The influence of similarity on visual working memory representations

Po-Han Lin
University of Iowa, Iowa City, IA, USA

Steven J. Luck
University of California, Davis, CA, USA

In verbal memory, similarity between items in memory often leads to interference and impaired memory performance. The present study sought to determine whether analogous interference effects would be observed in visual working memory by varying the similarity of the to-be-remembered objects in a colour change detection task. Instead of leading to interference and impaired performance, increased similarity among the items being held in memory led to improved performance. Moreover, when two similar colours were presented along with one dissimilar colour, memory performance was better for the similar colours than for the dissimilar colour. Similarity produced better performance even when the objects were presented sequentially and even when memory for the first item in the sequence was tested. These findings show that similarity does not lead to interference between representations in visual working memory. Instead, similarity may lead to improved task performance, possibly due to increased stability or precision of the memory representations during maintenance.

Decades of research shows that similarity between representations in memory typically leads to impaired memory performance. In the domain of verbal working memory, lists of phonetically similar words (e.g., “cad, map, man, cap, mad”) are usually more difficult to remember than lists of phonetically dissimilar words (e.g., “pit, day, pen, bar, few”—see Conrad, 1964). In the domain of episodic long-term memory, similar representations may lead to false memories by priming other related representations, as in the Deese/Roediger-McDermott paradigm (Roediger & McDermott, 1995). In addition, proactive interference can impair memory performance when...
materials from the same set are repeated over trials (Keppel & Underwood, 1962). Thus, similarity often engenders poorer memory performance, and it does not usually lead to better performance unless it can somehow be used as a retrieval cue (e.g., knowing that all the words in a list rhyme—see Gupta, Lipinski, & Aktunc, 2005).

In the context of visual working memory, similarity between representations might be expected to impair performance if the representations are competing for access to a limited supply of some representational medium (see, e.g., Alvarez & Cavanagh, 2004). For example, if a limited pool of colour-selective units is used to store colours in working memory, and two slightly different shades of red are being stored, these two representations may compete for access to the set of units, reducing the number of units available for each representation. Similarly, if the representations of the two shades of red are not well isolated from each other, the colour of one representation may blend with the colour of the other representation, distorting both representations. If the representations are segregated, however, then similarity should not influence performance as long as it cannot be used in some kind of strategic manner.

To examine these alternatives, Experiment 1 used a standard colour change detection task (based on the paradigm of Luck & Vogel, 1997) in which the colours on a given trial were either similar to each other or dissimilar from each other (see Figure 1). Surprisingly, performance was more accurate for similar colours than for dissimilar colours. Experiments 2 and 3 ruled out the possibility that this effect reflected the use of relational encoding or any other change in the initial formation of the memory representations. This surprising pattern of results suggests that visual working memory representations are created independently but may interact in maintenance or decision processes.

**EXPERIMENT 1**

This experiment involved a standard change detection task with items of similar colours in one condition and items of dissimilar colours in a second condition. To demonstrate the potency of the similarity manipulation, the

---

1 We are using the term *visual working memory* to refer to the working memory subsystem for storing object identity information; the subsystem for storing spatial information may operate according to different rules.

2 The effects of similarity on verbal working memory are typically found only when ordered recall is required. Because time in the auditory modality is often analogous to space in the visual modality, one might expect similarity to influence visual working memory only when the location of a visual feature must be remembered. The experiments reported in the present study encourage but do not strictly require memory for the location of a feature.
Figure 1. (A) Example stimulus sequences from the heterogeneous and homogeneous change detection tasks of Experiment 1. (B) Example stimulus sequences from the heterogeneous and homogeneous visual search tasks of Experiment 1. (C) Colour coordinates of the three sets of eight colours used in all experiments. The colours were measured with a Tektronix J17 LumaColor chromaticity meter using the 1976 CIE (Commission Internationale d’Eclairage) coordinate system. Luminances ranged between 8.3 and 22.5 cd/m² for the blue set, between 15.6 and 22.7 cd/m² for the red set, and between 54.7 and 61.8 cd/m² for the green set. To view this figure in colour, please see the online issue of the Journal.
same sets of similar and dissimilar colours were also used in a visual search task. Greater similarity between the target and distractors should lead to a steeper search slope.

Method

Participants. Twenty University of Iowa students received course credit for participating. All were between the ages of 18 and 30 and reported normal or corrected-to-normal visual acuity. Half performed the change detection task first and half performed the visual search task first.

Change detection task. The stimuli (see Figure 1A) were presented at a viewing distance of 57 cm on a video monitor with a black background and a continuously visible fixation point. Each trial consisted of a 200 ms sample array followed by an 800 ms delay interval and then a test array that was visible until the subject responded. The sample array consisted of three or four coloured squares (0.65° × 0.65°) distributed randomly across a 9.7° × 7.1° region. The colours of the items were selected from a set of eight shades of red, a set of eight shades of green, and a set of eight shades of blue (colour coordinates are shown in Figure 1C). On homogeneous trials, all items in the sample array were selected at random without replacement from a single set (all red, all green, or all blue). On heterogeneous trials, one colour was selected at random from each set, with a second item from one of the three sets when the sample array contained four items.

The test array was identical to the sample array on 50% of trials, and the colour of one item changed on the remaining 50%. When a colour changed, it changed to an unused value from the same set (e.g., from one shade of red to another). To produce a relatively consistent change magnitude, the new colour was always one of the three most distant members of the same set. Subjects made an unspeeded button-press response at the end of each trial to indicate whether or not a change had been detected.

Set size varied unpredictably from trial to trial, but all the trials within a given block were either homogeneous or heterogeneous. After a brief practice period, subjects performed six blocks of 96 trials, alternating between the homogeneous and heterogeneous conditions (the starting condition was counterbalanced).

To minimize contributions from verbal working memory, subjects performed a concurrent articulatory suppression task that effectively discourages visual-to-verbal recoding (Baddeley, 1986; Besner, Davies, & Daniels, 1981; Murray, 1968). Specifically, three randomly selected digits

---

3 As will be shown in Experiment 2, the same pattern of results is obtained whether colour homogeneity varies unpredictably or remains constant within each block.
were presented at the beginning of each trial block, and participants repeated these numbers aloud at a rate approximately three digits per second throughout the block.

**Visual search task.** The visual search stimuli (see Figure 1B) were identical to those used for the change detection task, except that half of the squares were rotated by 45° to create diamond shapes. The target was defined by its colour, and the target colour varied at random from trial to trial. Each trial began with a 1500 ms presentation of a target cue, which was a 0.65° filled circle presented at the centre of the monitor that indicated the target colour for that trial. The search array was then presented and remained visible until the response. The cue stimulus also remained visible so that subjects did not need to rely on memory for the cued colour. The search array consisted of one, two, or three items, each of which was a square or diamond (randomly determined). One item drawn in the target colour was always present, and subjects made a speeded response on one of two buttons to indicate if the item drawn in this colour was a square or a diamond.

Subjects received one homogeneous trial block and one heterogeneous trial block, in counterbalanced order. In the homogeneous trial block, the colours of the items were drawn at random without replacement from a single colour set. In the heterogeneous trial block, each colour was drawn from a different set. Sixty trials were tested at each set size in each trial block.

**Results and discussion**

Change detection accuracy is shown in Figure 2A. Performance was quantified with several measures, including percentage correct, $d'$, $A'$, and the Cowan-Pashler $K$ (Cowan et al., 2005; Pashler, 1988). All yielded the same pattern, so only percentage correct will be reported here. The data were analysed with a within-subjects analysis of variance (ANOVA) with factors of set size (3 vs. 4) and colour homogeneity (homogeneous vs. heterogeneous).

Accuracy was found to be significantly greater at set size three than at set size four, $F(1, 19) = 10.15, p < .01$. Accuracy was also significantly greater on homogeneous trials than on heterogeneous trials, $F(1, 19) = 15.09, p < .01$. The effect of heterogeneity was somewhat greater at set size 3, leading to a significant interaction, $F(1, 19) = 8.09, p < .05$.

Visual search results are summarized in Figure 2B. Mean RTs were virtually identical for homogeneous and heterogeneous trials at set size 1, for which the stimuli were identical. This suggests that subjects did not use wildly different strategies for the two conditions. However, the slope was
substantially higher when the distractors were from the same set (298 ms/item) than when they were from different sets (29 ms/item), which was a significant difference in a paired t-test, \( t(19) = 5.79, p < .01 \). Error rates were generally low, but increased as set size increased for homogeneous trials. An ANOVA on the error rates yielded a significant interaction between set size and condition, \( F(2, 38) = 5.41, p < .01 \). The direction of the error effects indicates that the response time (RT) results were not a result of
speed-accuracy tradeoffs. Together, the RT and accuracy results indicate that the similarity manipulation was highly potent.

This experiment demonstrates that change detection accuracy is not impaired when the objects being remembered are similar, but is instead improved. We have replicated this pattern several times. The effect is not large, but it is almost always statistically significant. This effect will provide an important constraint on models of visual working memory.

EXPERIMENT 2

Experiment 2 was designed to rule out explanations based on relational encoding during perception or during memory encoding. The simultaneous presentation of similar colours in Experiment 1 may have led to low-level contrast effects that made the colours more perceptually distinct. Alternatively, subjects may have encoded differences between the colours rather than the individual colours, and this may have been more effective when the colours were similar to each other. To rule out these possibilities, we presented the sample items sequentially rather than simultaneously (but at distinct locations, as in Experiment 1). A single test item was presented at the end of the trial at a location that had previously been occupied by one of the sample items, and the subjects were asked to determine whether this item was the same colour as the corresponding sample item. The set size was fixed at three items.

This procedure eliminates the possibility of perceptual contrast effects. Relational encoding in working memory is still possible, because subjects could potentially store the first item in memory and then use this as an anchor point for storing the second and third items in memory. However, this could not impact performance on those trials when the first item happens to be tested. That is, there is no way that the similarity of the second item could influence the encoding of the first item (although it could influence maintenance or decision processes, as discussed later). Two versions of the experiment were conducted, one in which the homogeneous and heterogeneous trials were tested in separate blocks of trials (Experiment 2A) and one in which they were randomly intermixed within blocks (Experiment 2B).

Method

The methods were identical to those of Experiment 1 except as follows. Eighteen new subjects were recruited for Experiment 2A, and twenty were recruited for Experiment 2B. As illustrated in panels A and B of Figure 3, a sequence of three sample items was presented on each trial, each lasting
200 ms and followed by a delay of 800 ms. A single item was presented at test, and its colour was equally likely to be the same as or different from the colour of the corresponding item from the sample sequence. As in Experiment 1, a changed colour was always drawn from the same set as the original colour. The first, second, and third sample items were equally likely to be tested, and this varied unpredictably from trial to trial. The ANOVA for each experiment included factors of homogeneity (homogeneous vs. heterogeneous) and serial position (whether the test item was presented at the location of the first, second, or third sample item).

Results and discussion

The results are shown in Figure 4A (Experiment 2A) and Figure 4B (Experiment 2B). Accuracy was greater at later serial positions, but this effect was significant only for Experiment 2B, $F(1, 19) = 13.10, p < .01$. Accuracy was significantly greater for homogeneous trials than for heterogeneous trials in both versions, $F(1, 17) = 11.77, p < .01$ for Experiment 2A, and $F(1, 19) = 17.53, p < .01$ for Experiment 2B. This effect was comparable in magnitude to the similarity effect observed at set size 3 in Experiment 1. The effect was approximately equal at all three serial positions in Experiment 2A, for which the interaction between serial position and homogeneity did not approach significance ($F < 1$). The effect of homogeneity declined for the third item in the sequence in Experiment 2B, leading
to a significant interaction, $F(2, 38) = 3.24, p < .05$, perhaps because observers could determine whether the trial was homogeneous or heterogeneous after the second item in the sequence. Planned comparisons indicated that accuracy was significantly higher for homogeneous trials
than for heterogeneous trials at the first serial position in both Experiment 2A, \( t(17) = 2.26, p < .05 \), and Experiment 2B, \( t(19) = 3.85, p < .01 \).

The finding of a homogeneity effect with sample items that were separated by a gap of 800 ms clearly rules out any effect of homogeneity on low-level sensory processing. The finding of a homogeneity effect for the first item in the sequence further rules out the possibility that homogeneity influenced the initial encoding of the colour in working memory. The finding of similar effects for blocked and mixed versions of the experiment rules out differences in strategy at the time of the encoding of the first item.

**EXPERIMENT 3**

Experiment 3 was designed to push the boundaries of the similarity effect by testing whether all items being maintained in working memory must be similar to each other to produce a benefit of similarity. That is, if two items are similar to each other and one item is different, is memory performance better when one of the two similar items is tested than when the dissimilar item is tested? Or does the effect of similarity require that all items in memory are either similar or dissimilar? This experiment also allowed us to test whether accuracy for the first item in the sequence would be more accurate when the second item in the sequence was similar (and hence the third item was dissimilar) compared to when the second item in the sequence was dissimilar (and the third item was similar). This comparison provides information about the temporal dynamics of the similarity effect.

**Method**

The methods were identical to those of Experiment 2 except that two of the three items in a sequence were selected from one colour set and the other item was selected from a different colour set (see Figure 3C). The sequential ordering of the two similar colours and the one dissimilar colour was random. There were not separate homogeneous and heterogeneous trials. A new group of 20 subjects participated. The ANOVA variables were serial position of the tested item (first, second, or third) and similarity (whether or not the tested item was similar to one of the other two items in the sequence).

**Results and discussion**

The results are shown in Figure 4C. Accuracy was significantly greater for later serial positions, \( F(2, 38) = 27.13, p < .01 \). Accuracy was also
significantly greater when one of the two similar items was tested than when the dissimilar item was tested, $F(1, 19) = 5.00, p < .05$. The effect of similarity was constant across serial positions, and the interaction did not approach significance ($F < 1$). A planned comparison demonstrated that accuracy was significantly greater for the similar items than for the dissimilar item at the first serial position, $t(19) = 2.33, p < .05$.

These results demonstrate that similarity can improve memory performance even when only a subset of the items in working memory are similar to each other, and even when the first item in the sequence is tested. This clearly rules out any explanations based on the nature of the initial memory encoding process. In addition, these results indicate that similarity can improve performance even when only a subset of the items in memory are similar to each other, further constraining the range of possible explanations. More specifically, any mechanisms that depend on the overall level of homogeneity amongst the items being held in memory could not operate in the present experiment because all trials contained two similar items and one dissimilar item. For example, one could imagine that the overall degree of lateral inhibition between representations might be adjusted in a top-down manner as a function of the overall degree of similarity, but such a mechanism could not explain the results of the present experiment. In addition, it was very clear to observers when all the items were similar or when all of the items were dissimilar in Experiments 1 and 2, and this may have facilitated the use different decision rules for these two types of trials. However, observers in the present experiment could not know whether the dissimilar item or one of the two similar items would be tested until the test stimulus appeared, at which time it would probably be too late to use a different decision rule.

In additional follow-up analyses, we found that accuracy for the first item in the sequence was significantly greater when the second item was similar (and the third item was therefore dissimilar) compared to when the second item was dissimilar (and the third item was therefore similar), with mean accuracy levels of 75.6% correct and 70.9% correct, respectively ($p < .05$). However, accuracy for the third item in the sequence was approximately the same when the second item was similar (and the first item was dissimilar) compared to when the second item was dissimilar (and the first item was similar), with mean accuracy levels of 81.7% correct and 81.9% correct, respectively ($p < .05$). Thus, the benefit of similarity on the first item in the sequence was essentially erased when a dissimilar item was presented before a similar item was presented, whereas temporal proximity did not matter when memory for the third item in the sequence was tested.

How might this pattern be explained? When a representation of the first item in the sequence is created, there are no other memory representations that might influence the precision or stability of this representation. When
the next item is presented, it can either help or harm the representation of
the first item depending on whether it is similar or dissimilar. At this point,
we do not know if the presentation of a dissimilar second item causes the
representation of the first item to become less accurate and/or stable than it
was before the presentation of the second item, or whether the presentation
of a similar second item causes the representation of the first item to become
more accurate and/or stable than it was before the presentation of the second
item. However, once the first item has been stored in memory, it can be used
to enhance the representation of subsequent items, even when dissimilar
items intervene between the similar items. Consequently, accuracy for the
third item does not depend on whether it the similar item was the first or
second item in the sequence (as long as one of these items was similar to it).

GENERAL DISCUSSION

This study has demonstrated a small but consistent improvement in
performance when the items being maintained in memory are similar to
each other compared to when they are dissimilar. Although this small
improvement is interesting, the most important finding is that there was no
decrement in performance when the items were similar to each other. This
finding falsifies the very plausible hypothesis that visual working memory
representations of similar feature values would be less accurate than
representations of dissimilar feature values. Consequently, any theory of
visual working memory cannot be correct if it implies that performance
should be impaired by similarity among the items being held in memory. It
should be noted, however, that the present results are limited to memory for
colour, and other principles may operate for other features (particularly for
features such as location and direction of motion, which do not appear to be
represented in the same brain systems used to represent surface features such
as colour—see Courtney, Ungerleider, Keil, & Haxby, 1996; Jonides et al.,
1993). Moreover, the present finding of a benefit for similar feature values
may depend on the particular range of similarity values used in these
experiments or on the use of a single set of integral dimensions (hue,
saturation, and brightness). Nonetheless, any model of visual working
memory representations must be able to explain, at a minimum, the finding
that colour similarity does not impair performance over the range of values
tested here.

In the following sections, we describe potential models of working
memory that are incompatible with the finding of no performance
decrement for similar items, and then we evaluate possible explanations
for the finding that performance is actually improved for similar items.
Implications of a lack of impairment for similar items

The finding no performance impairment when the items in memory are similar is incompatible with the hypothesis that visual working memory representations are simply activated values of units in the feature space, with a simple summation across simultaneously remembered feature values. This possibility is illustrated in Figure 5 (top), which shows each memory representation as a noisy distribution of activity over colour space. When similar colours are being remembered, they produce overlapping representations in colour space, and the summation of activity may lead to substantial activation levels in the region between two peaks (Figure 5B). If, on a change trial, the test item fell into this region, it would be difficult to determine that it was, in fact, a changed value. Thus, this sort of simple model would predict a decline in memory performance when the items being represented are similar, and it is not compatible with the finding of improved performance for similar items.

The present results are also incompatible with the hypothesis that visual working memory representations are segregated from each other but compete for a highly limited pool of feature units. In such a model, a precise feature value is represented by means of the pattern of activation across units with different tuning curves. If, for example, a given shade of

![Figure 5. Hypothetical distribution of activity over the two-dimensional colour space for heterogeneous and homogeneous trial types. (A) and (B) show a purely additive model, whereas (C) and (D) show a model that includes local inhibition within the colour space.](image)
green is represented by the pattern of activity across units with overlapping
colour tuning curves, then a representation of a particular shade of green
should compete with the representation of another shade of green for access
to units that code the green region of the colour space. In contrast, a
representation of green and a representation of red would not compete in
this manner (unless the tuning curves were exceptionally broad). If memory
performance were limited by the number of available feature units, then
accuracy should be lower when similar colours are being represented. The
finding that performance is not impaired when similar colours are
represented thus argues against the possibility that memory accuracy is
meaningfully limited by competition for feature units.

Potential explanations of improved performance for similar
items

The finding of improved performance when the items being represented in
memory are similar to each other is quite surprising. The key result, from
Experiments 2 and 3, is that similarity leads to improved performance even
for the first item in the sequence and even when the observer cannot know in
advance whether that first item will be similar or dissimilar to the subsequent
items. These results rule out any explanations that posit an effect of
similarity on the initial encoding of the item, such as the concept of bands of
attention that has been used to explain effects of similarity in absolute
identification tasks (Luce, Green, & Weber, 1976). Instead, similarity must
cause either a change in the representations as they are being maintained or
a change in the comparison and decision processes that are used to generate
a same/different response once the test item is presented. There is no obvious
way in which similarity would produce a change in the comparison or
decision processes that would lead to improved performance, so we will focus
on mechanisms that might change the nature of the representation as it is
being maintained.

It is essential to keep in mind that the similarity effects observed for the
first item in the sequence cannot be explained by any further extraction of
information about the colour of the first item in the sequence when
subsequent items are similar. The first stimulus (and any high-resolution
iconic memory of the first stimulus) is no longer available when the
subsequent similar stimuli are presented, so there is simply no way to
obtain more information. Instead, any explanation must posit that similarity
influences performance indirectly, such as by enhancing the stability of the
representation.

There are at least three plausible ways in which maintenance processes
could have been influenced by similarity. First, the simultaneous
maintenance of similar items may lead to local inhibitory interactions within the colour space, producing a sharpening of the memory representations. This possibility is illustrated in Figure 5 (bottom), which shows the effects of adding lateral inhibition within the colour space (for a model of spatial memory that uses this sort of inhibition, see Spencer, Simmering, Schutte, & Schöner, in press). In contrast to the simple summation model shown at the top of Figure 5, the inhibitory interactions in this model will lead to reduced activity in the region between similar representations, which could improve the ability to determine that a test value is different from the original sample values (when the colour of the test item happens to land in the region of colour space between two sample representations). This would also cause modest shifts in the centroids of the representations, but computer simulations suggest that these shifts would not have much influence on performance in the present change detection task (J. S. Johnson, personal communication, 1 April 2007). Different methods would be necessary to observe the predicted centroid shift.

A second and related possibility is that one memory representation can provide an anchor point for another representation to reduce drift in the representations over time. If we assume that this anchoring is more effective for similar representations than for dissimilar representations, it could explain the improved performance observed on homogeneous trials. Indeed, representations may tend to drift in the absence of some sort of anchoring from nearby values, perhaps toward or even away from category prototypes (evidence for drift of this sort has been observed in spatial working memory—see Schutte, Spencer, & Schöner, 2003; Spencer & Hund, 2002). For example, the representation of a single shade of red may tend to drift towards a prototypical red value unless a different shade of red held in memory can be used as an anchor point. Anchoring could be beneficial even if there are no prototypes that cause systematic drift of one of the representations. If, for example, each representation tends to drift randomly and independently, then tying them together would cause the drifts to partially cancel. This would be analogous to examining a random walk in two independent variables compared with a random walk in the average of the two variables; the individual random walks would drift faster than the mean of the two random walks.

A third possibility is that similarity influences the allocation of attention during the maintenance process. Previous research has demonstrated that spatial cues presented after encoding can influence visual working memory performance (Griffin & Nobre, 2003), apparently by protecting the representations from disruption or decay during the maintenance period (Matsukura, Luck, & Vecera, 2007). It is possible that a similar advantage can be obtained by focusing attention onto a portion of the colour space.
No matter what the explanation, the present results indicate that visual working memory representations interact after the encoding stage, but in a manner that leads to improved rather than impaired change detection performance. Thus, the representations of different objects may be completely independent, as a simple “slot” view of working memory might suppose (see, e.g., Zhang & Luck, 2008). A slot account might be expected to predict no effect of similarity rather than a benefit of similarity, but the benefit of similarity could be explained on the basis of strategic differences in processing rather than direct interactions between the representations. Alternatively, the representations may interact directly, but not in a simple additive manner that causes them to blend together. Instead, they may interact by means of a mechanism that actively segregates the representations so that they do not become blended (see, e.g., Raffone & Wolters, 2001). That is, working memory may involve an inhibitory mechanism that acts to segregate the representations so they do not merge, but this inhibitory mechanism may cause the representations to repel each other. In this manner, a mechanism that ordinarily promotes independence may yield weak interactions as a side effect.

REFERENCES


*Manuscript received May 2007  
Manuscript accepted September 2007  
First published online March 2008*