Perception, 2006, volume 35, pages 865 – 881

DOI:10.1068/p5526

Surface assignment modulates object formation for visual short-term memory

Árni Kristjánsson¶
Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1 3AR, UK
Received 4 March 2005, in revised form 10 September 2005; published online 30 June 2006

Abstract. Recent results have suggested that the operational units of visual short-term memory (VSTM) are whole objects, rather than features or the total amount of information to be remembered. Here, for the first time, the influence of surface assignment on object formation for VSTM was investigated. The observers had to memorize the features of four briefly presented (300 ms) two-part objects, followed by a mask and a cue indicating which object to report on. The experiments contrasted whether there were any apparent depth differences between the two parts of each object, and whether observers had to report on only one or both features of the post-cued target object. Depth differences induced with stereoscopic disparity, and with a pictorial depth cue (simple interposition of object features), interfered strongly with performance when both features of an object needed to be memorized, but aided performance when only a single feature needed to be remembered. Furthermore, there was considerable within-feature interference consistent with some previous findings, but contradicting others. The potential implications for conceptions of VSTM are discussed in the light of two hypothesized stages: an early feature-based stage, as well as a higher-level object-based stage where the depth manipulations exert their effects. The results argue for a strong modulatory influence of surface assignment on object formation for a VSTM task.

1 Introduction

If a visual representation of a stimulus is to be retained after it is no longer visible, some form of visual memory is required. Many lines of evidence suggest that we do indeed have a specialized short-term memory system for retaining visual stimuli for short periods of time, used for example in imagery tasks like mental rotation (Baddeley 1986; Baddeley and Lieberman 1980; Phillips 1974, 1983). This system is distinct from iconic storage (Sperling 1960), in that it survives masking, as well as interference from secondary tasks (Phillips 1974), and in that it leaves a trace for at least several seconds (Phillips and Baddeley 1971; Tanaka and Sagi 1998) while also having severe capacity limitations.

How is information in visual short-term memory (VSTM) organized? If, for example, a red cross is presented and is to be retained in memory, do we remember the color and the form as separate entities, or do we retain the red cross as a unitary object? Intuitively, many would adhere to the latter view. In an influential study, Luck and Vogel (1997) provided evidence for the view that the fundamental units of VSTM are indeed unified objects, even when we only need to retain the stimuli in memory for less than a second (see also Lee and Chun 2001; Vogel et al 2001).

Luck and Vogel presented two arrays of colored bars of a certain orientation for 100 ms each with a 900 ms blank screen interval between the two. On 50% of the trials there was a change in one of the objects from the first array to the next, whereas on the other 50% of trials the two arrays were identical. Observers indicated whether they thought there had been a change in the display between the two arrays or not. Luck and Vogel found that there was no difference in change-detection accuracy when observers had to retain both features of the objects in the array in memory (when there could be a change in either the color and the orientation of the bars) or

---

¶ Current address: Department of Psychology, University of Iceland, Oddi v. Sturlugötu, 107 Reykjavík, Iceland; e-mail: ak@hi.is just one of the features (when there could be a change in only one of the features within each block of trials, the color or the orientation). Thus they argued that sixteen features on 4 objects could be remembered just as well as only four features on the same 4 objects. Supporting this object-based account of VSTM, Lee and Chun (2001) showed that the number of spatial locations to be monitored in a change-detection task did not affect VSTM performance, whereas the number of objects clearly did. Luck and Vogel (1997; see also Vogel et al 2001) also presented evidence indicating that objects containing two different colors could be retained just as well as objects containing one color. Thus there was seemingly no within-feature nor between-features interference for objects in VSTM.

This result accorded rather well with the popular view that visual attention operates on unified objects rather than the individual features in each case. Irwin (1992; see also Irwin and Andrews 1996) reached a similar conclusion regarding the nature of memory for visual stimuli over saccadic eye movements. Such object-based, 'transsaccadic' memory has, according to Irwin, a capacity of about three to four items, a number similar to that suggested as the capacity of VSTM by Luck and Vogel (1997). Furthermore, in Allport (1971), participants were asked to report the color, the shape, or both features of three items presented on a screen. Reports of both color and shape were essentially as good as reports of color or shape alone which is broadly consistent with the results of Luck and Vogel (see also Wing and Allport 1972).

If objects are indeed the fundamental units of VSTM, this opens up the question of how these objects are formed. How does the visual system decide what shall be considered an object? Is object formation prioritized in visual processing? How early in the perceptual process does object formation take place? A popular view is that objects are bound together through the operation of visual attention (Irwin and Andrews 1996; Kahneman et al 1992; Treisman and Gelade 1980; Wolfe et al 1989). Under this scheme visual attention binds individual features together.

Wheeler and Treisman (2002; see also Magnussen et al 1991, 1996, for a related view) have argued for a somewhat different view of VSTM than the one proposed by Luck and Vogel (1997), incorporating the proposed role of visual attention in object formation. They argued that objects are not by necessity the fundamental units of VSTM. Their claim is that object formation is an effortful process and visual attention is needed to bind objects together (from elementary features) and visual attention is also needed to maintain those representations over time. Under this scheme, object-based processing is beneficial, but comes at a cost and requires effort. This view also entails that, should attentional resources be engaged in a different task, bound objects should fall apart into their constituent features at the same time (see eg Woodman and Luck 2004). Wheeler and Treisman also claimed that features from different dimensions can be stored in parallel, but there is a finite capacity for any given feature dimension. This view leads to the prediction that there should be considerably more within-feature interference than between-features interference on VSTM tasks (see also Xu 2002a, 2002b), contrary to the claims of Luck and Vogel (1997). So, for example, observers should find it harder to remember two colors within the same object than if the two features to be remembered came from different dimensions (such as shape and color).

This was indeed what the experiments of Wheeler and Treisman (2002) revealed, but note that this result is in direct contrast with Luck and Vogel's results described above, where they found no cost difference in having to retain an object that had two colors in memory or an object with two features from different dimensions. Part of the research described in this paper has some bearing on this issue (experiment 3, in particular). Olsson and Jiang (2002) have also argued for a 'weak' form of the object hypothesis, arguing that within-feature capacity is much less for a given object than the between-features capacity, but also arguing that object formation aids VSTM performance, other things being equal, which is indeed consistent with the proposal of Wheeler and Treisman (2002; see also Xu 2002b).

A third alternative formulation of VSTM was proposed by Alvarez and Cavanagh (2004). They presented evidence that retention in VSTM can most harmoniously be accounted for by appealing to limits in the total amount of information that can be stored, not only by the number of objects. They found that the capacity of VSTM could be anything from less than 2 objects to 4 to 5 objects, dictated by the nature of the stimuli in each case, which argues against the popular idea that VSTM has a fixed capacity of around 4 items (see, eg, Vogel et al 2001), although they did acknowledge that there seemed to be an upper limit of VSTM capacity of 4 to 5 items as Irwin (1992), and Luck and Vogel (1997) had proposed, but also that this capacity could be lower for relatively complex objects.

1.1 *A stage of surface analysis*

Nakayama and colleagues (1995; see also Driver and Baylis 1995), have suggested that surface analysis is a critical aspect of the perceptual process. Surface assignment can, for example, have dramatic effects on visual-search performance (He and Nakayama 1992; see also He and Nakayama 1995), which suggests that attention operates at a level following surface assignment, a stage perhaps akin to what Marr (1982) called the $2\frac{1}{2}$ dimensional sketch.(1) Furthermore, surface assignment can determine correspondence strength between different features of an ambiguous apparent-motion display (He and Nakayama 1994). There is also persuasive evidence, from studies of neurological patients with attentional deficits, that surface assignment is intact, at least to a considerable extent, in such patients (Driver et al 1992; Mattingley et al 1997; see Driver and Vuilleumier 2001 for a thorough discussion), again suggesting that surface analysis precedes the operation of attention.

If attention binds features to assemble objects, the object-formation process should follow surface assignment, since attentional effects appear mostly to follow surface analysis as was mentioned above (Driver and Baylis 1995; Driver and Vuilleumier 2001; He and Nakayama 1992; Nakayama et al 1995). Related to this, Xu (2002c; see also Xu and Nakayama 2003) reported results testing VSTM for objects on different surface planes induced by stereoscopic disparity. Xu found that when observers had to pay attention to only one depth plane at a time, performance was better than when they had to attend to both surfaces, even though the number and type of objects in the display were the same in each case. This result suggests that surface assignment can have a critical effect on performance in a VSTM task. The aim in the present paper is to address another aspect of surface assignment—the process of *object formation*.

The three experiments presented here were designed to investigate how the object-formation process is influenced by apparent depth differences. The main question that I wanted to answer is how the object-formation process is influenced by induced depth differences between parts of an object (see eg Hoffman and Singh 1997) when a VSTM task is used. The results of Xu (2002a, 2002b) have shown that there are indeed important constraints on what can count as an object for VSTM; features are better integrated in VSTM if they belong to the same parts of objects. The question here is whether depth differences between parts of objects can have similar crucial effects for VSTM tasks.

Note also that the current study presents, if not a methodological improvement, at least an alternative to the common change-detection technique used in most of the

---

(1) This conception does, of course, not deny that attentional effects upon early visual areas can occur, but simply that, functionally, many higher level effects such as visual pop-out and apparent motion seem to occur after surface assignment has taken place (see eg Nakayama et al 1995).

studies cited above (Alvarez and Cavanagh 2004; Lee and Chun 2001; Luck and Vogel 1997; Olson and Jiang 2002; Vogel et al 2001; Wheeler and Treisman 2002; Xu 2002a, 2002b), in that here observers are presented with 4 potential(2) objects that contain 2 features and a post-cue indicates which of them to report on. The post-cue coincides exactly in time with the presentation of a mask. The recollection of those objects is compared across two conditions: where only one of the features of an object must be retained and when both features must be retained. In other words, this is a very direct test of VSTM performance, perhaps more direct than the change-detection task used in most of the current literature, since it is unproven whether the best strategy to perform well in those tasks is to retain a visual representation of the stimuli in memory or simply to detect changes between two views of a particular scene.(3)

2 Experiment 1: Color and position

The main task of the observers in experiment 1 was to remember either single features (the *position* of a gray disk relative to a square or the *color* of that square) of potential objects (see footnote 2) that were presented for 300 ms or to remember *both* features of those objects, in different blocks of trials. The stimuli were immediately followed by a mask, and a post-cue that indicated which of the 4 items the observers should report on (see figure 1a). Judging by the results of Vogel et al (in press), 300 ms is a sufficiently long time for consolidation in VSTM. There were three different depth conditions. In the 'no-interposition/zero-disparity' condition, the potential objects appeared unified with no depth induced between the two parts of the object (see figure 1b). In the 'interposition/zero-disparity' condition there was zero disparity between the two parts of the objects (the square and the disk), but one was placed in front of the other, so the objects were separated only by a pictorial depth cue (interposition, see figure 1b). In the 'interposition and disparity' condition binocular disparity was manipulated so that the two parts of the potential object were separated in depth such that the square seemed to hover in depth in front of the disk (see figure 1b).

There were three different memory conditions: report *color only* within a block; report *position only* within a block; and report either *color or position* within the same block of trials (in this case observers did not know whether to report the color or position until the post-cue appeared along with the mask, after the presentation of the objects to be remembered). In the condition where both features were to be remembered, observers were strongly encouraged to try to remember the two parts of each object as unified objects. It is for this reason that the two-part stimuli are referred to as 'potential objects'; observers were encouraged to attempt to construct objects of the two parts (see footnote 2).

2.1 *Method*

2.1.1 *Stimuli and procedure.* The stimuli were presented on a 75 Hz CRT screen controlled by a G3 Macintosh computer. The VisionShell programming library for the C programming language was used for stimulus generation (http://www.visionshell.com). Two versions of the 4 potential objects to be remembered were presented on the screen, but they were then optically stereoscopically fused with a Wheatstone-type stereoscope consisting of a mirror prism and a mirror (one set for each eye) that projected one fused version of the stimuli to each eye, thus inducing the binocular disparity. Because

---

(2) The term 'potential' object is used here since observers were asked to try to construct coherent objects out of the two parts that were presented. It is thus not assumed beforehand that observers were actually successful in forming objects out of the parts. From here on in when 'objects' are referred to in the context of the present experiments it is, in this sense, as potential objects.

(3) Note that the current study is by no means the only study to study VSTM with the use of a post-cue paradigm; my point is simply that in the *recent* literature the change-detection paradigm has almost exclusively been used.

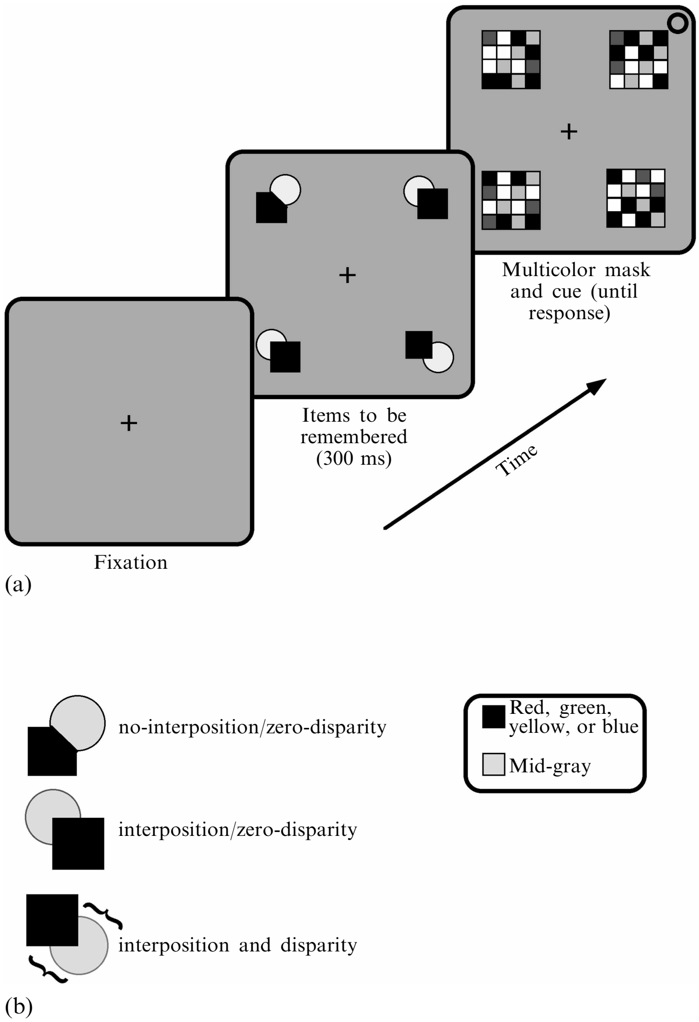

Figure 1. The design of the experiments in this paper. (a) Sequence of events in the experiments. Each trial started with the presentation of a fixation cross at the center of the screen, followed by the presentation of the four stimuli to be remembered (for 300 ms), followed by the masking stimuli and the post-cue indicating which of the four items to report on, and what feature of that item. (b) The three types of objects used in experiment 1 (see sections 2.1.1 and 3.1 for details on the type of stimuli used in those experiments). Note that the stimuli are not drawn to the scale used in the experiments (see the section listed above for details).

of this, observers effectively saw only 4 objects even though 8 objects were actually present on the screen. Each of the 4 objects could be one of the three types shown in figure 1b, and there was always at least one stimulus from each of the three depth conditions on each trial. In the 'interposition and disparity' condition the binocular disparity between each eye's view of the square in front was 18.8 min of arc, which made the square appear to 'float' in front of the disk. The viewing distance was 40 cm so the length on one side of the square was 2 deg and the radius of the disk was 1 deg.

There were three report conditions: observers were asked to remember the color of the square only, the position of the gray disk (top-left, top-right, bottom-left, or bottom-right relative to the square) only, or were required to remember both features of each object, and were encouraged to try to retain these features in memory as objects.

The observers were always told beforehand whether a 'both-features' or a 'single-feature' block of trials was coming up. A post-cue indicated which of the 4 objects to report on, and which feature. If the post-cue was a pink square, observers were to report the color of the square, while a gray circle indicated that the position of the gray disk was to be reported. On 'single-feature' blocks the cues always prompted a response on only one of the features, while on 'both-features' blocks it was determined randomly from trial to trial which feature observers were to report on. So on the single-feature trials observers only had to be concerned with reporting one of the features, and could ignore the other, while on 'both-features' blocks observers had to remember both features. The observers responded with a key-press on the numeric keypad on the right of a standard computer keyboard. If the observers had to report on the position of the circle, the appropriate response keys were 1 (disk at bottom-left), 3 (disk at bottom-right), 7 (disk at top-left), and 9 (disk at top-right). If they had to report the color of the square, the appropriate keys were 2 (green), 4 (yellow), 6 (blue), and 8 (red). Appropriately colored tape was placed on the response keys to distinguish the correct color response keys.

The squares were of one of four colors: yellow (124.2 cd m^{-2}; CIE coordinates: $x = 0.416$, $y = 0.495$), red (43.6 cd m^{-2}; CIE coordinates: $x = 0.604$, $y = 0.338$), green (52.8 cd m^{-2}; CIE coordinates: $x = 0.286$, $y = 0.592$), or blue (25.8 cd m^{-2}; CIE coordinates: $x = 0.147$, $y = 0.112$). The disks were always gray (67.4 cd m^{-2}; CIE coordinates: $x = 0.306$, $y = 0.334$). The stimuli were presented on a black (0.5 cd m^{-2}) background. A multicolor mask immediately followed the 300 ms presentation of the stimuli. The mask consisted of four squares of random dots (dot size = 12 min of arc) presented at the locations of the objects, covering the whole area where each object had appeared (see figure 1a). The colors of the dots were randomly determined for each dot, and the colors were picked from the same set of colors as those of the squares of the objects. The mask and post-cue stayed visible until the observer had responded with a key-press. A sound signal indicated whether the response was correct or not. The viewing distance was 40 cm. On any given trial, the position of the display items around the circle was random, with the constraint that each of the objects was 90 deg away from the others so that all were equidistant from each other. A light-gray fixation cross was present throughout, which helped ensure that proper binocular fusion was maintained and no extra time was needed for fusion before the next trial, after the stimuli were presented.

2.1.2 *Observers.* The six naive observers in the first experiment all had normal color vision and normal or corrected-to-normal visual acuity. They also underwent a test for stereoscopic acuity and all had maximum stereoacuity as measured by the test. Each of the observers participated in 800 trials; 2 blocks of 100 trials where only the position of the disk needed to be remembered; 2 blocks of 100 trials where only the color of the square had to be remembered; and 4 blocks of 100 trials where both the color of the square and the position of the disk needed to be remembered. The 8 blocks were run in counterbalanced order between the subjects. The observers also underwent 100 to 200 practice trials until they were confident how to perform the task, and about the correct key-response mapping.

2.2 *Results and discussion*

The average percentage of correct scores from experiment 1 is presented in figure 2. The first thing to note is that there is a large detrimental effect of depth differences when both features must be remembered for both the position and the color task (the two sets of three bars on the right). Furthermore, interposition and disparity seem to affect performance separately, since performance is worst with both interposition and disparity. It is, however, also possible that the disparity-induced depth accounts for all the differences between the 'no-interposition/zero-disparity' condition and the 'interposition

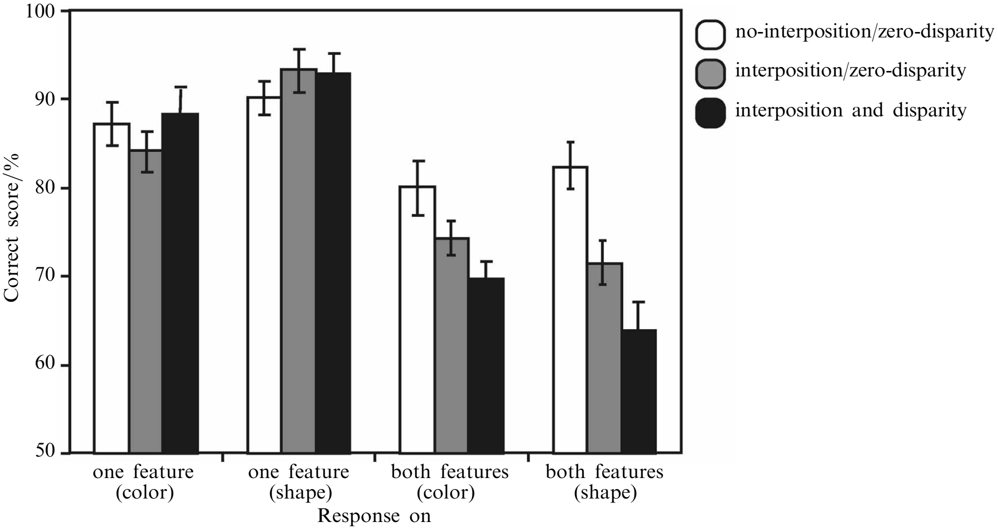

Figure 2. The results of experiment 1: the mean percent-correct scores for the six observers for each condition of the experiment—whether observers had to remember only one, or both features of each object, color of square or position of disk, and what the depth/disparity condition was on each trial (the different brightness levels of the bars). The error bars show the standard error of the mean.

and disparity' condition. No attempt is made here, however, to unconfound the two. When only one feature must be remembered, however, the depth differences seem not to have much effect on performance at all (the two sets of three bars on the left). There is also a large effect of whether one or two features of an object must be remembered, and note that the difference between the 'respond both features' and 'respond one feature' conditions cannot be accounted for by the effects of interposition or disparity, since performance is worse in the 'no-interposition/zero-disparity' condition when both features must be retained than when only one feature must be retained.

A 2 (report on both color of square and position of disk or only one or the other within a block) by 3 (the three depth conditions) repeated-measures ANOVA (thus with individual differences used as the error term) revealed a significant main effect of report condition ($F_{1,5} = 246.10$, $p < 0.001$) as well as a significant main effect of depth condition ($F_{2,4} = 11.90$, $p = 0.021$). The interaction between the two factors was also significant ($F_{2,4} = 12.16$, $p = 0.02$), showing that the effect of depth condition was different according to whether the observers had to report on only one of the features (position or color) or both of them. A posteriori t-tests (with Bonferroni-corrected critical α-levels) between the depth conditions, when both features were to be retained, revealed a significant difference between the 'no-interposition/zero-disparity' condition and the 'interposition/zero-disparity' condition ($t_5 = 4.05$, $p = 0.01$); and a significant difference between the 'interposition/zero-disparity' and 'interposition and disparity' conditions ($t_5 = 7.07$, $p < 0.001$) as well as a significant difference between the 'no-interposition/zero-disparity' and 'interposition and disparity' conditions ($t_5 = 5.95$, $p = 0.002$).

From these results it seems clear that depth differences, whether disparity-induced or induced with pictorial depth cues interfere with the object-formation process. A likely reason for this is that a stage of surface assignment, which is sensitive to the depth cues, precedes the processing level at which object formation can take place, which means that the induced depth prevents successful object formation.

3 Experiment 2: Color by color

Before proceeding further, a potentially important caveat must be acknowledged in relation to the results of experiment 1. It is quite possible that response confusion could have had an effect on the results (see for example Magnussen et al 1996). In the previous experiment, when only one feature was to be remembered, only one type of response was required within each block of trials—position or color. When both features were to be remembered, two types of responses were required within a block, and observers did not know from one trial to the next what type of response would be required on the next trial. Such differences between conditions can, of course, lead to considerable confusion between the two different types of responses (but note that an explanation along those lines would not explain why there is a difference between performance on the three different depth conditions on the two feature-type trials since the potential response confusion is equivalent across the three depth conditions). Nevertheless, to find out the potential effects of response confusion in this paradigm, a similar experiment to experiment 1, albeit where issues of response confusion could be avoided, was conducted.

3.1 *Methods*

In the second experiment, only one type of response (a 4-way color discrimination) was required. The objects used in this experiment were similar to those used in experiment 1, except that the disk in each case was also of one of the four colors used in experiment 1. The only response type that was required, then, was color. The post-cue indicated whether the observers were supposed to report the color of the disk or of the square. If the observers were to report the color of the square, the post-cue was a gray square, while the cue was a gray circle if the color of the disk was to be reported. Two conditions were compared, one where observers only needed to remember the color of either the square or the disk, and one in which the observers needed to remember both colors. Each of the six observers participated in 400 trials; 100 where only the color of the disk needed to be remembered; 100 where only the color of the square had to be remembered; and 2 blocks of 100 trials where the color of both parts of the object needed to be remembered. The 4 blocks were run in counterbalanced order between the subjects. In all other respects, methods were similar to those in the preceding experiment.

3.2 *Results and discussion*

Figure 3 shows the results of experiment 2, which were overall similar to those of experiment 1. When the colors of both parts of the potential object had to be remembered, induced depth disrupted performance to a similar degree as in experiment 1. On the other hand, when only the color of either the disk or the square had to be remembered within a given block there were no differences in performance between the three depth conditions. The main conclusions of experiment 1 are therefore still valid, even when effects of response confusion have been ruled out: depth differences between the two parts of the objects used here interfere with the retention of the colors of those parts, strongly suggesting that object formation can only take place after a stage responsive to the depth cues tested here, a likely candidate being a stage of surface analysis (Driver and Baylis 1995; Marr 1982; Nakayama et al 1995). As in experiment 1, interposition and disparity contributed separately to the observed results, since performance is worst when both are present in the stimulus (although it must be admitted that no attempt was made to unconfound the two).

A 2 (report on color of both square and disk or only one or the other within a block) by 3 (the three depth conditions) repeated-measures ANOVA revealed a significant main effect of depth condition ($F_{2,4} = 11.9$, $p = 0.014$) as well as a significant main effect of report condition ($F_{1,5} = 2823.8$, $p < 0.001$). The interaction between the two

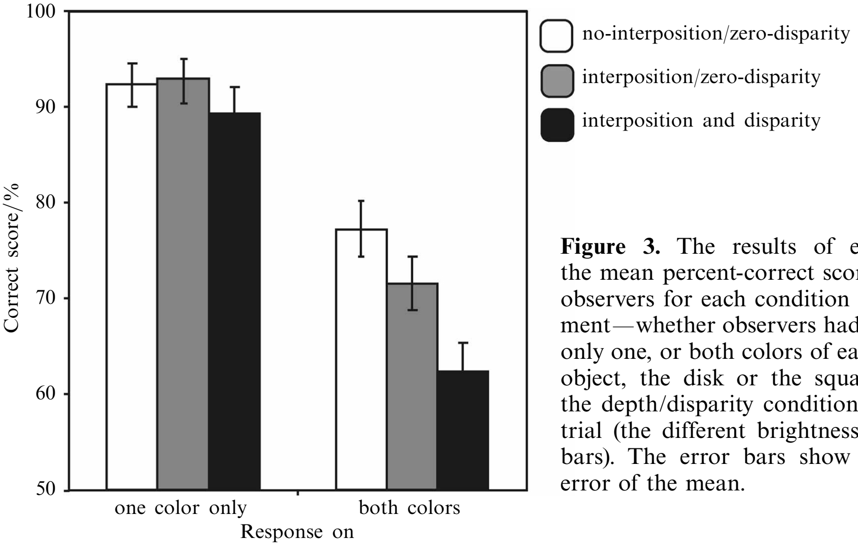

Figure 3. The results of experiment 2: the mean percent-correct scores for the six observers for each condition of the experiment—whether observers had to remember only one, or both colors of each part of the object, the disk or the square, and what the depth/disparity condition was on each trial (the different brightness levels of the bars). The error bars show the standard error of the mean.

factors was also significant ($F_{2,4} = 8.17$, $p = 0.039$), showing that the effect of depth condition differed depending on whether the observers had to report on only one of the colors (square or disk) or both of them. As in experiment 1, a posteriori t-tests (again with Bonferroni-corrected critical α-levels) between the three depth conditions, when both features were to be retained, revealed a significant difference between the 'no-interposition/zero-disparity' condition and the 'interposition/zero-disparity' condition ($t_5 = 3.06$, $p = 0.028$); and a significant difference between the 'interposition/zero-disparity' and 'interposition and disparity' conditions ($t_5 = 3.959$, $p = 0.011$), as well as a significant difference between the 'no-interposition/zero-disparity' and 'interposition and disparity' conditions ($t_5 = 5.77$, $p = 0.002$).

The results of experiment 2 have, potentially, some further interesting implications. As described in the introduction, Luck and Vogel (1997; see also Vogel et al 2001) reported that color-by-color conjunctions on the same object were just as easy to retain in VSTM as conjunctions from different dimensions, as, for example, shape and color. This result accorded well with their conjecture that objects rather than features were the fundamental units of VSTM, since there seemed to be no within-feature interference in their change-detection task. This result has not escaped controversy, however (see eg Olsson and Jiang 2002; Wheeler and Treisman 2002; Xu 2002a; but see Vogel et al 2001 for some responses to these criticisms). The results of experiment 2 here are relevant in that, contrary to Luck and Vogel's result, remembering two colors was a lot more difficult than remembering only one, even when there was no apparent depth difference between the different parts of the objects, in seeming contrast to Luck and Vogel's result.(4)

It is important to note, however, that the stimuli used in experiment 2 were quite different from the ones used by Luck and Vogel to establish this potentially very important result, since they used two colored squares, a large one and a small one, where the whole extent of the small one was embedded within the confines of the large one. Luck and Vogel (1997) showed that the relations between the two parts of the object can have important consequences for VSTM performance. Furthermore, the results of Xu (2002a, 2002b) revealed important constraints on what can count as an object for VSTM. Any arbitrary relation between two parts does not seem to be sufficient for

---

(4) Note, however, that the post-cue procedure used here is, of course, in many ways quite dissimilar to the change-detection procedure used in their study.

successful object formation. To test whether the fact that one part of the object was not embedded within the other part of the object would explain the discrepancy between the results of experiment 2 and the findings of Luck and Vogel (with respect to the within-color interference), performance was tested with stimuli similar to the ones used by Luck and Vogel (1997) where each object consisted of two squares; a small square completely located within the confines of a second larger one. Two depth conditions were compared: the 'disparity' condition where the smaller square appeared to hover in front of the larger square and the 'zero-disparity' condition where there was zero disparity between the two squares. One added benefit of trying this new version of the stimuli was that this could help generalize the findings regarding the depth manipulations of the current paper to new circumstances.

4 Experiment 3: Stimuli conducive to successful object formation

In experiment 3, performance was tested on a VSTM task with objects consisting of color-by-color conjunctions similar to those used by Luck and Vogel (1997). There were three main reasons for this:

(i) To find out whether a paradigm more similar to Luck and Vogel's would result in a similar pattern to the one that was observed in experiment 2, where color-by-color conjunctions were very hard to retain in VSTM, even in a 'no-interposition/zero-disparity' condition. Wolfe et al (1994; see also Wolfe et al 1990) showed that the degree to which one part of an object is embedded within another part of the same object can affect search times in a visual-search task, such that the search was easiest when one part of the object was within the confines of the other part. It is therefore quite possible that the stimuli used by Luck and Vogel (a small square within the confines of a larger one) will be more conducive to object-based processing for VSTM than the ones used in experiment 2 in this paper. This might then explain why considerable within-feature interference was observed in experiment 2. There, color-by-color conjunctions, when observers had to remember both colors, had a devastating effect on performance, for the most part irrespective of depth condition, although performance was even worse when apparent-depth differences were induced. This detrimental effect of having to remember both parts of an object was more pronounced in experiment 2 than in experiment 1, where the two features were position and color, even though it is quite possible that response confusion played a large role in the detrimental effect of having to remember two features in the first experiment, as explained above.

(ii) To investigate whether objects more conducive to object formation (Luck and Vogel 1997; Wolfe et al 1990, 1994) would not suffer as much from the disparity-induced depth differences as those stimuli that might not be as conducive to object-based encoding (see Xu 2002a, 2002b).

(iii) To attempt to find more evidence for the effects of apparent-depth differences on VSTM, thus obtaining more experimental support for the modulatory influence of surface assignment on VSTM performance.

4.1 *Methods*

In experiment 3 the stimuli were similar to those used by Luck and Vogel (1997; see also Vogel et al 2001; Wheeler and Treisman 2002), where a small square was embedded within a larger one. The two squares comprising each of the 4 stimuli were of different colors on each trial. Unlike in the previous two experiments there were only two depth conditions in this experiment, 'zero-disparity', and 'disparity-induced depth'. Interposition was present in both cases. The length of each side of the larger square was 4.7 deg while each of the sides of the smaller square was 3.35 deg. In the three blocked conditions, observers were supposed to remember the color of the small square only, the large square only, or the color of both squares (where the observers did not know which square was the one whose color was to be reported until the post-cue appeared). The four colors used were the same as those used in the previous experiments. In the 'zero-disparity' condition the large and the small square appeared to be at the same depth (except for that imposed by the pictorial depth cue of interposition) while in the 'disparity' condition the disparity between the two smaller squares was 18.8 min (as in the previous two experiments) so that when the stimuli were fused with the stereoscope, the smaller square seemed to hover in front of the larger square. On each trial there were 2 items from each disparity condition (4 in total). The post-cue was a gray square that was presented so that it was slightly peripheral to the mask (the masks appeared where the 4 items were before, as in the previous experiments; see figure 1); if the cue was a large gray square the observers had to report the color of the larger square, while a small gray square indicated that the color of the small square within the larger one should be reported. In all other respects the methods were similar to what was described for the previous two experiments.

As before, six naive observers participated in the experiment and underwent a test for intact stereovision and performed at ceiling, meaning that their stereoscopic acuity was normal, or higher. Each observer participated in 400 trials in all; 100 trials where they had to remember the color of the small square, 100 where the color of the larger square was to be remembered, and 200 trials where the color of both squares had to be remembered. The run order of the 4 blocks of 100 trials was counterbalanced across subjects.

4.2 *Results and discussion*

There are three issues regarding the results of the third experiment that I wish to highlight (see figure 4). First, the results again show the detrimental effect of disparity-induced depth differences between different parts of an object on VSTM performance seen in the first two experiments in this paper (compare the two black bars in figure 4). Second, in the present paradigm these same apparent depth differences actually improve performance when the color of only one of the parts must be remembered (compare the two leftmost bars in figure 4). The separation in depth seems, then, to result in less between-part confusion than the no-disparity condition. This may mean

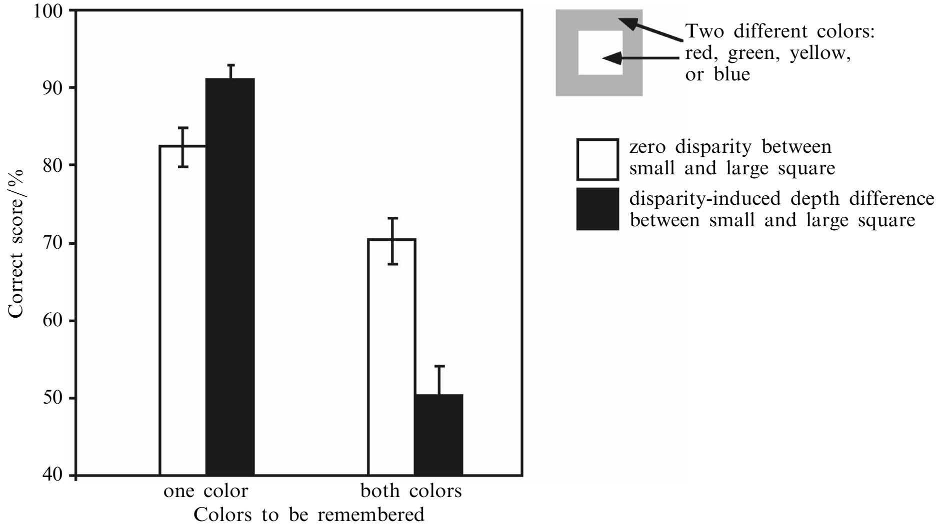

Figure 4. The results of experiment 3: the mean percent-correct scores for the six observers for each condition of the experiment—whether observers had to remember the color of only one of the two squares comprising each object, or the colors of both squares, and what the disparity condition was on each trial (the different brightness levels of the bars). The error bars show the standard error of the mean.

that the object-formation process in this case is at least partly automatic, when circumstances are conducive to that, as they appear to be in the zero-disparity condition, resulting in more within-object confusion about which color to report, even though object formation is not really 'needed' in order to perform the task well. In other words, the disparity separation may have helped observers to ignore the irrelevant square on each trial.(5)

A 2 (report on color of both the small square and the large square, or only one or the other within a block) by 2 ('disparity' or 'zero-disparity') repeated-measures ANOVA revealed a significant main effect of report condition ($F_{1,5} = 85.36$, $p = 0.001$) as well as a significant main effect of disparity condition ($F_{1,5} = 41.9$, $p = 0.001$). The interaction between the two factors was not quite significant ($F_{2,4} = 4.57$, $p = 0.085$) although the trend was clearly in that direction. In fact, an a posteriori t-test (again with Bonferroni-corrected critical α-levels) between the two depth conditions for the single-color report condition only, showed a quite significant difference between the two ($t_5 = 3.47$, $p = 0.018$) supporting the claim above, that the disparity-induced depth differences aided performance when observers only had to remember the color of one of the two squares. A similar a posteriori t-test between the two disparity conditions for the 'report color of both' condition also revealed a significant main effect of depth condition ($t_5 = 4.41$, $p = 0.007$) but here, of course, the effect was in the opposite direction in that non-zero disparity harmed performance.

The third interesting point regarding the results from experiment 3 is that they are seemingly inconsistent with the results of Luck and Vogel (1997; see also Vogel et al 2001), where they found no cost difference of using the same feature (color) on their change-detection task as compared to when the two features to be reported were from different dimensions.(6) According to our results, there was a large additional cost of having to remember the two colors as compared to when only one color was to be remembered. This was the case both when the two objects were in the same depth plane as when they were separated in depth by disparity. The condition where there were no depth differences is directly comparable to the stimuli that Luck and Vogel used, except that in their experiment they used a change-detection paradigm, rather than asking observers directly about what they had seen as is done here. It is, of course, quite possible that the discrepancy between the results can be explained by this difference in the two procedures, but such an explanation would also have to explain in what way the present paradigm does not test VSTM performance. The results are, on the other hand, in line with what has been reported by Olsson and Jiang (2002), Wheeler and Treisman (2002), and Xu (2002a) who in all cases used the change-detection procedure favored by Luck and Vogel (1997). Here this result is observed in a different experimental paradigm from the change-detection paradigm used in those studies, which further generalizes the finding of within-feature interference for VSTM. As mentioned before, it is by no means certain that retaining objects in memory is the best strategy for change detection rather than just monitoring changes in each case, while it is clearly a requirement in the paradigm used here, so it is possible that the current results constitute even stronger evidence for within-feature interference on a VSTM task than previously obtained with the change-detection paradigm.

The results also show that the within-feature interference found in experiment 2 is not explained by a putative difficulty observers had with forming objects out of those stimuli, since this same interference is clearly present in the current experiment.

(5) Note that observers did about as well on the task for the small and large squares so the data were pooled over the two sizes.

(6) Note that the experimental procedure used here involving a post-cue and mask is different from that used by Luck and Vogel (1997) so this should not be considered a non-replication of their results. Nevertheless, their argument would certainly have predicted a different result than was observed.

This result is problematic for a strong version of the 'objects as units for VSTM' hypothesis, which states that there should be no within-feature interference as long as the things to be remembered are within the confines of the same object (see Olsson and Jiang 2002 for a detailed discussion of this issue). It should, of course, be noted that considerable controversy has arisen over the Luck and Vogel claim of no within-feature interference, and some researchers have reported results that directly contradict the original results of Luck and Vogel (Olsson and Jiang 2002; Wheeler and Treisman 2002; Xu 2002a; but see also Vogel et al 2001), as has been discussed at length previously in this article.

5 General discussion

To summarize, the results in this paper show, for the first time, that there are large detrimental effects of apparent-depth differences on VSTM when both features of the objects must be remembered. Experiment 3 also revealed that these same depth differences can actually improve VSTM when only one feature must be remembered. It seems, then, that the depth differences interfere with object formation, resulting in much worse performance when both features of an object must be remembered. It also seems that, when two depth cues are combined (disparity and interposition), this detrimental effect is more pronounced than when only interposition is present. In other words, the two manipulations may contribute separately to the observed pattern of results. It is, of course, possible that disparity accounts for all the effects, when it is used, but the present data cannot address this question, since no attempt was made to 'unconfound' disparity and interposition in the experimental design used here.

The fact that performance on the single-feature condition in experiment 3 was better under the disparity condition than the no-disparity condition suggests that the object-formation process is in some sense automatic, since the two squares seemingly cannot be 'kept apart', and making objects out of the items to be remembered can actually harm performance when the color of only one of the squares must be retained. This result is in agreement with the object-based accounts of VSTM (Luck and Vogel), but the fact that there was also considerable within-feature interference (for color) is more supportive of the 'weak' object hypothesis of Olsson and Jiang (2002).

This is not the first example in the literature of how the different parts of a figure are less well retained if they are harder to define as a unitary perceptual object. Xu (2002b) has reported results that are consistent with the results here, showing that the benefits of object-based encoding of stimuli for a VSTM task are strongly modulated by the nature of the relations between the different parts of the objects to be remembered. For example, objects that have a point of negative minimum of curvature (see Kristjánsson and Tse 2001) between their parts are retained less well than objects that do not; and the same applies to objects containing parts separated into figure and ground (Xu 2002c; see also Xu 2002a). The present results add to these findings in that they show how inferred depth relations between objects influence how well objects can be grouped and retained in visual memory.

In light of the results of the three experiments presented here, I want to propose that the object-formation process follows a stage of surface assignment, which is in line with the conjecture of Wheeler and Treisman (2002) that the object-formation process is effortful, requiring attention, and that attention is also required to retain such a representation over time. Consistent with this, previous reports have indicated that attentional effects can only occur after a stage of surface assignment (see, for example, Driver and Vuilleumier 2001; Nakayama et al 1995), or what some have termed 'mid-level vision' (Marr 1982; Nakayama et al 1995). A potentially related finding is the finding of Vogel et al (in press), who showed that grouped items are remembered better than ungrouped ones in a VSTM task, demonstrating the importance of mid-level visual processes such as perceptual organization for VSTM performance, as does the study of Xu and Nakayama (2003) who showed that placing objects in different depth planes improves VSTM performance. It is, however, interesting that Xu and Nakayama (2003) reported that the benefits of depth were not due to perceptual grouping, and ruled this out on the basis of experiments where objects in the same depth plane did not benefit, in terms of VSTM performance, from being grouped by similarity; which seemingly contradicts the findings of Vogel et al (in press).

It is also of note that in the current studies brief masked displays with a post-cue were used, whereas most of the previous results depended upon experiments where the primary task was change detection, not retention of the presented stimuli as such. It is not certain that the most effective strategy for performing well in the change-detection task is to try to remember what was presented rather than detecting a change (identifying the change was commonly not required in these tasks). This fact is, perhaps, especially important in light of the current debate on whether there is within-feature interference in object-based VSTM, since the current results generalize the finding of within-feature interference beyond the change-detection paradigm.

5.1 *Two stages of VSTM?*

Wheeler and Treisman (2002) have, as mentioned above, argued that the object-formation process is effortful, requiring attentional resources, and that the maintenance of objects in VSTM is effortful as well. In light of this, it is quite possible that the present experiments probed VSTM performance at two different levels. In one case, VSTM may operate on unified objects. The disparity and interposition manipulations used in the present experiments would then adversely affect this stage. Disparity-induced depth differences make object formation harder. This object-based stage is effortful, requiring attentional resources to bind object representations together and maintain them in memory (Jiang et al 2000; Wheeler and Treisman 2002).

Wheeler and Treisman also argued for the existence of a memory process that does not operate on objects. This is, most likely, an earlier stage of perceptual memory more tied to the operation of relatively early visual areas (Allport 1971; Magnussen and Greenlee 1999; Tanaka and Sagi 1998). Decay of the memory trace in this system is relative slow; probably a lot slower than that of an object-based memory trace, even up to 15 to 30 ms (Magnussen et al 1990; Tanaka and Sagi 1998). This perceptual memory may not require attentional resources (at least not to the same degree as the object-based one), so the memory capacity is mainly limited by the capacity of each processing module. Thus, it has finite within-feature capacity, as suggested by Wheeler and Treisman (2002). Results consistent with this have also been reported by Magnussen and colleagues (Magnussen et al 1991, 1996; Magnussen and Greenlee 1999). They argued that these 'lower-level' memory traces could be related to what Tulving and Schachter called the "perceptual representation system" (Schachter and Tulving 1994; Tulving and Schachter 1990). Importantly, there is good evidence that on some memory tasks when only single features are to be retained, involving simple stimuli such as Gabor patches, observers perform better than they think they do. For example Magnussen et al (1990) presented two Gabor patches at two different times with a variable interval between the two to their observers. The task of the observers was to judge which of the Gabor patches had the higher spatial frequency, which effectively means that they had to retain the first Gabor patch in memory and compare it with the second one. Observers were able to do this quite well even with long delays, such as 10 to 30 s, and interestingly did quite well even when they felt that they were simply guessing. This indicates that a memory trace may remain in early visual areas that aids performance, even though it is not directly available for conscious processing (see Magnussen and Greenlee 1999 for a thorough discussion of the issues involved). Such memory mechanisms may be related to priming effects in vision (Kristjánsson et al 2002; see Kristjánsson, in press, for a review) as well as implicit learning in attention shifts (discussed, for example, in Kristjánsson et al 2001; Kristjánsson and Nakayama 2003; and in Nakayama et al 2004).

Furthermore, Tanaka and Sagi (1998) reported that a near-threshold Gabor patch left a memory trace that facilitated contrast sensitivity for up to 16 s. They hypothesized that the locus of this memory trace was the primary visual cortex, suggesting the existence of a memory mechanism tied to early visual processing at the cortical level. I want to propose here that performance on the tasks in the experiments here is partly mediated by these lower-level memory traces (in this case possibly tied to the operation of color-sensitive areas of the temporal lobe—Bartels and Zeki 2000; Hadjikhani et al 1998), whereas, if performance on the 'both-features' tasks in the current experiments is to be good, objects must be formed to ease retention, but this is strongly interfered with in-depth manipulations that I have used here. Such a conception of VSTM fits well with the finding of Alvarez and Cavanagh (2004) who reported a near-perfect correlation between the total amount of information load measured by visual-search slopes (or, in other words, search rate) of different types of objects and observers VSTM capacity for these items as measured by a change-detection task. This means that memory capacity is variable; depending on the stimulus dimension in each case. But Alvarez and Cavanagh also seemingly found an upper limit of 4 to 5 objects for VSTM independent of information content. In fact, they argued that their results might very well reflect that there are independent memory stores for different features, which is certainly an explanation along the lines that I have been presenting here, if one assumes that visual memory is partly based on activity in early-level visual areas which are modular, with distinct areas operating on different aspects of each visual stimulus.

These conclusions are undeniably speculative. Whether a dichotomous conception of VSTM like the one that I present here is an accurate representation of reality, and is indeed the reason why the present results were obtained, remains to be seen, but seems at present to be the most parsimonious account for the data. Further experimentation may, however, be necessary for more concrete conclusions on this issue.

6 Conclusions

I have presented the results from three experiments that have demonstrated strong modulatory influences of surface assignment with binocular and pictorial depth cues on the process of object formation for visual short-term memory. I propose that performance on visual short-term memory tasks probably reflects the operation of two distinct stages of visual processing: an early level tied to the operation of early visual areas and a later stage that operates on unified objects that are assembled through the operation of visual attention along the lines of the one proposed by Wheeler and Treisman (2002). It is this second, higher-level stage, that is influenced by the depth manipulations between the parts of objects, imposed in the current experiments.

Acknowledgments. The research in this paper was supported by fellowships from the Human Frontiers Science Program, the Fulbright Association, Harvard University, and the American–Scandinavian Society awarded to the author. Most of this research was carried out while the author worked at the Vision Sciences Laboratory at Harvard University.

References

Allport D A, 1971 "Parallel encoding within and between elementary stimulus dimensions" *Perception & Psychophysics* **10** 104–108

Alvarez G, 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 D, 1986 *Working Memory* (Oxford: Oxford University Press)

Baddeley A D, Lieberman K, 1980 "Spatial working memory", in *Attention & Performance, VIII* Ed. R Nickerson (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 521 – 539

Bartels A, Zeki S, 2000 "The architecture of the color centre of the human visual brain: New results and a review" *European Journal of Neuroscience* **12** 172 – 193

Driver J, Baylis G C, 1995 "One-sided edge assignment in vision: II. Part decomposition, shape description, and attention to objects" *Current Directions in Psychological Science* **4** 201 – 206

Driver J, Baylis G C, 1996 "Edge-assignment and figure – ground segmentation in short-term visual matching" *Cognitive Psychology* **31** 248 – 306

Driver J, Baylis G C, Rafal R D, 1992 "Preserved figure – ground segregation and symmetry perception in visual neglect" *Nature* **360** 73 – 75

Driver J, Vuilleumier P, 2001 "Perceptual awareness and its loss in unilateral neglect and extinction" *Cognition* **79** 39 – 88

Hadjikhani N, Liu A K, Dale A M, Cavanagh P, Tootell R B, 1998 "Retinotopy and color sensitivity in human visual cortex area V8" *Nature Neuroscience* **1** 235 – 241

He Z J, Nakayama K, 1992 "Surfaces versus features in visual search" *Nature* **359** 231 – 233

He Z J, Nakayama K, 1994 "Perceived surface shape not features determines correspondence strength in apparent motion" *Vision Research* **34** 2125 – 2135

He Z J, Nakayama K, 1995 "Visual attention to surfaces in 3-D space" *Proceedings of the National Academy of Sciences of the USA* **92** 11155 – 11159

Hoffman D D, Singh M, 1997 "Salience of visual parts" *Cognition* **63** 29 – 78

Irwin D, 1992 "Memory for position and identity across eye movements" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **18** 307 – 317

Irwin D, Andrews R V, 1996 "Integration and accumulation of information across saccadic eye movements", in *Attention and Memory XVI* Eds T Inui, J C McClelland (Cambridge, MA: MIT Press) pp 125 – 155

Jiang Y, Olsson I, Chun M M, 2000 "Organization of visual short term memory" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **26** 683 – 702

Kahneman D, Treisman A, Gibbs B J, 1992 "The reviewing of object files: object-specific integration of information" *Cognitive Psychology* **24** 175 – 219

Kristjánsson Á, 2006, "Rapid learning in attention shifts—A review"' *Visual Cognition* **13** 324 – 362

Kristjánsson Á, Mackeben M, Nakayama K, 2001 "Rapid, object-based learning in the deployment of transient attention" *Perception* **30** 1375 – 1387

Kristjánsson Á, Nakayama K, 2003 "A primitive memory system for the deployment of transient attention" *Perception & Psychophysics* **65** 711 – 724

Kristjánsson Á, Tse P, 2001 "Curvature discontinuities are cues for rapid shape analysis" *Perception & Psychophysics* **63** 390 – 403

Kristjánsson Á, Wang D, Nakayama K, 2002 "The role of priming in conjunctive visual search" *Cognition* **85** 37 – 52

Lee D, Chun M M, 2001 "What are the units of visual short-term memory, objects or spatial locations?" *Perception & Psychophysics* **63** 253 – 257

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

Magnussen S, Greenlee M W, Asplund R, Dyrnes S, 1990 "Perfect short-term memory for periodic patterns" *European Journal of Cognitive Psychology* **2** 245 – 262

Magnussen S, Greenlee M W, Asplund R, Dyrnes S, 1991 "Stimulus-specific mechanisms of visual short-term memory" *Vision Research* **31** 1213 – 1219

Magnussen S, Greenlee M W, Thomas J P, 1996 "Parallel processing in visual short-term memory" *Journal of Experimental Psychology: Human Perception and Performance* **22** 202 – 212

Magnussen S, Greenlee M W, 1999 "The psychophysics of perceptual memory" *Psychological Research* **62** 81 – 92

Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)

Mattingley J B, Davis G, Driver J, 1997 "Preattentive filling-in of visual surfaces in parietal extinction" *Science* **275** 671 – 674

Nakayama K, He Z J, Shimojo S, 1995 "Visual surface representation: A critical link between lower-level and higher-level vision", in *Visual Cognition: An Invitation to Cognitive Science* volume 2, 2nd edition, Eds S M Kosslyn, D N Osherson (Cambridge, MA: MIT Press) pp 1 – 70

Nakayama K, Maljkovic V, Kristjánsson Á, 2004 "Short term memory for the rapid deployment of visual attention", in *The Cognitive Neurosciences, III* Ed. M Gazzaniga (Cambridge, MA: MIT Press) pp 397 – 406

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

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

Phillips W A, 1983 "Short-term visual memory" *Philosophical Transactions of the Royal Society of London B* **302** 295 – 309

Phillips W A, Baddeley A D, 1971 "Reaction time and short-term visual memory" *Psychonomic Science* **22** 73 – 74

Schachter D L, Tulving E, 1994 "What are the memory systems of 1994?", in *Memory Systems 1994* Eds D L Schachter, E Tulving (Cambridge, MA: MIT Press) pp 1 – 38

Sperling G, 1960 "The information available in brief visual presentations" *Psychological Monographs* **74** 498

Tanaka Y, Sagi D, 1998 "A perceptual memory for low-contrast visual signals" *Proceedings of the National Academy of Sciences of the USA* **95** 12729 – 12733

Treisman A, Gelade G, 1980 "A feature-integration theory of attention" *Cognitive Psychology* **11** 97 – 136

Treisman A, 1988 "Features and objects: The Fourteenth Bartlett Memorial Lecture" *Quarterly Journal of Experimental Psychology A* **40** 201 – 237

Tulving E, Schachter D L, 1990 "Priming and human memory systems" *Science* **247** 301 – 306

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

Vogel E K, Woodman G F, Luck S J, in press, "The time course of consolidation in visual working memory" *Journal of Experimental Psychology: Human Perception and Performance*

Wheeler M E, Treisman A M, 2002 "Binding in visual short term memory" *Journal of Experimental Psychology: General* **131** 48 – 64

Wing A, Allport D, 1972 "Multidimensional encoding of visual form" *Perception & Psychophysics* **12** 474 – 476

Wolfe J M, Cave K R, Franzel S L, 1989 "Guided search: An alternative to the feature integration model for visual search" *Journal of Experimental Psychology: Human Perception and Performance* **15** 419 – 433

Wolfe J M, Friedman-Hill S R, Bilsky A B, 1994 "Parallel processing of part – whole information in visual search tasks" *Perception & Psychophysics* **55** 537 – 550

Wolfe J M, Stewart M I, Friedman-Hill S R, Yu K P, Shorter A D, 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

Woodman G F, Luck S J, 2004 "Visual search is slowed when visuospatial working memory is occupied" *Psychonomic Bulletin & Review* **11** 269 – 274

Xu Y, 2002a "Limitations of object-based feature encoding in visual short-term memory" *Journal of Experimental Psychology: Human Perception and Performance* **28** 458 – 468

Xu Y, 2002b "Encoding color and shape from different parts of an object in visual short-term memory" *Perception & Psychophysics* **64** 1260 – 1280

Xu Y, 2002c "The interaction of surfaces and visual short-term memory capacity" *Abstracts of the Psychonomic Society* **7** 471

Xu Y, Nakayama K, 2003 "Placing objects at different depths increases visual short-term memory capacity" *Journal of Vision* **3** 27a http://journalofvision.org/3/9/27/

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

PERCEPTION

VOLUME 35 2006

www.perceptionweb.com