Cognition 108 (2008) 719–731

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Visual working memory capacity for objects from different categories: A face-specific maintenance effect

Jason H. Wong *, Matthew S. Peterson, James C. Thompson

Department of Psychology, George Mason University, MS 3F5, 4400 University Drive, Fairfax, VA 22030, USA

ARTICLE INFO

Article history:
Received 31 March 2008
Revised 3 June 2008
Accepted 18 June 2008

Keywords:
Visual memory
Visual attention
Object recognition
Face perception

ABSTRACT

The capacity of visual working memory was examined when complex objects from different categories were remembered. Previous studies have not examined how visual similarity affects object memory, though it has long been known that similar-sounding phonological information interferes with rehearsal in auditory working memory. Here, experiments required memory for two or four objects. Memory capacity was compared between remembering four objects from a single object category to remembering four objects from two different categories. Two-category sets led to increased memory capacity only when upright faces were included. Capacity for face-only sets never exceeded their nonface counterparts, and the advantage for two-category sets when faces were one of the categories disappeared when inverted faces were used. These results suggest that two-category sets which include faces are advantaged in working memory but that faces alone do not lead to a memory capacity advantage.

1. Introduction

Working memory is an important component of cognitive function, as it aids in the storage and manipulation of information. Working memory can be split into visual and verbal components (Baddeley, 1992), and visual memory can be split into memory for locations and memory for objects (Smith et al., 1995). With any memory system, it is important to determine how much information can be held at any one time, and a debate in the literature has focused on the capacity visual object working memory. Luck and Vogel (1997) demonstrated that visual object working memory (VOWM) has a maximum capacity of four objects, even when each object is comprised of multiple features. As with much research in this area, a change detection task paradigm was used. Participants detected the presence or absence of a change between two displays of objects, separated by a delay period. This task allows for the measurement of memory accuracy and capacity (Cowan, 2001). The view that object working memory can hold a maximum of four objects runs parallel with other research showing a similar capacity of four stimuli in other memory domains, such as spatial working memory used in multiple object tracking (Pylyshyn & Storm, 1988) or in verbal working memory (Cowan, 2001).

However, not all research supports the view that visual object working memory stores discrete objects. Alvarez and Cavanagh (2004) demonstrated that fewer objects are stored in memory as the visual information load of each object increases. Visual information load was defined as the per item processing rate during a visual search task, so more complex items took longer to identify during search. However, this effect may be due to the amount of similarity between complex objects (Awh, Barton, & Vogel, 2007). More complex objects often contain more features, which may make comparison difficult.

Object complexity may not be the only factor that influences the capacity of object working memory. Olsson and Poom (2005) argued that capacity can vary significantly depending on the categories of objects in the display. Categories, in this respect, refers to objects that are visually

* Corresponding author. Tel.: +1 703 993 1714; fax: +1 703 993 1330.
E-mail address: jwong1@gmu.edu (J.H. Wong).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2008.06.006

similar. Memory sets from earlier studies consisted either of objects from different categories, such as differently colored blocks, or a small number of similar stimuli, such as Landolt *C*'s. Olsson and Poom (2005) required memory for objects from similar categories (e.g., ovals of various sizes or one shade of various colors) and memory capacity was reduced to only a single object. This one-object memory capacity contradicts the majority of other research in this area and demonstrates the importance of examining object categories in the memory set.

Curby and Gauthier (2007) extended the idea of examining object categories in working memory by demonstrating a memory advantage for faces. Their change detection task included upright and inverted faces along with other object categories. Both set size and encoding duration were varied, and results showed an advantage for upright faces over inverted faces and other complex objects at encoding times of 1500 ms and beyond. This face-specific effect was hypothesized to result from the holistic processing of faces (Tanaka & Sengco, 1997), enabling the features of faces to be 'chunked' into larger units that contain more information than other complex objects and thus increasing storage capacity.

It is clear that categories play a role in the capacity of object working memory, but research has not been performed that examines whether holding objects from multiple categories can affect memory capacity. Previous studies have used objects that were either all dissimilar or all similar. For example, many studies of visual working memory require participants to hold colored blocks (Luck & Vogel, 1997; Vogel & Machizawa, 2004) or lines of different orientations (Luck & Vogel, 1997; Vogel, McCollough, & Machizawa, 2005) in memory, so each object is dissimilar. In contrast, other studies have required memory for similar intracategorical objects such as shaded cubes or Chinese characters (Alvarez & Cavanagh, 2004; Olsson & Poom, 2005; Woodman, Vogel, & Luck, 2001).

It has long been known that phonologically similar sounds are more difficult to rehearse in working memory (Conrad & Hull, 1964) and phonological similarity affects the capacity of the phonological loop. Likewise, semantic similarity affects working memory capacity, with semantically similar items leading to greater interference (Wickens, Dalezman, & Eggemeier, 1976). The current studies examine whether object similarity will affect storage capacity of visual working memory in a similar fashion by using different categories of complex, real-world objects. In order to accomplish this, we examined whether capacity differs between memory sets consisting of objects from a single category (e.g., faces or houses only) and sets consisting of the same number of objects but from different categories (e.g., faces and houses together). Our change detection task consisted of three conditions: two objects from a single category, four objects from a single category, and four objects consisting of two objects from two separate categories. These conditions will allow for the determination of whether holding two-category sets in memory leads to an increase, decrease, or no change in capacity compared to one-category sets.

2. Experiment 1

In Experiment 1, memory for two-category sets was examined using faces and houses. These two categories are ideal for the initial study because evidence suggests that faces and houses might be represented in different neural regions (Kanwisher, McDermott, & Chun, 1997). The critical comparison will lie in the memory capacity of four objects from a single category and four objects from two different categories. The default hypothesis is that object similarity has no effect on capacity. A second hypothesis is that working memory may have to coordinate representations in these separate neural areas instead of one single area, and attentional limitations may lead to decreased performance in the two-category condition compared the one-category condition. Alternatively, a lack of representational overlap between the two categories could lead to an increase in capacity in the two-category condition, possibly due to decreased interference between the two object categories in memory.

2.1. Method

2.1.1. Participants

A total of 24 George Mason University undergraduates participated (6 males, 18 females). The naïve observers received partial class credit in exchange for their participation. The mean age of participants was 19.8 years, and all participants had normal or corrected-to-normal vision.

2.1.2. Apparatus and stimuli

A Power Macintosh G4 (450 MHz) equipped with a 17-in. monitor operating at 85 Hz at a resolution of 1024×768 was used to display stimuli. Participants were seated 60 cm away from the monitor. The system was running custom software in order to present the stimuli, control the timing of experimental events, and record participants' responses.

Two different classes of objects were used in this experiment. The face database was provided by the Max-Planck Institute for Biological Cybernetics in Tuebingen, Germany (Troje & Bülthoff, 1996). Only forward-facing faces were used. Faces were made grayscale and resized to $3.1^\circ \times 3.1^\circ$ of visual angle. A total of 90 faces were used in the experiment. The houses were originally published in Harley, Dillon, and Loftus (2004). Houses were also made grayscale and resized to $3.1^\circ \times 3.1^\circ$. Color information was removed from all objects to increase encoding difficulty and reduce individual distinctiveness. One hundred-and-twenty houses were used. Two sample faces and houses can be seen in Fig. 2.

Each memory display consisted of objects on a black background. The display was divided in twelve positions in a 4×3 grid. Each grid space was $4.0^\circ \times 4.0^\circ$ and could contain one object. The display was constructed by randomly choosing two or four grid spaces and loading objects into those locations. A randomly generated mask that was used between the memory and test array was generated by drawing one hundred white lines between random points on a black background (see Fig. 1).