

Object perception is selectively slowed by a visually similar working memory load

Alan Robinson

Department of Cognitive Science, University of California,
San Diego, USA



Alberto Manzi

Second University of Naples, Italy, &
Cognitive Electrophysiology Laboratory, New York State
Psychiatric Institute, USA



Jochen Triesch

Frankfurt Institute for Advanced Studies, Germany



The capacity of visual working memory has been extensively characterized, but little work has investigated how occupying visual memory influences other aspects of cognition and perception. Here we show a novel effect: maintaining an item in visual working memory slows processing of similar visual stimuli during the maintenance period. Subjects judged the gender of computer rendered faces or the naturalness of body postures while maintaining different visual memory loads. We found that when stimuli of the same class (faces or bodies) were maintained in memory, perceptual judgments were slowed. Interestingly, this is the opposite of what would be predicted from traditional priming. Our results suggest there is interference between visual working memory and perception, caused by visual similarity between new perceptual input and items already encoded in memory.

Keywords: visual working memory, perceptual judgements, object recognition, object classification, face processing, priming

Citation: Robinson, A., Manzi, A., & Triesch, J. (2008). Object perception is selectively slowed by a visually similar working memory load. *Journal of Vision*, 8(16):7, 1–13, <http://journalofvision.org/8/16/7/>, doi:10.1167/8.16.7.

Introduction

Visual working memory is a critical part of human visual cognition. While much research has addressed the capacity limits of visual working memory, little is known about how using visual working memory might influence other processing in the visual system. In this work we investigate if maintaining an item in visual working memory interferes with concurrent object perception. This question is motivated by research on the neural underpinnings of visual working memory.

Single unit recordings in monkeys suggest that two brain areas contribute to visual working memory representations: inferior temporal cortex and prefrontal cortex (Miller, Erickson, & Desimone, 1996; Miller, Li, & Desimone, 1993). In inferior temporal cortex neurons that respond to the initial display of a visual stimulus thereafter have changed response profiles while the monkey maintains that stimulus in memory. These changed response profiles are hypothesized to be critical to maintaining these items in visual memory. Since these neurons play an important role in object recognition, we hypothesized that their involvement in memory encoding could negatively impact perceptual abilities. fMRI research with humans also finds activity in prefrontal cortex and inferior temporal cortex during visual memory tasks (Druzgal & D'Esposito, 2003; Ranganath, DeGutis,

& D'Esposito, 2004), suggesting that humans also activate perceptual areas to remember the identity of an object. Jonides, Lacey, and Nee (2005) have suggested this is true for spatial memory as well, and advanced the hypothesis that, as a general rule, perceptual regions may play a role in all types of working memory storage. Pasternak and Greenlee (2005) have also reviewed evidence that supports the hypothesis that many types of working memory depend on activity in perceptual areas.

If perceptual areas play a supporting role in visual memory, how might perception and memory interact? Some research has been conducted on the relationship between visual memory and visual attention. The biased-competition theory (Desimone & Duncan, 1995) suggests that visual working memory provides top-down modulation of early visual areas, biasing attention toward visual stimuli that match the contents of memory. This has been explored in visual search experiments, where subjects must remember an item for a memory test and concurrently search for a target among distractors. If holding an item in memory increases the likelihood of it being attended in the search array then reaction time (RT) should slow when one of the distractors matches visual memory. The results vary somewhat; Downing and Dodds (2004) found no interaction between the contents of memory and search performance. Soto, Heinke, Humphreys, and Blanco (2005) found that when the memorized item matched the search target, RT was faster,

and when the memorized item matched a distractor, RT was slower. In contrast, when the memorized item always matched a distractor, Woodman, Boucher, Schall, and Luck (2004) found that subjects were faster. These results suggest that the contents of memory can either positively or negatively bias the selection of stimuli for attention and processing, depending on the task demands.

Additional evidence has been gathered outside of the domain of visual search. Kim, Kim, and Chun (2005) found that the Stroop effect was decreased by a working memory load, so long as the memory load was in the same modality as the task irrelevant, distracting part of the stimuli. This suggests that a memory load can reduce the salience or level of processing of stimuli in the same modality.

In this paper we investigate whether a memory load may also have the general effect of slowing object recognition and classification. This perceptual cost of maintaining an item in memory might be found whenever the contents of memory and the stimuli to be judged are similar enough to cause overlap in the neural populations, which support the recognition and maintenance of those items. If the two stimuli are in the same modality, then there might be some interference, but even greater interference should be found if the stimuli are of the same basic class.

Note that object recognition and classification are likely very closely linked in the brain; for conciseness hereafter we will refer to object classification when we mean the process of identifying an object and determining what class it belongs to.

Experiment 1

Experiment 1 investigated if a visual working memory load would slow object classification by measuring how quickly subjects could recognize if a novel human face was male or female. A delay in this perceptual judgment was taken as evidence that object classification had been slowed. The male/female perceptual judgment was made under three conditions: while holding a human face in working memory, a Fribble (a computer generated abstract 3D object), or under no memory load. If there is overlap in the brain areas that support perception and visual working memory, then we reasoned that the most perceptual slowing would be found in the *face memory* condition, where the memory stimuli and judgment stimuli were of the same visual class. We predicted relatively less slowing in the *Fribble memory* condition because of the low visual similarity between Fribbles and faces. To ensure subjects used working memory, they were given a same/different memory test at the end of each trial in the memory conditions. The *gender only* condition was included to estimate the speed of the gender judgment task when there was no memory load.

Methods

Participants

A group of 24 right-handed undergraduate students with normal or corrected-to-normal vision completed the experiment for course credit.

Stimuli

The stimuli used for the gender judgments were identical across conditions, only the memory load differed. Faces were generated using 3D modeling software (Poser 5.0, Curious Labs, Santa Cruz, California). Gender was manipulated by changing the geometry of the face, not its coloration or amount of hair. To generate “different” faces for the same/different memory task the geometry of the jaw, chin, mouth and lips, or nose was modified. Thus, both the gender and memory tasks relied on the overall shape of the face. Likewise, Fribble memory pairs differed primarily in the shape of their different parts. Each trial used a new memory stimulus not previously seen before, and for half of the trials unchanged stimuli were shown for the same/different test. Faces that were memorized were never used for the gender judgment task. All images were presented on a black background, at an average visual angle of $13^\circ \times 17^\circ$, using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Procedure

The procedure for the three conditions tested is shown in **Figures 1a–c**. In all conditions each task began with a one-word instruction, shown for 700 ms, which was included to make sure subjects did not become confused as to what kind of response was required on the next screen. In the *face memory* condition, trials began with the word “Memorize”, followed by the image of a face for 3500 ms. Then the gender judgment task began, signaled by the word “Gender”, shown for 700 ms. A face would be displayed for up to 1500 ms, during which the subject would indicate if it was male or female by pushing M or L on a keyboard. Both speed and accuracy were stressed. This process would then repeat once more for a different face. We included two gender judgments between the memorization and recall tasks to measure two different effects:

1. a task switching cost (Rogers & Monsell, 1995), incurred by switching from memorizing a stimulus to making a perceptual judgment, and
2. the cost thereafter of maintaining a stimulus in memory while making additional perceptual judgments.

After the gender task ended the subjects saw the word “Recall” for 700 ms, followed by the image of a face for

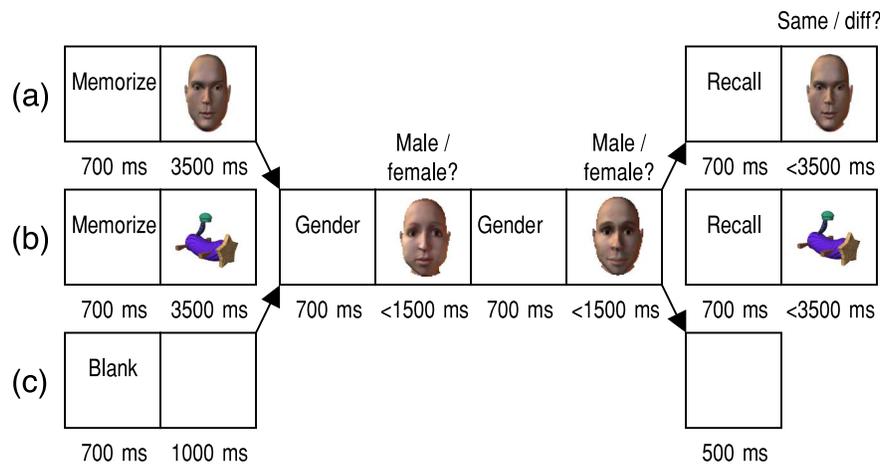


Figure 1. The three conditions in Experiments 1 and 2. Each box represents a screen displayed to the subjects during a trial, with the number of milliseconds of display listed below each box. < symbol indicates a response was required from the subject and the display would terminate after a keypress. (a) *Face memory* condition. (b) *Fribble memory* condition (same timing as *face memory* condition, but different memory stimuli). (c) *Gender only* condition, with no memory load, but the same gender task.

up to 3500 ms. Subjects were to indicate if this was the same face as seen at the beginning of the trial after the “Memorize” label, or a different face, by pushing the A or X key on a keyboard.

In the *Fribble memory* condition subjects memorized a Fribble instead of a human face, but still conducted the same gender judgment task as in the *face memory* condition. The timing of stimulus presentation in these two conditions was identical.

In the *gender only* condition subjects just judged the gender of two faces and had no other task. Since the *gender only* condition had no memory task, we shortened the total display time for the screens before and after the gender judgments, so that the overall pacing of the task would not be dramatically different between conditions.

Subjects practiced all 3 conditions before experimental trials were collected. Each condition was run separately in blocks of 32 trials. Subjects completed 6 blocks (2 of each condition), the order of which was counterbalanced to minimize any learning effects. Gender was counterbalanced within each block, such that all possible combinations of gender of the memorized face, and gender for the first and second gender judgments were tested an equal number of times. This meant that knowing the gender of the memorized face had no predictive utility for the first or second gender judgment, and likewise, knowing the answer to the first gender judgment did not predict the gender on the second gender judgment.

Results and discussion

To measure if maintaining an item in visual memory slowed face classification, we examined RT on the gender judgment task. We analyzed RT from all trials, independent of the correctness of gender or memory response. An

incorrect response on one of the tasks might indicate that a subject chose to ignore that task for that trial, but we reasoned that if subjects were ignoring one of the tasks this should only decrease interference between tasks and therefore reduce the effect measured. Furthermore, some of the faces are relatively ambiguous in gender, with some subjects indicating they were male and some indicating they were female, making it difficult to determine what an incorrect response should be. In any case, however, the results do not change significantly if we only include trials where the responses were correct.

A repeated measures ANOVA on condition (*face*, *Fribble*, or *gender only*) and judgment order (*first* or *second* face) found a main effect of condition, $F(2,46) = 28.8, p < 0.0001$; a main effect of judgment order, $F(1,46) = 72.7, p < 0.0001$; and an interaction between these two, $F(2,46) = 40.7, p < 0.0001$. We then conducted six paired *t*-tests between conditions to determine which conditions drove the main effects (Figure 2a). To control for multiple comparisons we report the Bonferroni–Holm corrected *p* value (as described in Aickin & Gensler, 1996, see also Holm, 1979), which keeps the probability of a type-I error (false positive) below 0.05. Both the first and second gender evaluations in the *face memory* condition were slower than the corresponding judgments in the *Fribble memory* condition (by 41 ms; $p < 0.01$, and 37 ms; $p < 0.007$, respectively). This indicates that holding an item in memory slows perceptual judgments for visually similar items. Comparing the *Fribble* condition to *gender only*, however, only the first gender evaluation was found to be significantly slowed (by 113 ms; $p < 0.0001$). This suggests that there is a task-switching cost incurred when switching from memorizing an object to making a perceptual judgment, but after that switching cost additional judgments are not slowed when the item in memory is a different type of object.

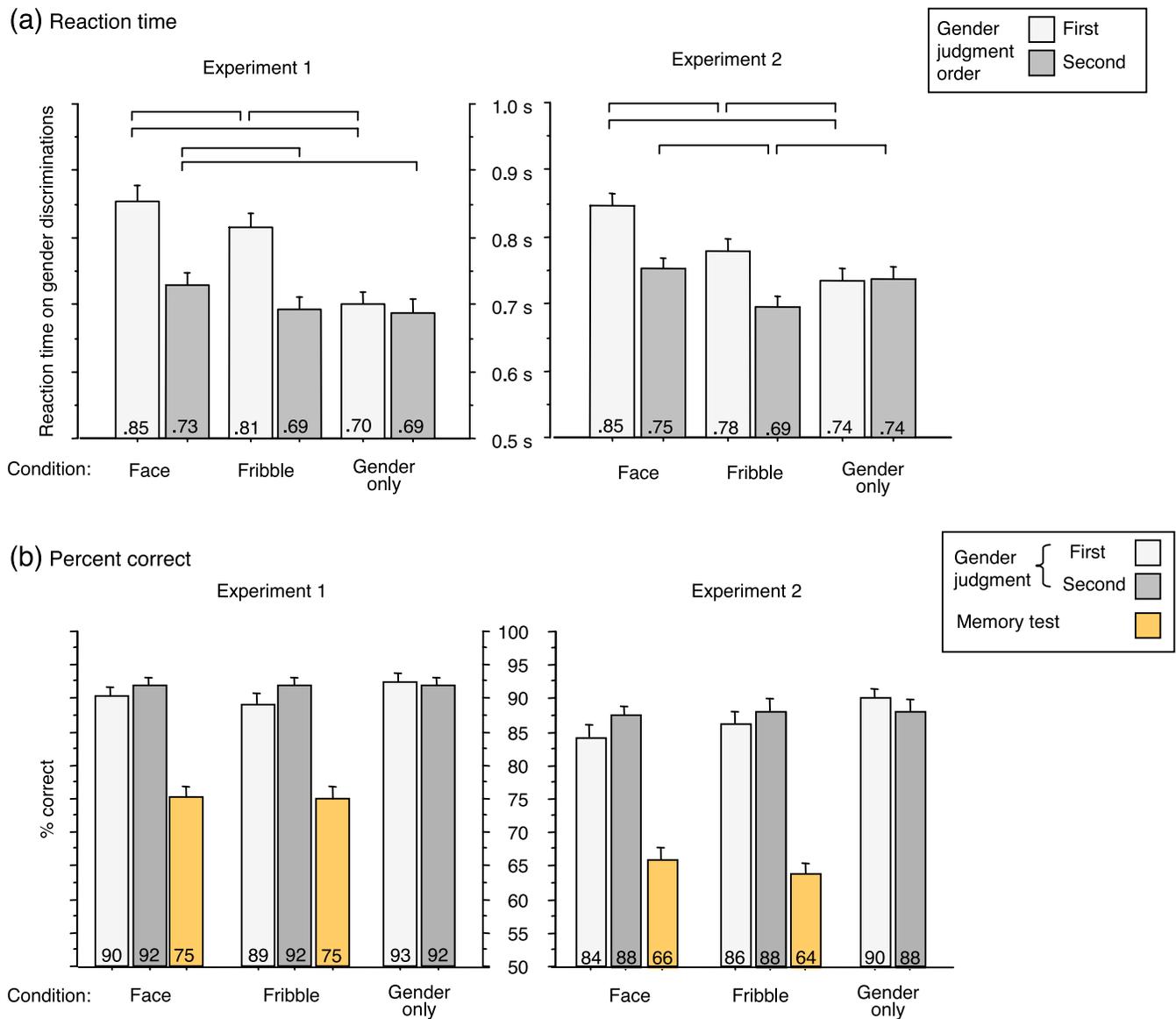


Figure 2. (a) Average reaction times on gender judgment task for the first face (white bar) and second face (gray bar) for each trial, split by condition. Differences marked with horizontal lines are statistically reliable at an alpha level of 0.05 or better, after applying a Bonferroni–Holm correction for multiple comparisons using paired *t*-tests. (b) Accuracy on gender judgements for all 3 conditions and on the memory test in the first 2 conditions. The number at the bottom of each bar indicates the actual numerical value for that condition. All error bars represent 1 *SEM*.

In conclusion, this experiment suggests two kinds of costs of using visual working memory:

1. A task-switching cost, which is incurred only once and does not depend on the visual similarity between the object being memorized and being judged, and
2. a persistent slowing of visual judgments while maintaining a similar visual working memory load.

While the task-switching cost is larger, it is only incurred once, so the persistent slowing could be more significant overall, depending on the task and the number of

judgments that must be made while the working memory load is maintained.

Could these results be due to differing levels of difficulty of memorizing Fribbles and faces, and not because of visual similarity? Jolicoeur and Dell’Acqua (1998, 1999) demonstrated that consolidating (transferring) an item into working memory can slow RT on a second task when performed simultaneously. This effect was found even though the two tasks were in different modalities (visual and auditory) suggesting it was caused by a modality independent central bottleneck. The timing of our task, however, suggests consolidation is unlikely to explain our findings. We used a fairly long interval

between the onset of the memory stimulus and the onset of the gender judgment stimulus (4.2 s), whereas Jolicoeur & Dell’Acqua’s work suggests that consolidation is complete after 2 s (see also Woodman and Vogel (2005) for evidence that consolidation can complete in as little as 100 ms). Furthermore, consolidation costs increase with the difficulty of memorizing a stimulus, so to explain our results it would be necessary for the slowest gender recognition to occur under the most difficult memory condition. Accuracy on the memory test, however, suggests that faces and Fribbles were of equal difficulty to memorize (both 75% correct for the same/different test; see Figure 1b). RT on the memory test, if anything, suggests that Fribbles were the more difficult stimuli (an average of 1.54 s for faces, and 1.81 s for Fribbles). Thus the difficulty of the memory test does not provide a compelling explanation of the RT effects in the gender judgment task.

We also examined accuracy on the gender judgments. Average accuracy was only 91%, probably because there were a number of reasonably ambiguous faces generated by using the texture of one gender combined with the shape of another. There was no effect of memory condition on gender accuracy.

Experiment 2

The RT difference between *face memory* conditions and *Fribble memory* conditions in Experiment 1 supports the hypothesis that maintaining an item in visual memory slows object classification. Experiment 1 did not, however, control for use of phonological encoding. Different parts of the *Fribbles* could be given names, whereas subtle variations in facial features are harder to represent verbally. Perhaps in the *Fribble* condition subjects used phonological working memory, whereas in the *face* condition they used visual memory. Thus the effect in Experiment 1 might indicate that maintaining visual memory interferes with object classification, but maintaining phonological memory does not. In Experiment 2 we tested this hypothesis by disrupting subjects’ ability to use phonological encoding.

Methods

Participants

A new group of 24 right-handed undergraduate students with normal or corrected-to-normal vision completed the experiment for pay.

Procedure

Experiment 2 used the same stimuli and procedure as Experiment 1 except that subjects were given a

phonological distracter task. For the duration of the experiment subjects were instructed to verbally recite three letters or numbers continuously (‘234’, ‘RST’, etc.). To prevent habituation, subjects switched to a different verbal sequence at the end of each block. This type of task has been shown to greatly decrease the ability to use phonological encoding (Baddeley, 1992).

Results and discussion

As in Experiment 1, a repeated measures ANOVA on condition (*face*, *Fribble*, or *gender only*) and judgment order (*first* or *second* face) found a main effect of condition, $F(2,46) = 17.1$, $p < 0.0001$; a main effect of judgment order: $F(2,46) = 49.8$, $p < 0.0001$; and an interaction between these two, $F(2,46) = 58.9$, $p < 0.0001$. We conducted six paired *t*-tests between conditions to determine which conditions drove the main effects (Figure 2a), using Bonferroni–Holm corrected *p* values. As in Experiment 1, both the first and second gender judgments in the *face memory* condition were significantly slower than the corresponding judgments in the *Fribble memory* condition (by 69 ms; $p < 0.0001$, and 58 ms; $p < 0.0001$, respectively). This suggests that the cost of working memory measured in Experiments 1 and 2 is due to visual similarity and not due to reliance on different working memory subsystems.

As in Experiment 1, only the first gender judgment in the *Fribble* condition was significantly slowed as compared to the *gender only* condition (by 43 ms; $p < 0.03$). In contrast to Experiment 1, however, the second gender judgment in the *Fribble* condition was actually 44 ms faster than the second judgment in the *gender only* condition ($p < 0.02$). This would suggest that, relative to no memory load, holding a *Fribble* in memory would actually speed face judgments. Note that this was not the case in the first experiment, so perhaps it is due to inter-subject variability. It is worth considering, however, that the overall task demands in the *gender only* condition are very different from the memory conditions. The time between making pairs of gender judgments is much shorter, and perhaps this causes fatigue for some subjects, leading to slower performance. We will revisit this issue in Experiment 3.

Accuracy on the memory test was an average of 66% for faces and 64% for Fribbles (Figure 2b), and RT for the memory test was 1.28 s and 1.43 s, respectively. As in Experiment 1, this suggests that the effects on gender judgment RT are not due to memory task difficulty. It is interesting to note that memory accuracy was lower in Experiment 2. This may reflect a speed/accuracy tradeoff, since subjects were also faster to make the same/different response in Experiment 2. Average accuracy on the gender task was also slightly reduced (87%) suggesting an overall decrease in performance across tasks, probably due to the addition of the phonological distracter task.

Experiment 3

Experiments 1 and 2 found that memorizing a face slows face classification more than memorizing a Fribble. It could be, however, that the slowing is not due to the memory load but occurs because the subjects are looking at a face for a while before they switch to doing the gender task. For instance, staring at a face has been shown to cause visual adaptation, which can influence categorical perception of faces (Webster, Kaping, Mizokami, & Duhamel, 2004). Perhaps visually adapting to a face can also cause slowed perceptual processing on the next task. Thus, memorizing a face would slow perception, but only because it leads to adaptation. In Experiment 3 we test this hypothesis with a *just attend* condition, which is visually identical to the *face memory* condition except that subjects are not asked to memorize the face. If it is the memory load that causes the perceptual slowing, then gender judgments in the *just attend* condition should be faster than in the *face memory* condition. Furthermore, the *just attend* condition can be taken as a more appropriate estimate of the speed of the gender judgments when there is no memory load, since the trial timing is exactly the same between the *just attend* and *face memory* conditions. We also included the *gender only* condition from previous experiments so we could compare it to the *just attend* results.

Methods

Participants

A new group of 24 undergraduate students with normal or corrected-to-normal vision completed the experiment for course credit.

Procedure

Experiment 3 used the same *face memory* and *gender only* conditions from Experiment 1. A new condition was added, which was visually almost identical to the *face memory* condition, but with a different initial task (Figure 3). In this condition (*just attend*) subjects were instructed to attend to the first face in the trial but not memorize it. To ensure that subjects attended to the face for the entire 3.5 s that it was shown, subjects were given

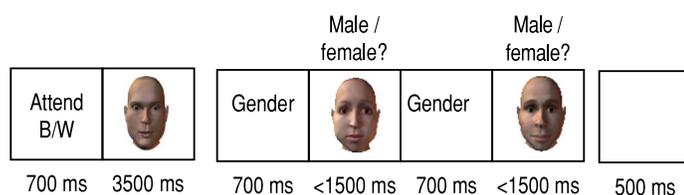


Figure 3. The *just attend* condition in Experiment 3.

the task of monitoring for the face to change from color to grayscale. This would occur for 500 ms and could happen at any point within the 3.5 s. When this occurred subjects were to verbally report to the experimenter that the item had changed. The *just attend* condition included 4 extra catch trials in which this happened; no gender judgments were collected for these trials. As in the first two experiments, conditions were run in blocks of 32 trials, with the order of blocks counterbalanced to minimize learning effects.

Because there was no memory task in the *just attend* trials, there was no recall face shown at the end of the trial. In addition, because subjects were instructed to attend, but not memorize the first face, the instruction text at the beginning of the trial read “attend b/w”. In all other ways, however, the *just attend* and *face memory* trials were visually identical and also had identical timing.

Results and discussion

All subjects detected all of the grayscale catch trials. In addition, only one subject, on one trial, reported that a face changed to grayscale when this did not in fact happen. Thus we conclude that the grayscale task successfully caused the subjects to attend to the initial face in the *just attend* condition, throughout the 3.5 s it was displayed.

A repeated measures ANOVA on condition (*face*, *just attend*, or *gender only*) and judgment order (*first* or *second* face) found a main effect of condition, $F(2,46) = 19.2$, $p < 0.0001$; a main effect of judgment order: $F(2,46) = 55.3$, $p < 0.0001$; and an interaction between these two, $F(2,46) = 13.4$, $p < 0.0001$. We conducted six paired *t*-tests between conditions to determine which conditions drove the main effects (Figure 4a), calculating Bonferroni–Holm corrected *p* values. Both the first and second gender judgments in the *face memory* condition were significantly slower than the corresponding judgments in the *just attend* condition (by 82 ms; $p < 0.0001$, and 41 ms; $p < 0.001$, respectively). Since the two conditions were as alike as possible except that *just attend* contained no memory load, this strongly suggests that the slowing observed is due to the cost of using visual working memory, rather than a side effect caused by viewing the stimuli in order to memorize them.

As in Experiment 1, RT was slower on the first and second faces in the *face memory* condition than in the *gender only* condition (by 98 ms; $p < 0.0001$, and 25 ms; $p < 0.037$, respectively). This replication provides further support for the hypothesis that a memory load slows face classification.

We also found that the RT for the *gender only* condition was slower than in the *just attend* condition for the second gender judgment in a trial, although only by 20 ms, supporting the idea that giving subjects a string of gender judgments with minimal pauses in between actually leads

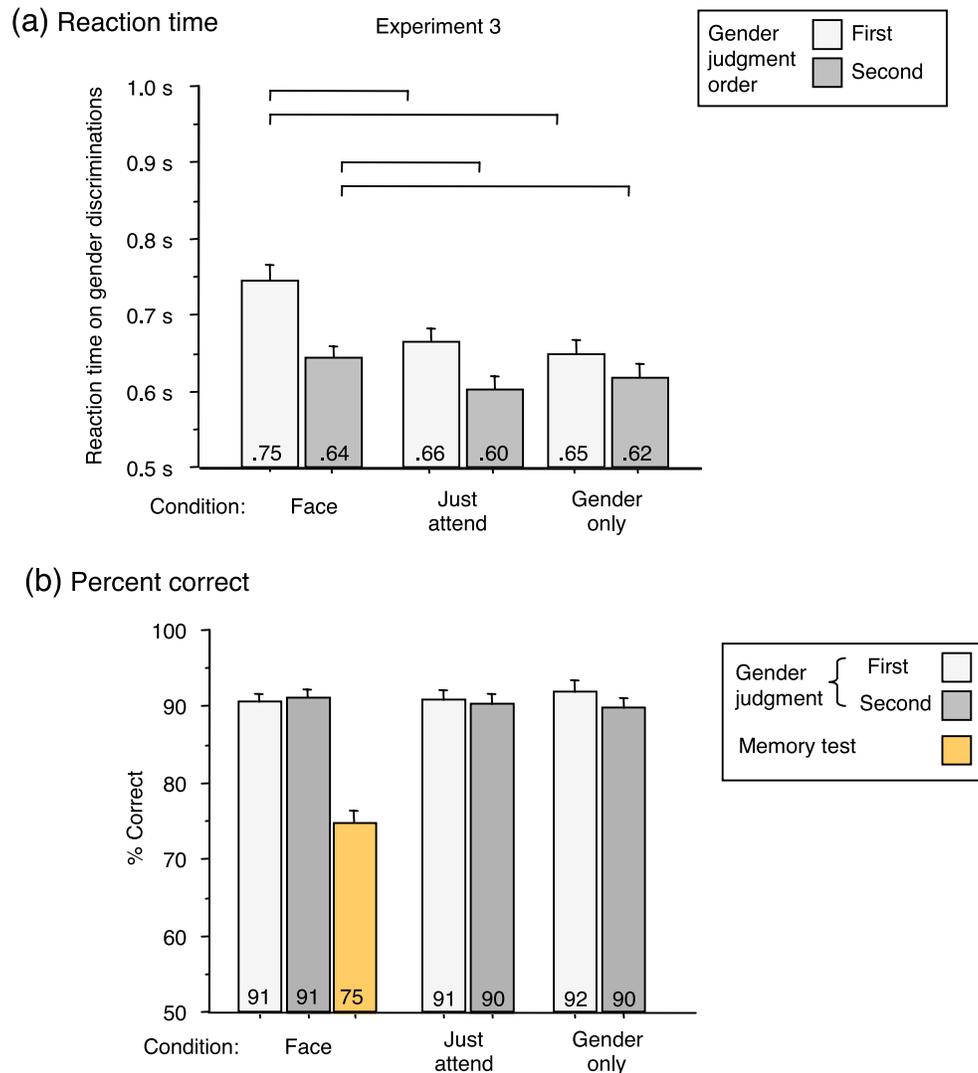


Figure 4. (a) Average reaction times on gender judgment task for the first face (white bar) and second face (gray bar) for each trial, split by condition. Differences marked with horizontal lines are statistically reliable at an alpha level of 0.05 or better, after applying a Bonferroni–Holm correction for multiple comparisons using paired *t*-tests. (b) Accuracy on gender judgements for all 3 conditions, and on the memory test in the face condition. Error bars represent 1 *SEM*.

to slower performance. This also helps explain why in [Experiment 2](#) the second face judgment while holding a Fribble in memory was faster than the corresponding no-memory load face judgment.

Mean accuracy on the face memory test was 75%, and mean accuracy on the gender task was about 90%, across conditions ([Figure 4b](#)). As in previous experiments, this shows that subjects completed both the gender and memory tasks.

Experiment 4

In the prior 3 experiments we repeatedly found evidence that face classification was slowed when a face was held

in memory. In [Experiment 4](#) we investigated if this memory-based slowing is limited to the domain of faces, or if it generalizes to other visual categories. For instance, perhaps holding a face in memory is particularly detrimental to object classification for all object classes, and that is why we saw greater slowing in the face memory conditions. Another possibility is that only face perception suffers from a working memory load. To rule out these explanations, we tested a different perceptual judgment.

We selected the task of judging if a human body pose was natural (within the range of normal human movement) or unnatural (difficult or impossible to achieve unless a person was hyperflexible). While this judgment might include non-visual components, it requires that the subject be able to recognize the relative locations and orientations of the limbs of the body. If this recognition process is slowed, we reasoned that the judgments of

natural or unnatural pose would also be slowed. We selected this particular task because it was similar to the male/female task in that it was (1) a binary decision, (2) did not require extensive training, (3) would require attention to position and configuration of high-level visual features and could not be done just using low-level components such as the relative power at different spatial frequencies.

Methods

Participants

A new group of 24 right-handed undergraduate students with normal or corrected-to-normal vision completed the experiment for course credit.

Stimuli

We reused the faces generated for the first 3 experiments. For the *body memory* conditions and for the body judgments, we generated new images using 3D modeling software (Poser 5.0, Curious Labs, Santa Cruz, California). The bodies were fully clothed males and females, with the head removed at the neck (to prevent any interaction with the face stimuli). The bodies were viewed from a variety of different angles, and the positions of the legs, arms, and torso were varied randomly.

Half of the figures were in ‘natural’ poses, and half were in ‘unnatural’ poses. Unnatural poses were created by setting joint angles to be outside of the possible range for a typical person. We conducted a small norming study with 6 subjects to ensure that the majority of subjects were in agreement as to the ‘naturalness’ or ‘unnaturalness’ of the poses. Note that while there was general agreement, there were some individual differences in the ratings subjects gave.

For the memory test using the body stimuli, subjects always viewed the same figure from the same angle in the memorize and recall stages. If the pose was different between memorize and recall, it differed in terms of a single joint angle, either for the foot, arm, leg, shoulder, or torso. Half of the memory stimuli were in unnatural poses; if the joint angle changed between memorize and recall, the change never switched the pose from natural to unnatural, or vice versa.

Procedure

There were four conditions, all with identical timing (Figure 5). In the *face face* condition, subjects memorized a face, and judged the gender of two faces per trial. This condition was identical to the *face memory* condition used in the first 3 experiments. In the *face body* condition, the subjects memorized a face, and then conducted two body judgments per trial (“Is the pose natural or unnatural?”). The *body face* condition and the *body body* condition likewise tested the effect of memorizing a body pose on face judgments or body judgments, respectively. The timing of the four conditions was the same as the memory conditions in Experiment 1. For the body judgments, subjects responded by pushing the N key (natural) or K key (unnatural) on a keyboard.

Each condition consisted of 32 trials, and the order of the trials was randomized. The order of the conditions was counterbalanced between subjects to minimize practice effects.

Results and discussion

The results are shown in Figure 6. We conducted a repeated measures ANOVA on RT, with factors of

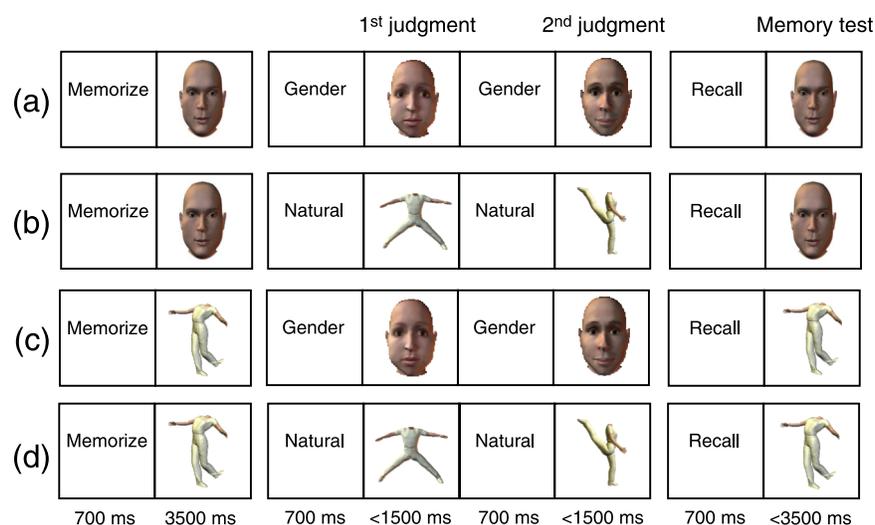


Figure 5. The four conditions in Experiment 4.

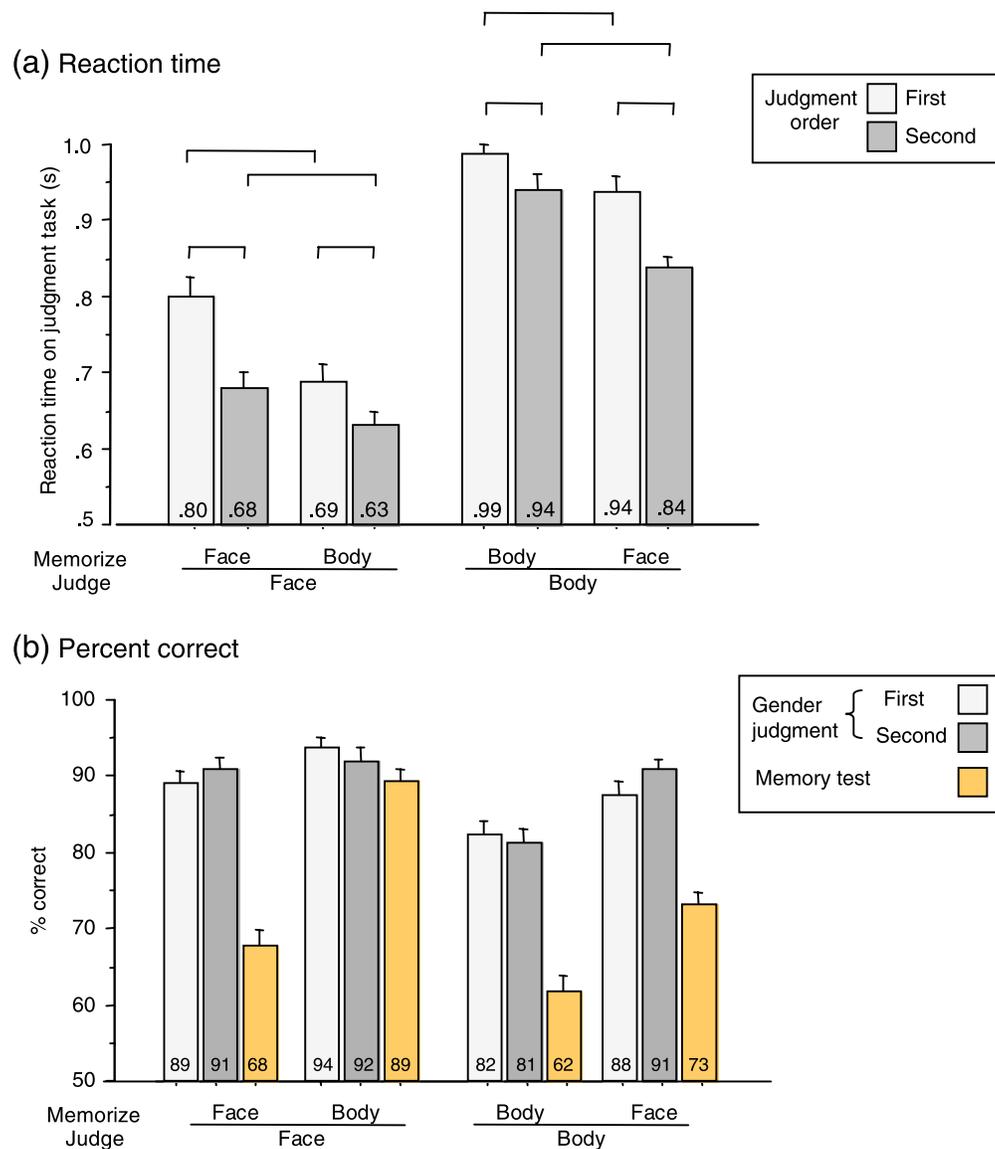


Figure 6. (a) Average reaction times for the first judgement (white bar) and second judgement (gray bar) for each trial, split by condition. Differences marked with horizontal lines are statistically reliable at an alpha level of 0.05 or better, after applying a Bonferroni–Holm correction for multiple comparisons using paired *t*-tests. (b) Percent correct on the first and second gender judgements, and on the memory test, split by condition. Error bars represent 1 *SEM*.

memory type (*face* or *body*), judgment type (face gender or body posture), and judgment order (first or second judgment). We did not find a main effect for the memory type factor: $F(1, 23) = 0.06$, $p < 0.8$, indicating that the type of stimuli memorized did not have an overall effect on judgment reaction time. This suggests that there is nothing special about holding a face (or body) in memory that slows object perception. The judgment type factor was significant: $F(1,23) = 279$, $p < 0.0001$, indicating that subjects were significantly faster at judging the gender of faces than the naturalness of body poses. There was also a main effect of judgment order: $F(1,23) = 32$, $p < 0.0001$, indicating that subjects were significantly slower at making the first judgment in a trial than the second. This is evidence of the task-switching cost of switching from

memorizing a stimulus to recognizing and classifying a new stimulus. Importantly, there was an interaction between the memorize factor and the judgment factor: $F(1,23) = 23$, $p < 0.0001$. This indicates that the effect on reaction time of memorizing a face or a body depends on whether you are judging a body or a face, with slower judgments made when the memorized stimulus was of the same class as the judged stimulus. This supports our hypothesis that object classification is slowed by a concurrent memory load, if it is of the same visual class.

Because our conditions were fully crossed (we tested the effect of each type of memory load on the reaction time for both types of judgments), it is possible to test for our main hypothesis using an ANOVA test, as we have just described. We next considered if we could use paired

t-tests to determine which conditions drove the main effects and interactions found with the ANOVA. To control for multiple comparisons we report the Bonferroni–Holm corrected *p* value.

We found that when the memory stimulus matched the type of the judgment stimulus, subjects were slower at making the judgment, for all conditions (by 111 ms for the first face judgment ($p < 0.0001$), by 49 ms for the second face judgment ($p < 0.042$); by 49 ms for the first body judgment ($p < 0.045$), and by 102 ms for the second body judgment ($p < 0.0001$)). Thus, it seems that both the face judgment and body judgment conditions, for both the first and second faces in a trial drove the interaction effect revealed by the ANOVA.

The accuracy data (Figure 6b) shows that subjects found the memory tests harder when the memory stimulus was of the same class (face or body) as the judgment stimulus (Figure 6b). When subjects had to memorize the same class as they judged, face memory accuracy dropped from 73% to 68% and body memory dropped from 89% to 62%. The fact that performance was higher when there was less potential for interference, however, suggests that subjects were still attending to both the judgment task and the memory task, and that the lower performance is just due to the increased task difficulty.

Interestingly, we found that there was a slight decrease in accuracy on the judgment tasks when the memory stimulus was of the same class as the judgment task. This was exceedingly slight for the gender judgments (a drop from 93% to 90%) but was more noticeable for the body judgments (a drop from 89% to 81%). Our experiments were not explicitly designed to test for this kind of effect, but our results do suggest that future research may find that the costs of maintaining visual working memory include both slowed reaction time and increased errors.

In conclusion, Experiment 4 supports the generality of the effect repeatedly measured in the first three experiments. Holding an item in visual working memory appears to slow down object classification, when both stimuli come from the same class. This experiment shows that this is true not just for faces, which have often been argued to have specialized recognition mechanisms in the brain, but also for body poses, a very different category of natural objects.

General discussion

In two separate experiments we found that gender identification was slower when a face was held in memory than when a Frizzle was held in memory. In a third experiment we found that gender identification was slowed when a face was held in memory, relative to a visually identical condition that did not require visual working memory. In a fourth experiment we found that judgments about human bodies were slower when a

human body was held in memory than when a face was held in memory (and vice versa). From these results we conclude that object classification is slowed by a visual working memory load when the contents of memory are similar to the object being classified. We conjecture that this may be due to the overlap of the neural populations involved in memorizing an object and perceiving an object. This suggests that the classification might be even slower if the memorized object is highly similar to the one to be perceived. For example, subjects may perceive a face particularly slowly if it is similar to the face held in memory. In order to test this hypothesis, we re-analyzed the data from our first three experiments (those which had enough face memory trials to allow this analysis) in the following way. We reasoned that faces of the same gender will tend to have a higher overlap in their neural representations than faces of opposite gender. Thus, when holding a face in memory, subjects may be particularly slow at judging faces of the same gender. This is exactly what we found (Figure 7).

We conducted separate repeated measures ANOVAs on the RT for the gender judgments in the *face memory* conditions in Experiments 1, 2, and 3, with two factors: (1) *gender match*: did the gender of the face judged match the gender of the face memorized at the beginning of the trial, and (2) *order*: was the face the first or second one judged in the trial. For all experiments there was a main effect of *gender match*, with faster judgments made when the gender of the face judged was *different* from the gender of the face memorized (by 35 ms in Experiment 1, $F(1,23) = 18.5$, $p < 0.0003$; by 45 ms in Experiment 2, $F(1,23) = 35.1$,

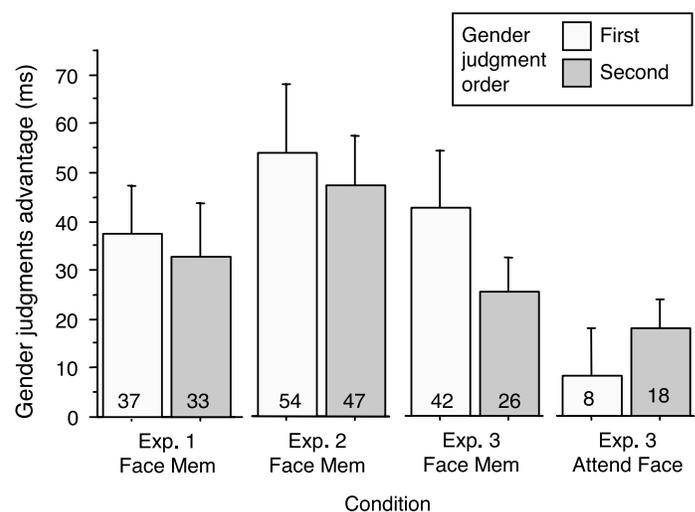


Figure 7. Average reaction time advantage on gender judgment task when the gender of the first face seen in the trial (memorized or just attended) does not match the gender of the face being judged. Calculated by subtracting RT on trials where the gender does not match from trials where it does. Results are split by the first judgment (white bar) and second judgment (gray bar) for each trial. Error bars represent 1 SEM.

$p < 0.0001$; and by 34 ms in [Experiment 3](#), $F(1,23) = 20.6$, $p < 0.0001$). There was also a main effect of *order* for all experiments, because subjects were faster to judge the second face, as already discussed in previous results sections. Interestingly, however, there was no interaction effect between *order* and *gender match* for any of the experiments ([Experiment 1](#), $F(1,23) = 0.1$, $p < 0.74$; [Experiment 2](#), $F(1,23) = 0.14$, $p < 0.71$; [Experiment 3](#), $F(1,23) = 1.9$, $p < 0.18$). This suggests that the effect of the memorized stimulus did not decrease much over time or intervening visual displays.

Note that this effect is the opposite of what one might expect from priming. Priming would predict that responses would be faster if the gender of the memorized face matches the gender of the face to be judged. Our analysis, however, reveals the reverse: gender matches actually lead to *slower* gender judgments. Previous studies have found reverse priming (referred to as the *negative compatibility effect*) for stimuli that are masked so that subjects cannot identify them (Eimer & Schlaghecken, 1998; but see Lleras & Enns, 2004, for methodological concerns). In all such experiments, however, the negative compatibility effect goes away when the prime is detectable, and traditional priming is found instead. Since none of our stimuli were masked, the negative compatibility effect cannot explain our results.

Furthermore, there is no statistically significant evidence that this gender mismatch advantage decreases over time. If the effect was somehow due to priming, then there should be less priming observed on the second face judgment than on the first.

Another possible cause for the gender mismatch advantage is visual adaptation. Adapting to a face with a clearly defined gender can shift the perception of a gender neutral face in the opposite direction of the gender that the subject adapted to (Webster et al., 2004). Though the effect of adaptation on reaction time with non-gender-neutral faces has not been explored, it is plausible that slight shifts in category boundaries would influence reaction time. For instance, if adaptation to a female face made a novel male face appear even more masculine, people might be slightly faster to identify that the face is masculine, and slightly slower to identify the face correctly as female if it was actually feminine.

The posited effect of adaptation could explain the gender mismatch advantage in [Experiments 1](#) and [2](#). When the subject views a face (in order to memorize it) gender selective neurons in visual areas might adapt slightly, shifting the subject's perception of all faces slightly away from the gender of the face being memorized. This categorical shift would speed later judgments when they did not match the memorized face's gender, and slow them when they matched. Again, however, if this were to explain the effect, one would expect that the adaptation would dissipate with time, and by the second gender judgment the effect would be significantly reduced or eliminated.

Furthermore, if adaptation is playing a role, then its effects should be present in the *just attend* condition in [Experiment 3](#), in addition to the *face memory* conditions of [Experiments 1](#), [2](#), and [3](#). We therefore conducted a repeated measures ANOVA on the *just attend* condition, with two factors: (1) *gender match* between the attended but not memorized face and the judged faces, (2) *judgment order*. There was a main effect of *gender match* ($F(1,23) = 4.8$, $p < 0.04$) and *order*, but no interaction between these two factors ($F(1,23) = 0.8$, $p < 0.39$). Gender mismatches were faster, just as in the *face memory* conditions of [Experiments 1](#), [2](#), and [3](#). The effect, however, was much smaller, especially for the first face judgment (only 8 ms), which is when the effect of adaptation should have been maximal. Thus, it seems possible that visual adaptation could be playing a minor role, but it cannot explain most of the gender mismatch advantage. Based on our results, it does not appear that the slowing of object recognition and classification can be explained by priming or adaptation. The theory that perception is slowed by the presence of visually similar items in working memory, however, does account for all the effects measured. We speculate that visual similarity is important because it increases the likelihood that the same neural populations that helped identify an object and then maintain it in memory will be needed to recognize objects seen during the maintenance period. Our current results suggest this underlying cause, but further neurophysiological experiments will be necessary to directly test this hypothesis. In a potentially related work, Soto, Humphreys, and Rotshtein (2007) found that having subjects memorize a cue at the beginning of each trial in a visual search task would speed RT when that cue appeared around the target item, even though on average the cue was not predictive of target location. A similar, but smaller effect was found when subjects only had to attend to the cue, but not memorize it, though fMRI scans suggested that these two conditions lead to distinctly different patterns of activation. While it seems that their result depends on spatial attention, and that ours does not, as we use no distractors, it would be very interesting to compare the results of their study to an fMRI version of our [Experiment 3](#), with a *face memory* and a *just attend* condition. In another related fMRI and behavioral experiment, Jha, Fabian, and Aguirre (2004) showed that memory performance is reduced when a visual stimulus is presented during the retention period, with greater reduction in performance when the object to be remembered is of the same class as the distractor stimulus presented during the retention period. When using face stimuli as both the memory stimulus and the distractor, they found elevated activity in prefrontal cortex and the fusiform face area, which may suggest that the prefrontal cortex plays a role in allowing concurrent perception and visual memory retention (note, however, that no task was performed on the distractor, so the elevated activity could have just been related to ignoring the distractor).

Our work has demonstrated that a visual memory load slows perception of objects of the same class, for both bodies and faces. Further work will be necessary to verify that the effect generalizes to other perceptual tasks and domains. We hypothesize that the finding will generalize as long as the visual areas that drive the perceptual judgments also play a role in maintaining visual working memory. Thus, the cost of using visual memory may be frequently encountered during everyday, natural tasks. This suggests that an optimal strategy during natural tasks that require perception and memory is to trade off visual working memory capacity for perceptual speed. Indeed, research on natural tasks suggests that people frequently encode just one item at a time (Gajewski & Henderson, 2002), or even just select aspects of one item (Droll, Hayhoe, Triesch, & Sullivan, 2005; Hayhoe, Bensinger, & Ballard, 1998; Triesch, Ballard, Hayhoe, & Sullivan, 2003). Such memory usage is significantly less than the estimated capacity limits of visual working memory of about 4 items (Luck & Vogel, 1997; but see Alvarez & Cavanagh, 2004, for a smaller estimate of memory capacity). The use of visual working memory may, in addition to slowing perception, cause perceptual errors. Change detection experiments, such as Droll et al. (2005) suggest that during natural behavior the contents of memory can sometimes override perception, leading to missed changes, even for the object of central interest in a task. Whether these errors are due to the lack of attending to perceptual input, or actual errors in the process of recognition and classification should be investigated in future research.

Our work also has implications for the literature on how visual search and visual memory interact. Woodman, Vogel, and Luck (2001) found that filling visual memory to capacity with the color of squares or the orientation of “C”s did not slow the rate of search for similar items on an item-by-item basis (the increase in RT as a function of search set size was the same across different memory loads). They did, however, find a slowing of RT by a fixed amount, independent of set size, which, due to the timing of their experiment was probably caused at least partially by memory consolidation. Given our results, one might also expect an item-by-item slowing of visual search. One possible explanation is that the slowing we observed occurs when the interfering stimuli are first presented visually. The delay might be the signature of a process that protects the contents of memory from being overwritten or degraded by the new input. Once this process is complete there might not be any further delay in processing the current input. The same process might be required for each new visual display, however, such as for the second of the two faces shown in our gender judgment task. In Woodman et al. (2001) subjects searched just one display per trial and were instructed to not make eye movements. Because visual input was fairly constant, each trial might have triggered the delaying process just once, independent of set size.

Outlook

Our work suggests that future research on visual working memory should consider how memory is used, and what potential tradeoffs there are to using memory during natural behavior. Furthermore, future work should explore how the same neural circuits can maintain the trace of a visual item in memory while at the same time recognizing new objects, even those that are very similar to the memorized item.

Acknowledgments

We thank Christof Teuscher, Mary Hayhoe, Hal Pashler, Emo Todorov, and Geoffery Woodman for comments on an earlier draft; Pepper Williams and Michael Tarr for donating the Fribble stimuli; and Tara Sears for assisting with subjects. This work is supported by a grant from the University of California Academic Senate. Alan Robinson was supported by NSF Grant DGE-0333451 to GW Cottrell/VR de Sa and NSF CAREER Award 0133996 to VR de Sa. Alberto Manzi was supported by funding from the Second University of Naples. Jochen Triesch was supported by the Hertie Foundation and by EU Marie Curie Excellence Center Grant MEXT-CT-2006-042484.

Commercial relationships: none.

Corresponding author: Alan Robinson.

Email: robinson@cogsci.ucsd.edu.

Address: Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515, USA.

References

- Aickin, M., & Gensler, H. (1996). Adjusting for multiple testing when reporting research results: The Bonferroni vs Holm methods. *American Journal of Public Health, 86*, 726–728. [PubMed] [Article]
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111. [PubMed]
- Baddeley, A. D. (1992). Working memory. *Science, 255*, 556–559.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436. [PubMed]
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222. [PubMed]

- Downing, P., & Dodds, C. (2004). Competition in visual working memory for control of search. *Visual Cognition, 11*, 689–703.
- Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1416–1438. [PubMed]
- Druzgal, T. J., & D’Esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *Journal of Cognitive Neuroscience, 15*, 771–784. [PubMed]
- Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: Behavior and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1737–1747. [PubMed]
- Gajewski, D., & Henderson, J. (2002). Minimal memory in a scene comparison task. 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.
- Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research, 38*, 125–137. [PubMed]
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.
- Jha, A. P., Fabian, S. A., & Aguirre, G. K. (2004). The role of prefrontal cortex in resolving distractor interference. *Cognitive, Affective & Behavioral Neuroscience, 4*, 517–527. [PubMed]
- Jolicoeur, P., & Dell’Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*, 138–202. [PubMed]
- Jolicoeur, P., & Dell’Acqua, R. (1999). Attentional and structural constraints on visual encoding. *Psychological Research, 62*, 154–164.
- Jonides, J., Lacey, S. C., & Nee, D. E. (2005). Processes of working memory in mind and brain. *Current Directions in Psychological Science, 14*, 2–5.
- Kim, S. Y., Kim, M. S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 16524–16529. [PubMed] [Article]
- Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General, 133*, 475–493. [PubMed]
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279–281. [PubMed]
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience, 16*, 5154–5167. [PubMed] [Article]
- Miller, E. K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience, 13*, 1460–1478. [PubMed] [Article]
- Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews, Neuroscience, 6*, 97–107. [PubMed]
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442. [PubMed]
- Ranganath, C., DeGutis, J., & D’Esposito, M. (2004). Category-specific modulation of inferior temporal activity during working memory encoding and maintenance. *Brain Research: Cognitive Brain Research, 20*, 37–45. [PubMed]
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207–231.
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 248–261. [PubMed]
- Soto, D., Humphreys, G. W., & Rotshtein, P. (2007). Dissociating the neural mechanisms of memory-based guidance of visual selection. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 17186–17191. [PubMed] [Article]
- Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision, 3*(1):9, 86–94, <http://journalofvision.org/3/1/9/>, doi:10.1167/3.1.9. [PubMed] [Article]
- Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*, 557–561. [PubMed]
- Woodman, G., Boucher, L., Schall, J., & Luck, S. (2004). Do the contents of visual working memory automatically influence attentional selection during visual search? Poster presented at the Society for Neuroscience 2004 Annual Meeting, San Diego, CA.
- Woodman, G. F., & Vogel, E. K. (2005). Fractionating working memory: Consolidation and maintenance are independent processes. *Psychological Science, 16*, 106–113. [PubMed]
- Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science, 12*, 219–224. [PubMed]