature conditions.

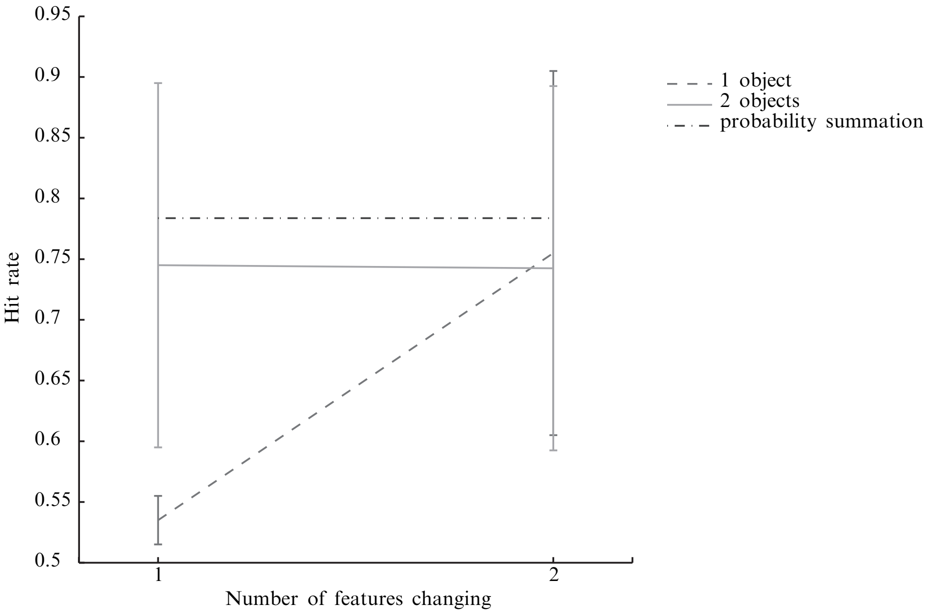

Figure 9. How the hit rate varies with the number of featural changes and number of objects changing. The level of performance expected from probability summation is shown. Error bars represent ±1 SE.

4.6 *Discussion*

The results indicate that multiple changes do not summate beyond the level expected from probability summation. This indicates that the detection of multiple changes can be accounted for by a simple model of parallel feature detection, where independent channels each have a certain probability of detecting a feature, and so performance for detecting multiple features can be predicted by probability summation alone. Furthermore, the lack of any difference between the two double-feature conditions (two features changing within an object or across two objects) supports the prediction that performance wold be determined by the number of features changing rather than the number of objects changing.

This result goes against the 'strong-object' assertion of studies such as Luck and Vogel (1997) that performance in vSTM tasks is based on the number of objects changing, rather than the number of features. If this assertion were correct, we would expect no difference between the detectability of a single object changing one feature and a single object changing two features. However, the criticism of Luck and Vogel's (1997) interpretation by Wheeler and Treisman (2002) and their results in favour of the feature-based interpretation of vSTM are more in agreement with the results of this experiment as are the results of Wilken (2001) experiments 1–4—it is the number of features changing, rather than the number of objects, that dictates performance. It is possible, however, that the difference between two features changing across two objects and two features changing within a single object is small enough that it needs a more sensitive measure than the one provided by this experiment.

Another possibility is that, for the current task, observers employed a serial-like search strategy in which only one object is searched at a time. If this were the case, multiple changes across objects would create no facilitation of detection performance, because only one of the changing objects was detected. Similarly, when a single object changes (regardless of whether it was on one or two features), it will be detected with the same level of performance as in the two-object change condition. However, it is important to note that multiple changes to a single object did improve performance substantially in comparison to single changes to a single object. This means that these changes somehow add to increase the salience of the single-object target. It is possible then that observers are employing an object-based search strategy, where multiple changes make a single object easier to detect, perhaps by improving the quality of its representation in vSTM. This account is similar to the 'weak-object' hypothesis proposed by Olson and Jiang (2002), whereby vSTM may be limited by the features changing (as opposed to the objects changing), but that this limitation can be alleviated by conjoining the features into objects. However, if this were the case, two featural changes to a single object should yield greater performance than a single feature changing on each of two separate objects, as this represents a comparison of a 'conjoined' double-feature condition with its equivalent unconjoined (ie features belonging to different objects) condition.

5 Discussion

The experiments reported in this paper address the question of how information is represented in vSTM by using a change-detection paradigm in which one or more objects changed one or more features (colour, size, or speed). In experiment 1 we examined the effect of change magnitude and set size on the detectability of a single object changing a single feature. A particular level of set size (4) and change magnitude that produced an equivalent level of detectability across the different features was chosen from experiment 1 to use in experiment 2. These fixed conditions are used in experiment 2 to compare one or two changes taking place in one or two objects.

5.1 *Experiment 1. Set-size effects and equivalent levels of detectability*

Experiment 1 yielded similar set-size effects and effects of change magnitude across the three types of change. This implies that (i) the detectability of one or more changes across the three stimulus variables was comparable; (ii) the range of change increments chosen from initial pilot testing were appropriate in that they produced roughly equivalent levels of detection. On comparing the effects of set size it became evident that a set size of 4 yielded the most similar levels of detectability across the different change types. Looking at the effects of change magnitude (ie number of steps) at a set size of 4, it was found that a change magnitude of 3 yielded the most similar levels of detectability across the different change types. Furthermore, this level of delectability (hit rate) was close to 0.75, which is halfway between the levels expected from chance responding and from perfect responding. This makes it ideal as a baseline level of change magnitude to use when comparing single and multiple changes.

In section 1, several studies were examined that looked at the detectability of one or two changes in one or more objects. None of these studies, however, controlled the magnitude of each individual featural change so that they were somehow equivalent. This control is important to quantitatively compare changes to different features and changes between single and multiple features. In the current study, this control was employed. However, experiments 1 and 2 each used a different set of participants, with a large difference in their mean ages. Ideally, the same participants should be used in both experiments. Also, individual detection thresholds could be set for each observer, rather than using a mean threshold from all observers. Given that the standard error sizes were small for the effects of set size and change magnitude, this is unlikely to be a significant issue for the current set of results. It is an issue that should be given consideration in future research, however.

5.2 *Experiment 2. The effect of multiple changes*

In experiment 2, on half of the trials, one or two objects changed one or two features. Performance was examined as a function of the number of feature types changing and the number of objects changing. It was found that performance could be predicted solely on the basis of the number of features changing, and that when two features changed (within a single object, or across two objects), this level of performance could be predicted by a simple model of probability summation for independent featural channels (ie based on the baseline performance for a single object changing a single feature). This result can distinguish between various hypotheses about how information is represented in vSTM, at least for the current task. A 'strong-object' hypothesis would predict equal performance for a single object changing one feature and a single object changing two features, which would be below that for two objects. A 'weak object' hypothesis would predict that multiple features changing within one object (ie conjoined features) would be more detectable than two features changing across two objects (ie disjoined features). Neither of these models matched the observed performance. Instead, it appears that the results are best explained by a multiple-resources account. Furthermore, given that the multiple-resources account suggests that information for multiple-feature dimensions is processed independently, a probability summation model for independent channels should predict the performance for the conditions where multiple features are changing. This was also observed in the current study.

While these results support a multiple-resources account, it is important to note the differences between this study and studies supporting other accounts, as the interpretation of how vSTM codes information may change depending on the task employed. For instance, several studies that support object-based accounts of vSTM have presented a probe stimulus at the end of each trial to the observer, and asked the question: 'Did this object change?'. The use of a probe stimulus may help to focus the observer's attention on a particular part of the sample/match display and may therefore necessitate the binding of features into objects (Wheeler and Treisman 2002). However, as shown by Bays and Husain (2008), the results favouring the strong-object hypothesis can also be produced by a multiple-resources model. This issue could be addressed by using the current task with a range of different response types, and looking at the pattern of data for each response type. This will be attempted in a future study.

Another issue that lends itself to future research is the question of how changes on different featural dimensions compare in terms of their relative contributions to the multiple-change effects. That is, even though each featural change in experiment 2 was equated in terms of its detectability (from the data of experiment 1), do changes in different dimensions interact in different ways? Unfortunately, the data collected from experiment 2 do not lend themselves to this analysis. This is therefore an issue that will be dealt with in a later study.

One of the problems with the data from experiment 2 is the variability of the conditions where two features change (on one or two objects). This indicates that there could be a difference between these conditions that the measures from this experiment were not sensitive enough to pick up. However, given that the magnitude of changes was equated on the basis of data from experiment 1, we can have greater confidence that this is not the case.

5.3 *vSTM and capacity*

This study supports the idea that vSTM is limited in capacity by the number of features to be remembered rather than the number of objects. It supports a multiple-resource account over a strong-object or even a weak-object account of the vSTM representation. However, the task is somewhat unique in the literature on vSTM and this may account for the differences. For instance, this task is a simple detection task in which observers must detect a change across two very similar displays. Observers are not required to do anything more complex, such as identify or locate the change. In this situation, featural information could be sufficient for the detection of changes, while in a task requiring a more complex response or a response to a probed item, it may be necessary for featural information to be integrated into complete objects. This would create a different pattern of data that may support, say a strong-object or a weak-object hypothesis of the vSTM representation. If this were the case, however, it obviously means that the interpretation of the vSTM representation changes depending on the task under consideration. Therefore, we may be looking at a more flexible system than previously thought, where under some conditions it acts as if constrained by a multiple-resources model and under some conditions it acts as if constrained at the level of complete objects. We do propose, however, that at present a multiple-resource account is better able to account for the findings of the various vSTM experiments examined in section 1, as well as the results reported here. Further research is required to determine if the nature of vSTM stimulus encoding is dependent on the task and the required response type (ie identification, localisation, or detection).

6 Conclusion

In conclusion, in this study we looked at the effect of changing one or two features across one or two objects, where the detectability of each featural change had been equated, on the basis of an experiment that manipulated both the set size and the magnitude of change. We found that the detectability of changes in the various conditions was adequately explained by a multiple-resources account of vSTM and could not be explained by either the strong or weak-object hypotheses. We suggest that it may be the case that the interpretation of the vSTM representation will change depending on the task and that this would require a re-evaluation of current models of vSTM in favour of more flexible accounts. We suggest further that the multiple-resource account is the most flexible of the current models, as it can explain data previously taken to be evidence of the strong- or weak-object hypotheses.

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

Baddeley A, 2003 "Working memory: looking back and looking forward" *Nature Reviews Neuroscience* **4** 829 – 839

Bays P M, Husain M, 2008 "Dynamic shifts of limited working memory resources in human vision" *Science* **321** 851 – 854

Duncan J, 1993 "Similarity between concurrent visual discriminations: Dimensions and objects" *Perception & Psychophysics* **54** 425 – 430

Grimes J, 1996 *Perception. Vancouver Studies in Cognitive Science* chapter "On the failure to detect changes in scenes across saccades" (New York: Oxford University Press) pp 89 – 109

Luck S J, 2008 *Visual Memory* chapter "Visual short-term memory" (New York: Oxford University Press) pp 43 – 86

Luck S J, Vogel E K, 1997 "The capacity of visual working memory for features and conjunctions" *Nature* **390** 279 – 281

McConkie G W, Currie C, 1996 "Visual stability across saccades while viewing complex pictures" *Journal of Experimental Psychology: Human Perception and Performance* **22** 563 – 581

Olson I R, Jiang Y, 2002 "Is visual short-term memory object based? Rejection of the 'strong-object' hypothesis" *Perception & Psychophysics* **64** 1055 – 1067

O'Regan J, Rensink R A, Clark J J, 1999 "Change-blindness as a result of 'mudsplashes'" *Nature* **398** 34

Palmer J, 1994 "Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks" *Vision Research* **34** 1703 – 1721

Palmer J, 1998 *Visual Attention* chapter "Attentional effects in visual search" (New York: Oxford University Press) pp 348 – 388

Palmer J, Ames C T, Lindsey D T, 1993 "Measuring the effect of attention on simple visual search" *Journal of Experimental Psychology: Human Perception and Performance* **19** 108 – 130

Palmer J, Verghese P, Pavel M, 2000 "The psychophysics of visual search" *Vision Research* **40** 1227 – 1268

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

Reisbeck T E, Gegenfurtner K R, 1998 "Effects of contrast and temporal frequency on orientation discrimination for luminance and isoluminant stimuli" *Vision Research* **38** 1105 – 1117
Rensink R, 2010, personal communication
Rensink R A, 2000 "The dynamic representation of scenes" *Visual Cognition* **7** 17 – 42
Rensink R A, 2002 "Failure to see more than one change at a time" *Journal of Vision* **2**(7) 245a
Rensink R A, O'Regan J K, Clark J J, 1997 "To see or not to see: The need for attention to perceive changes in scenes" *Psychological Science* **8** 368 – 373
Scott-Brown K C, Orbach H S, 1998 "Contrast discrimination, non-uniform patterns and change blindness" *Proceedings of the Royal Society of London, Section B* **265** 2159 – 2166
Simons D J, 2000 "Current approaches to change blindness" *Visual Cognition* **7** 1 – 15
Simons D J, Levin D T, 1997 "Change blindness" *Trends in Cognitive Sciences* **1** 261 – 267
Simons D J, Rensink R A, 2005 "Change blindness: Past, present and future" *Trends in Cognitive Sciences* **9** 16 – 20
Vogel E K, Woodman G F, Luck S J, 2001 "Storage of features, conjunctions, and objects in visual working memory" *Journal of Experimental Psychology: Human Perception and Performance* **27** 92 – 114
Wheeler M E, Treisman A M, 2002 "Competition and sharing of processing resources in visual discrimination" *Journal of Experimental Psychology: Learning, Memory and Cognition* **131** 1603 – 1616
Wilken P, 2001 "Capacity limits for the detection and identification of change: Implications for models of visual short-term memory" unpublished PhD thesis, Department of Psychology, The University of Melbourne, http://theassc.org/files/assc/patrick_wilken_thesis.pdf
Wing A, Allport D A, 1972 "Multidimensional encoding of visual form" *Perception & Psychophysics* **12** 474 – 476
Wolfe J M, Yu K P, Stewart M I, Shorter A D, Friedman-Hill S R, Cave K R, 1990 "Limitations on the parallel guidance of visual search: Color × color and orientation × orientation conjunctions" *Journal of Experimental Psychology: Human Perception and Performance* **16** 879 – 892
Xu Y, 2002 "Limitations of object-based feature encoding in visual short-term memory" *Journal of Experimental Psychology: Human Perception and Performance* **28** 458 – 468
Zhang W, Luck S J, 2008 "Discrete fixed-resolution representations in visual working memory" *Nature* **453** 233 – 235

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 40 2011

www.perceptionweb.com