Title
A Species Difference in Visuospatial Working Memory: Does Language Link “What” with “Where”?

Permalink
https://escholarship.org/uc/item/5wf7b7fw

Journal
International Journal of Comparative Psychology, 20(1)

ISSN
2168-3344

Authors
Washburn, David A
Gulledge, Jonathan P
James, Frances
et al.

Publication Date
2007-12-31

Peer reviewed
A Species Difference in Visuospatial Working Memory: Does Language Link “What” with “Where”?  

David A. Washburn, Jonathan P. Gulledge, Frances James, and Duane M. Rumbaugh  
Georgia State University, U.S.A.

A computerized version of a popular children’s memory game (“Concentration”) was used to test the role of language in visuospatial working memory of humans, apes, and monkeys. Participants were required to find matching pairs of pictures by “flipping over” computer-generated cards, and to remember which images had been seen and where each was hidden. All participants were able to locate the pairs of stimuli, but the nonhuman animals were consistently and significantly worse than the human adults. When humans could not use language, performance declined. When the stimuli were meaningful symbols from the chimpanzees’ language keyboards, performance improved. These data suggest that language provides an important function even in visuospatial working memory, linking “memory for what” with “memory for where.”

The capacity of verbal working memory improves with development, largely because articulation rate increases from childhood into the adult years (Hitch, Halliday & Littler, 1993; Hulme, Thomas, Muir & Lawrence, 1984). However, developmental changes are less clear for visuospatial working, which may benefit from aspects of cognitive maturation other than language development (Hamilton, Coates, & Heffernan, 2003; Pickering, 2001). Indeed, for at least one measure of visuospatial memory there is both anecdotal and scientific evidence indicating that young children perform at least as well as adults (Baker-Ward & Orenstein, 1988; cf. Gellatly, Jones, & Best, 1988; Schumann-Hengsteler, 1996a). Parents frequently report losing to their children on a popular visuospatial memory game (called “Concentration” among many other names). This familiar game requires pairs of cards to be turned over in an effort to locate matching images. If the pictures are not identical, the cards are returned to the face-down position. Thus, players have to remember what images have been revealed and where each is located in the array of cards in order to find the matching pairs most efficiently. The game has proven to be a useful test in research on visuospatial working memory development (Arnold & Mills, 2001; Eskritt, Lee & Donald, 2001; Schumann-Hengsteler, 1996b).

This memory game would also seem to be an ideal test of visuospatial memory for nonhuman primates, given the simplicity of the game, its nonverbal nature, and the fact that human children find it enjoyable. Additionally, the memo-
ry demands of the game can easily be altered by varying the number and nature of pairs of stimuli to be remembered and located. Accordingly, rhesus monkeys were successfully tested on a computerized version of this memory game and were able to find two to six pairs of computer-graphic stimuli on thousands of problems (Washburn & Gulledge, 2002; Washburn, Gulledge & Martin, 2003). Unlike humans however, the monkeys tended to perseverate on errors and as a consequence performed at or even worse than chance on the task. In a series of experiments, we demonstrated that the monkeys did understand that nature of the task; that is, the monkeys were attempting to locate the pairs of matching images, despite their poor overall performance on the task. The animals’ poor performance was not reflective of limits in visual memory (i.e., they could remember what images they had seen) or of limits in spatial memory (i.e., they could remember where they had been). On a computerized test of purely spatial memory (similar to a radial-arm maze task), the monkeys could remember, albeit imperfectly, which locations had versus had not been visited on each problem. However, when the task required visuospatial memory—just remembering “what” or “where” but remembering “what was where”—the animals performed very poorly.

Thus, we have established and replicated a species difference in visuospatial memory performance between humans and macaques. The question for the current study was “Why do monkeys perform so poorly on the game-like task?” In the present experiments, we tested the hypothesis that language provides for human adults and children a means for linking “what” and “where” memories in visuospatial memory, an advantage not available to nonhuman animals. That is, we hypothesized that linguistic coding does provide mnemonic support even for non-verbal, visuospatial memory.

Four experiments were conducted. The methods in each study were very similar; consequently, the studies will be described and reported together.

**Method**

**Participants**

In these experiments, humans and nonhuman primates were tested under a variety of conditions on a computerized version of the memory game “Concentration.” Human participants were undergraduate students who volunteered in partial satisfaction of research-participation requirement. The six adult male rhesus monkeys (Macaca mulatta), two adult female orangutans (Pongo pygmaeus), and three adult chimpanzees (Pan troglodytes; two males and one female) had learned prior to this study to respond to computer-generated stimuli by manipulating a joystick in accordance with the demands of a variety of tasks (Hopkins & Washburn, 2002; Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989; Rumbaugh & Washburn, 2003). The chimpanzees had also participated in years of language training and as a result could communicate with one another and human caretakers using a computerized language keyboard (Rumbaugh, 1977; Savage-Rumbaugh, 1986).

All nonhuman animals were tested in their home cages with continuous access to water. No animal was deprived of food or reduced in body weight for purposes of testing. The macaques were individually housed with continuous visual and vocal access to conspecifics. Each monkey had a dedicated test station that provided access to the task so the monkeys could work *ad libitum*. The two orangutans lived together in a large indoor/outdoor cage. A test station provided access to the task during daytime hours; however, the orangutans shared this one test station. Thus, although observation of the animals confirmed that both contributed approximately equally to the test data, it was not possible to determine which orangutan completed which trial. The chimpanzees also live together as a social group in large indoor/outdoor cages. These animals are separated from one another, with constant visual and vocal access, for brief periods each day for testing. Individual computerized test stations are made available during these test periods.
Apparatus

The nonhuman primates and students were tested with comparable apparatus, a computerized test system similar to the one described by Rumbaugh et al. (1989). Standard MSDOS-based computers were connected to 13-inch monitors were used to present all stimuli and record all responses. The apes and monkeys manipulated a joystick to respond, whereas the humans used the arrow keys on the keyboard. In each case, these keyboard or joystick responses controlled the movements of a computer-generated cursor on the screen. The computer program registered a response whenever this cursor was brought into contact with another computer-generated stimulus. The students received only auditory feedback for correct responses, whereas the nonhuman primates also received nutritive rewards for successes.

Procedure

For the first three experiments, each problem of the memory task began with 2 to 6 pairs of computer-generated “cards” located randomly around a plus-sign cursor (see Figure 1 inset). Moving the cursor (via the joystick or keyboard) into contact with any card caused it to “flip over” revealing an image. The nature of this image varied by experiment, as described below. If matching images were revealed on consecutive cards, auditory feedback and nutritive rewards were delivered. If the images did not match (an error), a buzzing noise was presented as the cards flipped back to conceal the image. Subsequently, the subject was permitted to pick two cards again. Each problem continued in this fashion until all pairs had been located, whereupon a new problem was presented with different randomly selected images concealed under cards in random locations.

Figure 1. Configuration of apparatus for the chimpanzees. Sherman (Pan troglodytes) manipulated a joystick (not visible in the photograph, but located within the port that can be seen near the center of the bottom of this figure) to respond to computer-generated stimuli that were presented on the monitor. The inset (bottom right) shows a typical screen configuration from Experiment 3. Matching lexigrams representing “coffee” and “melon” have been located, and the subject has revealed one image of “chow” and will touch one of the remaining active cells to locate a matching image.
Experiment 1 Procedure. The first study was designed to replicate and extend the findings from previous research with this task (e.g., Washburn & Gulledge, 2002). Human participants (28 undergraduate volunteers (mean age = 20.1 years; 18 females) and all 11 nonhuman primates completed at least 250 problems each of the memory task. Each problem presented 2 to 6 pairs of images, each of which was a clip-art image (see Figure 2, top). To our knowledge, these are the first data collected with a version of the Concentration task for the two ape species. It was important to obtain performance by these nonhuman primates to determine whether the species difference previously reported in visuospatial memory (and described in the Introduction) was limited to macaques. Additionally, these data promised the opportunity for comparison across species of changes in visuospatial memory as a function of differences in brain size and complexity. (Because the language-trained chimpanzees work in shorter daily sessions than do the monkeys and orangutans, the chimpanzees produce fewer trials each day—and of course test sessions with language-trained animals are in great demand. For these reasons, we elected not to administer problems with 5 or 6 pairs, which take a long time to complete, to the chimpanzees in this or subsequent experiments.)

Experiment 2 Procedure. In a follow-up study, each of these same human participants and rhesus monkeys from Experiment 1 completed at least 250 problems with random-dot patterns like the ones at the bottom of Figure 2. The three chimpanzees each completed at least 100 problems with these stimuli. Unlike the easily labeled simple shapes in Experiment 1, these dot patterns were designed to be difficult (albeit not impossible) for the undergraduate students to recode with verbal labels. Thus, it was anticipated that these stimuli would resist linguistic coding and verbal rehearsal and thus that human participants would be required to retain these stimuli as purely visual images—in much the same way that nonhuman primates presumably do irrespective of stimulus. To confirm that the stimuli were evenly discriminable, each of the monkeys and the students also performed 50 matching-to-sample trials with the same dot-pattern images.

Experiment 3 Procedure. In Experiment 3, the monkeys and the three language-trained chimpanzees (but no human participants) each completed 150 problems of the memory task using stimuli like the one in the inset of Figure 1. These stimuli, called lexigrams, are the communicative symbols used on the chimpanzees’ keyboards (see Rumbaugh, 1977; Savage-Rumbaugh, 1986). That is, the chimpanzees can use these symbols accurately to name the objects to which they refer (e.g., if shown an apple, the chimpanzees can pick the lexigram that means “apple”), or can select the appropriate object represented by any of the symbols (e.g., if shown the lexigram for apple, the chimpanzees can select an apple from an array of objects). The lexigrams had no meaning for the monkeys. Each memory-task problem in Experiment 3 consisted of 2, 3, or 4 pairs of these stimuli in random locations.

Experiment 4 Procedure. The fourth experiment involved only undergraduate volunteers (N = 40, mean age = 23.3 years; 18 Caucasians, 17 African-Americans; 5 other or unidentified; 24 females), each of whom completed six 6-pair problems with clip-art stimuli. As with Experiment 2, the goal of this study was to provide a measure of visuospatial memory performance for humans that was uninflated by verbal-memory processes. In the present investigation, the participants were required to engage in articulatory suppression (Baddeley, 1990) throughout three of the problems. The other three Concentration problems were baseline trials like to those administered in Experiment 1. In the articulatory suppression condition, participants were required to repeat a syllable (“the, the, the, the, the, the…” continuously, beginning before the first card was contacted and ending only after the last pair had been located. Articulatory suppression has been used in many studies to prevent verbal recoding or rehearsal of to-be-remembered information (e.g., Frick, 1985) without adding general demands on attention that themselves might disrupt performance (Baddeley, 1997). It occupies the vocal and subvocal speech mechanism that would otherwise service rehearsal and recoding, but unlike effortful distractor tasks (e.g., having participants subtract backward by threes), articulatory suppression does so without competing for attention or adding to memory load.

Results

The results of these four experiments are summarized in Table 1, which shows the mean number of errors as a function of experiment, species, and number of pairs. Chance was determined by Monte Carlo computer simulations which solved the similar problems using random selections, thereby mapping the population-level likelihood of performance at various levels. Binomial tests were used to
determine whether performance differed from chance, and an alpha level of $p < .05$ was used for all statistical tests. In Experiment 1, humans performed significantly better than each of the nonhuman primate species, and rhesus monkeys performed significantly worse than chance, replicating and extending the findings of previous studies (Washburn & Gulledge, 2002; Washburn et al., 2003). As in these previous studies, many of these errors were perseverations, or repetitions of a previous (frequently recent) error. In this experiment, the monkeys averaged 9.83 perseverative errors on 4-pair problems, where perseverations in this case include selection of a specific nonmatching pair of cards more than twice in a problem. In contrast, no human participant repeated a 4-pair error more as many as three times. The apes produced intermediate levels of errors, but only the chimpanzees’ accuracy levels were significantly better than chance. The orangutans averaged 2.70 perseverative errors on 4-pair problems, per the operational definition above. The chimpanzees averaged less than one such perseveration (mean = 0.6) per problem, but were observed on some problems to use a relatively inefficient strategy of searching systematically for the match to a specific card (ignoring other matches revealed along the way while flipping, for example, card 1 and 2, card 1 and 3, card 1 and 4, and so forth). Again, these are the first tests of chimpanzees and orangutans on any version of this memory task.

Table 1

*Mean number of errors on the memory task as a function of experiment (stimulus type), difficulty (number of pairs of images), species, and condition. Chance levels apply to all of the experiments, and were determined by simulating performance on the task using random selection of cards.*

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Species</th>
<th>2 pairs</th>
<th>3 pairs</th>
<th>4 pairs</th>
<th>5 pairs</th>
<th>6 pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Chance</td>
<td>2 errors</td>
<td>6 errors</td>
<td>12 errors</td>
<td>20 errors</td>
<td>30 errors</td>
</tr>
<tr>
<td>1 (clip-art)</td>
<td>Monkey</td>
<td>2.16</td>
<td>7.53</td>
<td>19.24</td>
<td>30.23</td>
<td>40.92</td>
</tr>
<tr>
<td></td>
<td>Orangutan</td>
<td>2.08</td>
<td>4.77</td>
<td>12.91</td>
<td>20.09</td>
<td>37.95</td>
</tr>
<tr>
<td></td>
<td>Chimpanzee</td>
<td>1.38</td>
<td>4.53</td>
<td>10.23</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Human</td>
<td>0.09</td>
<td>0.73</td>
<td>1.46</td>
<td>3.75</td>
<td>5.33</td>
</tr>
<tr>
<td>2 (dot-patters)</td>
<td>Monkey</td>
<td>1.96</td>
<td>5.90</td>
<td>14.89</td>
<td>25.04</td>
<td>33.34</td>
</tr>
<tr>
<td></td>
<td>Chimpanzee</td>
<td>1.61</td>
<td>5.03</td>
<td>11.95</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Human</td>
<td>1.03</td>
<td>3.14</td>
<td>5.97</td>
<td>7.04</td>
<td>14.86</td>
</tr>
<tr>
<td>3 (lexigrams)</td>
<td>Monkey</td>
<td>2.10</td>
<td>7.50</td>
<td>22.00</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Chimpanzee</td>
<td>1.18</td>
<td>2.43</td>
<td>6.48</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>4 (clip art)</td>
<td>Human: Baseline</td>
<td></td>
<td></td>
<td></td>
<td>9.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human: Articulatory Suppression</td>
<td></td>
<td></td>
<td></td>
<td>17.62</td>
<td></td>
</tr>
</tbody>
</table>
In Experiment 2, humans again performed significantly better than macaques. However, performance by humans was reliably worse with the random-dot stimuli than with the clip-art stimuli from Experiment 1. In contrast, the monkeys actually performed better in Experiment 2 than Experiment 1, although still not significantly better than chance. This improvement in performance for the monkeys reflected a decrease in the number and severity of perseverative errors (mean on 4-pair problems = 4.97). Conversely, task performance by humans and by chimpanzees was clearly compromised by the dot-pattern stimuli. Perhaps this was because the stimuli were more difficult to discriminate from one another; however, both the monkeys and the humans were able to match-to-sample with the stimuli with better than 80% accuracy. Thus, it seems clear that both species were capable of distinguishing the dot-pattern stimuli from each other; however, they could not efficiently remember where these random-dot images were located in the array of cards, and thus humans and nonhuman animals made frequent errors. We suggest that the reason that the humans performed relatively poorly in this experiment is because the stimuli resisted verbal recoding (although no data were collected to confirm that humans did not try to name the complex forms), and consequently verbal memory mechanisms could not be brought to bear on the visuospatial task. This interpretation is not inconsistent with the monkeys’ improvement in performance in this experiment, as the (nonlinguistic) monkeys were still not performing at levels better than chance.

In Experiment 3, performance by the monkeys returned to the levels characteristic of the first study, and was both significantly worse than chance and significantly worse than the level observed for chimpanzees. As in Experiment 1, the chimpanzees performed significantly better than chance. Note that the chimpanzees performed reliably better with meaningful lexigrams as stimuli than with the clip-art images that were used in the first experiment.

Experiment 4 data revealed significantly fewer errors in the baseline condition than in the articulatory suppression condition, \( t(37) = 19.41, p < .05 \). On average, 53% of the responses were correct in the baseline condition, but only 36% of the responses were correct when verbal recoding was prevented by concurrent articulation. Performance in both conditions was significantly better than would be expected by chance alone. Note that the higher number of errors in the baseline condition of this experiment relative to the data from Experiment 1 may reflect the fact that only three baseline problems per participant were collected in this study, compared to about 50 6-pair problems per participant in Experiment 1. Accordingly, the Experiment 1 mean is less variable.

General Discussion

Rhesus monkeys have provided a valuable model for neuropsychological studies of human cognition and behavior (see, for example, recent reviews by Easton & Emery, 2005; Farah, 2000; Gazzaniga, 2004; Ghazanfar, 2006; Krasnegor, Lyon & Goldman-Rakic, 1997; Parker, Wilding & Bussey, 2002; Posner, 2004). Our understanding of how the brain-behavior relations that correspond to perception, attention, and particularly memory is immeasurably indebted to monkeys’ ability to perceive, attend, remember, and so forth. Although monkeys are not humans, of course, they have generally proven able to perform a wide variety of human-like cognitive tasks outside of the language domain. Consequently, it seemed reasonable to anticipate that rhesus monkeys would perform well on the present
memory task—inspired as it was on a children’s game that does not require language and on which children can perform comparably to adults (Baker-Ward & Orenstein, 1988; cf. Gellatly, Jones, & Best, 1988; Schumann-Hengsteler, 1996a). Moreover, there is ample evidence that avian species can remember what foods are cached in different locations (Clayton & Dickinson, 1998), and Menzel (1999) demonstrated that a language-trained chimpanzee can recall what foods are hidden and where they are concealed. Surprisingly however, macaques and even great apes did not perform comparably to humans but rather struggled to perform the task better than chance.

Why was this the case? One might suggest that the animals did not understand the nature of the task, but previous analyses undermine this possibility (Washburn & Gulledge, 2002). Perhaps the demands of the task exceed the animals’ memory capacity for visual or spatial stimuli, but again previous studies with a variety of memory tasks (Washburn et al., 2003) impugn this suggestion. We suggest that this memory game is unique in that it requires coordination of the “what” and “where” memory systems, producing memory of “what is where.” It appears that nonhuman primates can certainly remember “what is where” (i.e., they do eventually find the pairs on all problems—without responding randomly—and on many problems they find the matching stimuli with human-like efficiency). However, performance on this task is based predominantly on a familiarity mechanism that is susceptible to perseverative errors. The monkeys solve these problems by looking for previously seen stimuli in previously visited locations, apparently without an efficient mechanism for precisely linking exactly what image is in what location. Consequently, the monkey would occasionally get stuck in a problem, repeatedly making the same error or small group of errors, dozens of times in a row. Once stuck in a perseverative loop, the monkeys frequently had to stop working altogether to break the cycle and complete the problem.

These perseverations were not observed in the data from human participants, although reports of perseverative errors on this memory game have been reported for aged adults (Portman, Feldstein, Davis, & Durham, 1998). Chimpanzees and orangutans were also generally able to inhibit perseverations, so that repetitions of errors were less frequent and when they did occur, they did not last long. Thus, it appears that the species difference frequently observed in the executive function of response inhibition (e.g., Washburn, 1994) may account for some of the between-species variability we observed in memory performance. It is noteworthy that the macaques’ best performance came in Experiment 2, where the number of perseverations declined, presumably because the complexity and similarity of the stimuli compromised the familiarity-based mechanism that could underlie perseverations.

However, the present data suggest another cognitive process as important for visuospatial memory on this task. Typically, human adults and children have language skills that may provide a redundant and efficient means of representing “what was where.” According to Baddeley’s (e.g., 1990) influential theory of working memory, memory for visual or spatial information is retained and manipulated in a separate slave system (called the visuospatial sketchpad) from the memory system used to retain verbal information (called the phonological or articulatory loop). Other researchers have also argued for the separability of verbal from visuospatial memory (Shah & Miyake, 1996). Activities that interfere with verbal
memory tend not to disrupt visual or spatial memory; activities that impede visuospatial memory tend to leave verbal memory relatively intact (Baddeley, 1990). Indeed, further research suggests that visual memory and spatial memory may themselves constitute separate working memory systems (Pickering, Gathercole, & Hall, 2001). However, the present findings suggest greater cross-talk between the working memory subsystems. Whereas memory for what (visual memory) and where (spatial memory) may generally operate separately from verbal storage, the verbal recoding functions of the phonological loop and long-term symbol knowledge appear to be critical for accurate memorial representations of what stimuli are located where spatially.

Thus, when a task requires precise linkage between visual and spatial information, there are benefits to the use of language to sustain memory. When the students could use language to name the stimuli, as in Experiment 1, they made 51% to 65% fewer errors than when the stimuli could not easily be recoded verbally, as in Experiments 4 and 2, respectively. When chimpanzees had linguistic understanding of the symbols, as in Experiment 3, they made 33% fewer errors than when the stimuli were familiar (they were used on repeated problems of this study, and clip art is often used in other studies with the apes) but nonmeaningful clip-art images. These findings are consistent with previous reports that humans perform better on the Concentration game when the images are meaningful symbols (Eskritt et al., 2001). That the monkeys did not also improve with the stimuli in Experiment 3 indicates that lexigrams are not simply easier to discriminate or to remember. That the monkeys’ performances do not improve across thousands of problems of training with the same stimuli (Washburn & Gulledge, 2002; Washburn et al., 2003) indicates that familiarity alone does not make “what is where” easier to remember. Additionally, the disruption in humans’ performance in Experiment 4 is consistent with “language” and not the “familiarity” explanations. Of course, it remains possible that chimpanzees but not monkeys benefit from the familiarity of stimuli, and because meaningful symbols are also inescapably familiar, the present study does not completely eliminate the possibility that familiarity plays a role. It does seem unparsimonious however to suggest that humans and chimpanzees but not monkeys benefit from familiarity. In contrast, it seems clear that some of the images were symbolically meaningful for humans and chimpanzees but not monkeys. Accordingly, we conclude that the meaningfulness of the symbolic stimuli provided critical mnemonic support to visuospatial memory, both for humans and for language-trained nonhuman primates. Language appears to provide an economical means for coding what was hidden where—and monkeys’ performance on the Concentration task apparently suffers for lack of this symbolic representational medium, in combination with the executive-function deficits discussed earlier.

References


