Is visual memory verbal?

Marian Zhen
2839 Bedford Avenue
Brooklyn, NY 11210
Midwood High School Advisor: Alan Stack

Supervising Scientists: Professor Denis Pelli and Katharine Tillman
Professor of Psychology and Neural Science
New York University
Department of Psychology
6 Washington Place, Room 959
New York, NY 10003

M. Zhen (2007) Is visual memory verbal? Intel Science Talent Search.
http://psych.nyu.edu/pelli/highschool.html#2007

ABSTRACT

Is visual memory purely visual, or is it mediated by words? Our ability to remember visual objects may depend on our ability to verbally name them. We tested this idea by looking for an effect of an extraneous verbal task on visual memory. If visual memory is purely visual, there should not be any interference from the verbal task. Instead, we find that having to remember a verbally-presented telephone number impairs visual memory. The effect is as strong on remembering a picture as on remembering a word, the name of the picture. Thus the visual and verbal memory tasks are competing for a common resource. Furthermore, we find that memory performance decreases with the number of syllables in the name of an object, even if observers are only tested with a visual representation of the object. This is strong evidence that visual memory is mediated by verbal naming.

INTRODUCTION

Suppose you witnessed a hit and run. When asked by the police what you saw, what do you say? Was it a Nissan Maxima or a Toyota Camry? Visual memory is the ability of a person to indicate what was seen. Visual memory tasks test how well a person remembers what he has seen. We need to remember many things, everything from phone numbers to names to faces to the directions to a friend's house. We are frustrated when our memory cannot do these simple tasks. However, unlike in the computers we build, memory is scarce and expensive in biological systems in general, and in people in particular. Human visual memory is especially limited. "Change blindness" experiments show that people can recall only a few details from a complex scene. [1](Wolfe, 1998). When asked to remember pictures made of random pixels, human observers can only remember about 8 pixels, in contrast to cheap digital cameras that capture 8 million pixels in each snapshot (Pelli & Farell, 1992).

Faced with the challenges of day-to-day life, people may find that they use many strategies to help increase their memory. Some people think of themselves as being more "spatial," while others consider themselves more "verbal." Indeed, the brain is separated into different areas, and language and vision rely on different areas. One might wonder to what extent these areas help each other. If visual memory is so limited, yet people can still recall what seems like much more than a few pixels, then perhaps in most visual memory tasks, people are using another resource.

If asked to remember a printed word, the 8 pixels of our visual memory could represent at most one character, but people get a whole word per fixation as they read (Pelli, Burns, Farrell, & Moore, 2006). In reading, there is evidence that we convert the visual image to verbal language, letters to phonemes (Rayner, 2001). Here we ask whether this process applies to visual tasks other than reading. Perhaps, whenever the visual objects have meaning, our "visual" memory is mediated by their names. The simple name starts to stand for the complex object, a much more compact representation. Such compression of information can help us to remember more items at once (Miller, 1956).

It has been shown that memory span is affected by the phonological complexity of the words to be remembered (Wilson, 2001). So, following this logic, if pictures are

[1] Change blindness paradigm is when an observer is shown a scene and an object is removed from the scene, the observer is asked to identify the object that was removed.

encoded like words, the number of syllables in their names should have an effect on how many of them we can remember. Here we will compare observers' memory for words and pictures and test the effect of the number of syllables on visual memory. Our goal is to see whether there is a connection between phonemic knowledge and visual memory.

METHODS: EXPERIMENT 1

Remembering a picture and reciting a phone number

Overview

Two tasks were conducted, using a partial-report paradigm. In the first task, the observer was briefly shown a set of objects (either pictures or words) followed by a delay period and then another set of objects in which one of the stimuli shown previously was missing. The observer was required to choose the missing word or picture from a display showing all 17 possible stimuli. The second experiment was identical to the first except that the observer had to perform the task while also retaining a 7-digit phone number read to her by the experimenter at the beginning of each trial.

Participants

The sample consisted of 5 participants ranging in age from 17 to 23 years old. The procedure was explained to the participants before testing. They were recruited through friends. They consented to the experiment and were compensated for their participation.

Apparatus

The program to test the observers was written in MATLAB (Mathworks, Massachusetts) with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997), running on a Macintosh computer.

Stimuli

The stimuli used in this experiment consisted of 17 animal silhouettes ("pictures") based on photographs from a children's book, the 17 corresponding names of those animals, displayed in the Courier font ("words"), and in task two, 40 7-digit phone numbers. The phone numbers were randomized and no phone number was repeated twice in a run.

The pictures or words were distributed along a 3 deg radius (invisible) circle around the fixation point. The diameter of the pictures was 0.8° and the letter size in the words was 1.0°.

Task

Observers were seated 60 cm away from the computer screen. The observers were told to fixate on the dot in the middle of the screen at each trial.

On each trial, the computer program presented either 3, 4, or 5 pictures or words for 500 milliseconds (Fig. 1A,B). Then this screen would disappear and there would be an inter-stimulus pause of 1 second. After this pause, another screen with the same objects would show up, with one object missing (Fig. 1C,D). This screen in turn disappears. On the final screen the observer would see all the possible stimuli. The observer's task was to click on the object that she believed was missing (Fig. 1E,F). If the observer clicked on the correct answer, the computer would produce a high pitched beep. If the observer chose the wrong answer, the computer would produce a lower pitched beep. Then the next trial began. Each run consisted of 40 trials. After the 40th trial, the computer program reported the number correct out of 40. Each observer completed 240 (6 x 40) trials of task 1.

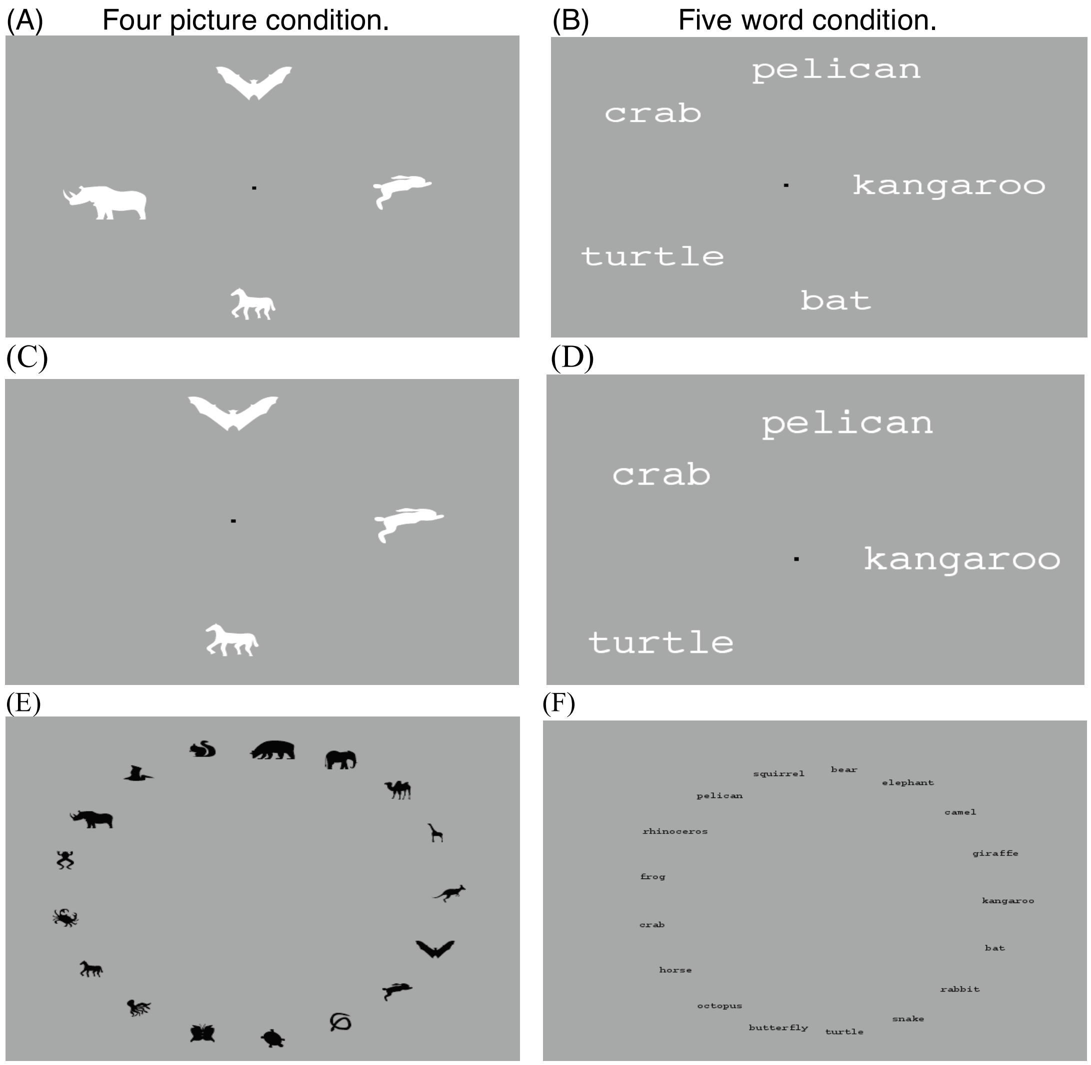


Figure 1: **(A,B)** This is the first screen each observer sees. This screen comes before the inter-stimulus pause. All the objects appear on the screen. **(A)** shows the screen of pictures with four items and **(B)** shows the screen of words with five items. **(C,D)** This shows the screen after the inter-stimulus pause. Note that one object is missing. **(E,F)** This is the final screen that the observer sees. In this screen, all the possible choices of the words or pictures are shown. The observer is asked to choose the item that was missing.

In the telephone number conditions (task 2), a different phone number was recited to the observer at the beginning of each trial (before the picture or word stimuli were displayed). The observer was not allowed to see the telephone number. Thus, the telephone number task was purely auditory and the animal task was purely visual. After each trial, the observer was required to recite the telephone number. Only trials where the observer correctly recited the telephone number and chose the correct word or picture were counted. For each run, percent correct out of 40 trials was measured. Each observer completed 240 (6 x 40) trials of task 2.

All 4 conditions were randomized.

RESULTS: EXPERIMENT 1

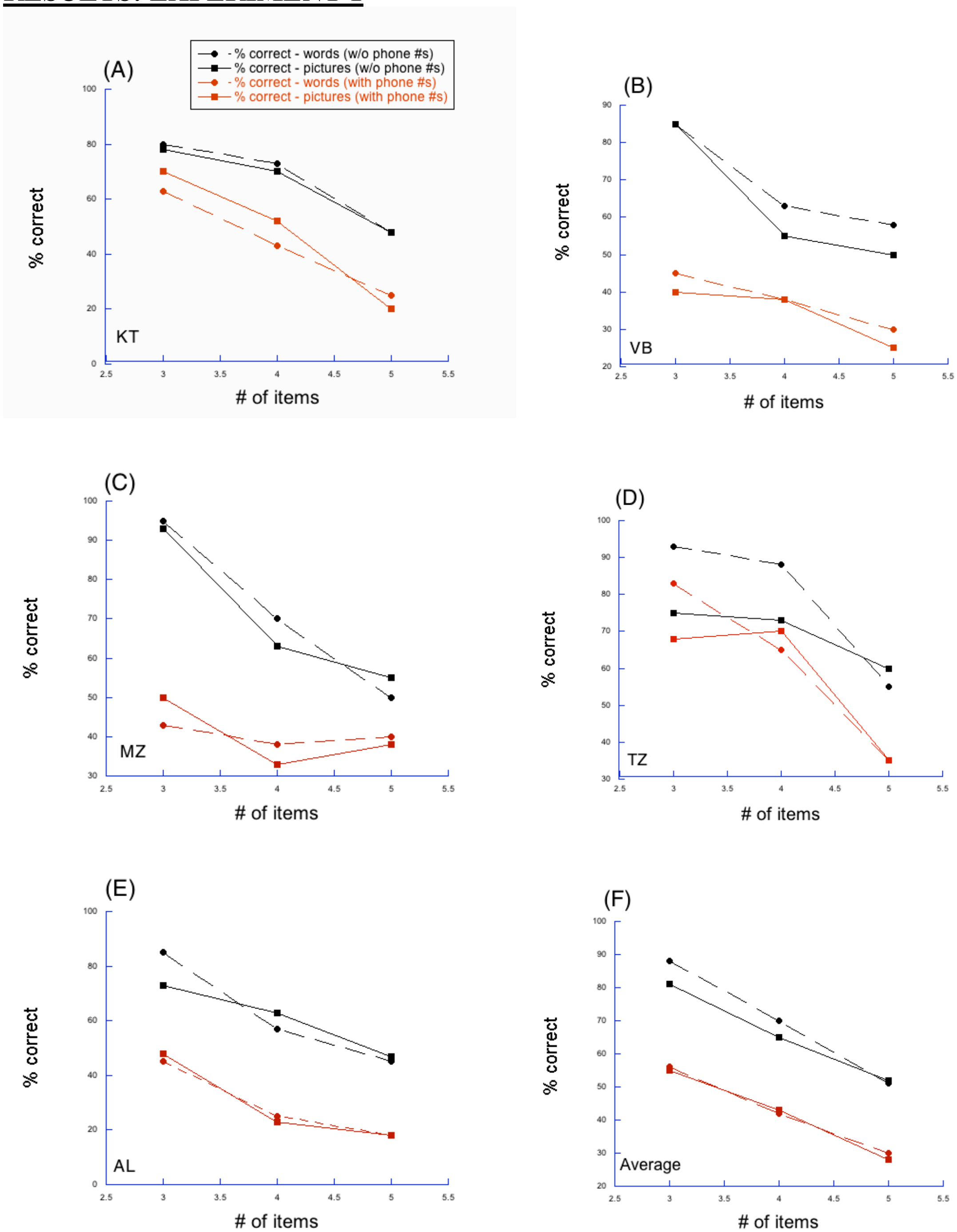


Figure 2: **Visual memory for pictures and words.** Each graph (one for each of the 5 observers) shows the proportion correct with and without phone number recitation.

Figure 2 shows the percent correct as a function of the number of items. Dashed lines represent trials with words. Solid lines represent trials with pictures. Red curves represent trials on which the observer had to remember the phone number, and black curves represent trials on which she did not. In all conditions, percent correct dropped as the number of items presented increased. For observers KT, MZ, VB, and AL, (Fig. 2A-C, E) on both tasks, there is very little difference in performance (less than 10%) between the two stimulus types, suggesting that words and pictures are remembered in the same way. Observer TZ (Fig. 2F) shows more of an advantage for remembering pictures than words in the 3-item condition, but not in the 5-item condition.

When required to remember a phone number while remembering the objects, all the observers did worse in both the picture and word trials. Each observer's visual memory was impaired when verbal stimuli were added to the strictly visual task. The average across the observers (Fig. 2F) shows very little difference between the two stimulus types. However, the average does show a significant decrease (about 30% for words and pictures) when the observers have to remember a phone number while doing the visual memory task. Each observer's visual memory was impaired by the verbal task. This shows that the two tasks are competing for a common memory resource.

METHODS: EXPERIMENT 2

Remembering pictures with short and long names

Overview

We wanted to see if there was an effect of name length on visual memory. We separated pictures by the length of their names and followed the procedure from experiment one, task one, to test observers' memory.

Participants

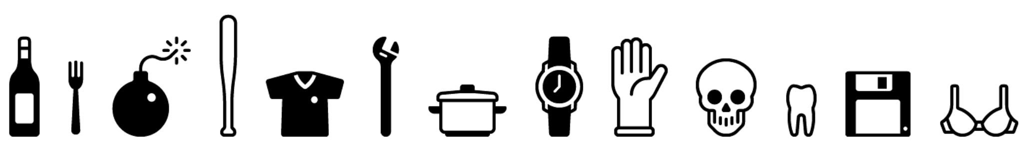
Recruitment techniques were the same as those employed in Experiment 1. There were 4 participants. 3 of the 4 observers participated in Experiment 1. They consented to the experiment and were compensated for their participation. The procedure was explained to the participants before testing.

Apparatus

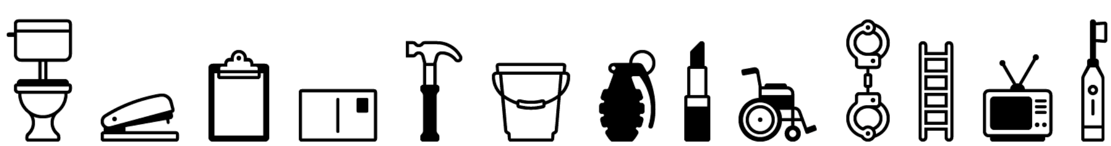
As in Experiment 1, the program to test the observers was written in MATLAB (Mathworks, Massachusetts) with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997), running on a Macintosh computer.

Stimuli

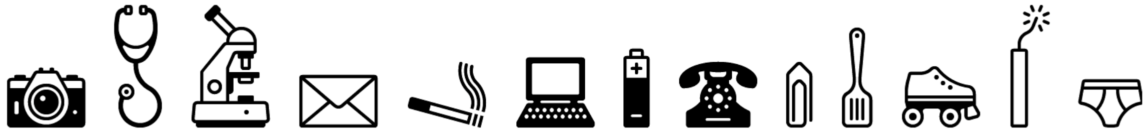
The stimuli used in this experiment were icons sampled from the Poppi font. The pictures were divided into three groups. The groups consisted of pictures with 1-syllable names, 2-syllable names and 3-syllable names. In each group, there were 13 objects:



One syllable: wine, fork, bomb, bat, shirt, wrench, pot, watch, glove, skull, tooth, disk, bra



Two syllables: toilet, stapler, clipboard, postcard, hammer, bucket, grenade, lipstick, wheelchair, handcuffs, ladder, TV, toothbrush



Three syllables: camera, stethoscope, microscope, envelope, cigarette, computer, battery, telephone, paperclip, spatula, roller skate, dynamite, underwear

Task

Participants were asked to follow the same procedures as those assigned to them in Experiment 1, task one. The only difference in Experiment 2 was the stimulus set used; instead of animal pictographs and names, pictures selected from the Poppi font were used in this experiment.

Before running the experiment, observers were shown the stimuli and asked to name them. In cases where items were incorrectly named, the correct name was taught to the observer.

RESULTS: EXPERIMENT 2

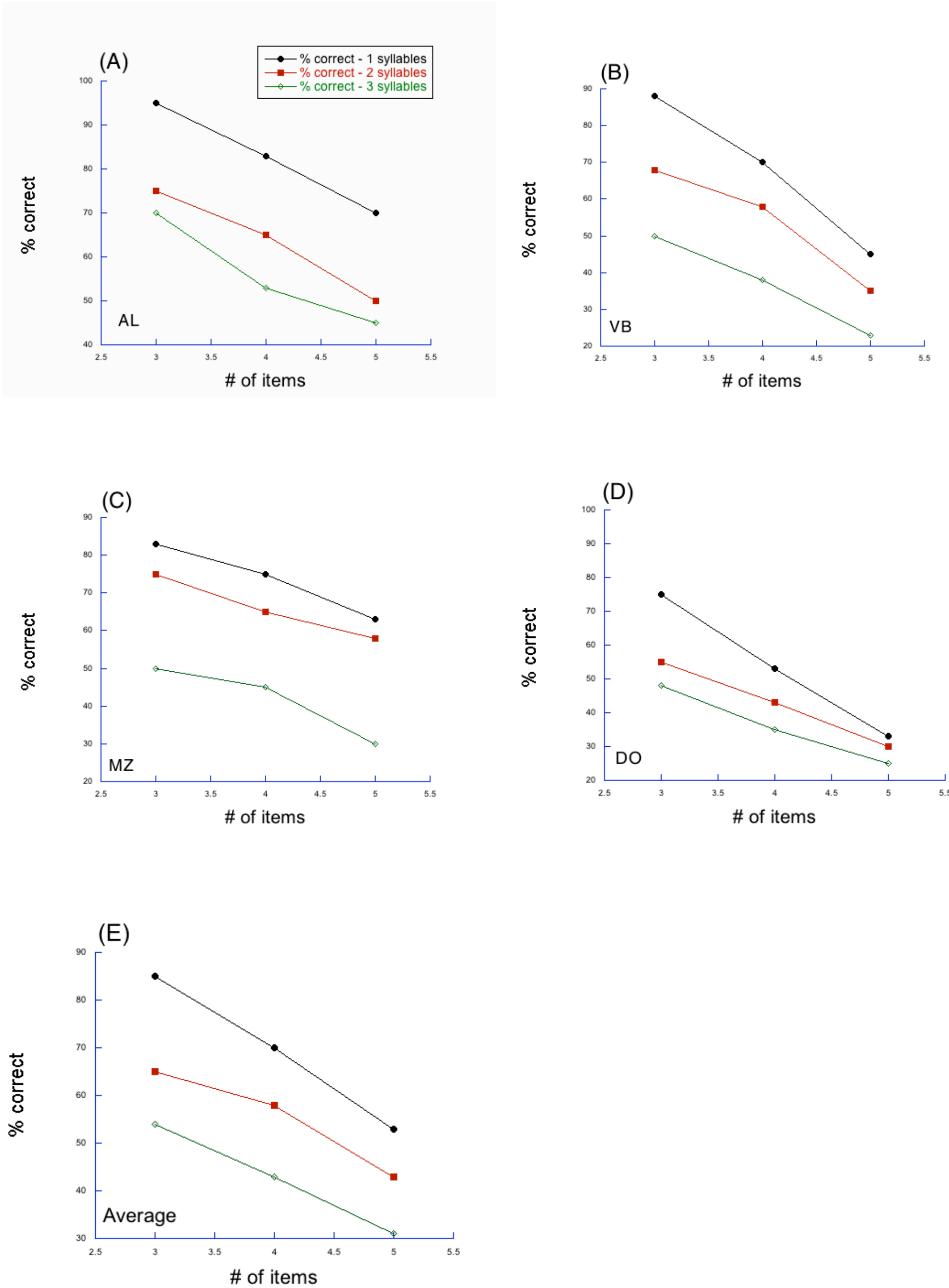


Figure 3: **Visual memory for short and longer words. (A-D)** Each observer's data is displayed. The black line shows items that were 1 syllable, the red line shows items that were 2 syllables, and the green line shows items that were 3 syllables. **(E)** shows the average across all four observers.

M. Zhen (2007) Is visual memory verbal? Intel Science Talent Search.
http://psych.nyu.edu/pelli/highschool.html#2007

Figure 3 plots the effect of name length on visual memory (for pictures). As the names of the pictures increased in number of syllables from 1 to 3, every observer's percent correct decreased. On average, performance decrease from 1- to 3-syllable names was about 30% in the three and four item conditions, and about 20% in the five item condition.

DISCUSSION

The object of this study was to see if there was a connection between visual memory and phonemic knowledge. We know that when people read, they are decoding graphic symbols to speech (Gibson, 1965). You see the letters, but you can "hear" the word. We wondered if, perhaps, people do this in all visual tasks, not just in reading. When we see an object, are we silently naming it to ourselves? Do we only remember the objects that we name?

Figure 2 shows that observers perform similarly on a memory task with picture stimuli as they do on a memory task with word stimuli. This suggests that they might, in fact, be "reading" the pictures: that is, decoding them into the auditory phonemes that make up their names.

Furthermore, requiring observers to hold an additional seven-digit phone number in memory impaired performance on the picture task just as much as the word task. We know that a person's short-term memory capacity is limited to seven plus or minus two items (Miller, 1956). The phone number was read aloud to the observer, so we know that it must be stored verbally (auditorily). Because we see an effect of the phone number task on the "visual memory" task, we know that the two are competing for the same resource. Both tasks appear to be competing for space in our verbal short-term memory. If observers were capable of relying on a nonverbal memory store to do the task, surely they would have done so. We must conclude that if there is a purely visual memory system, it is too weak to be of much help to observers in our task. This finding is in line with previous findings that visual memory can hold a mere 8 bytes of information (Pelli and Farrell, 1992).

If it is true that in order to remember objects in their visual environment, observers must name them, we would expect to find similar phonological effects with pictures than have been found with words. Baddeley (1986) found that when asked to recall lists of five words, observers' performance dropped by 40% as the number of syllables in the words was increased from 1 to 5. Here we show that when asked to remember five pictures, observers' performance drops by 20% as the number of syllables in the names of the items increases from 1 to 3.

If we use the verbal memory route to do these visual memory tasks, it makes sense that objects with longer names would take a longer time to encode. Since we present objects to observers for only a very brief period of time, the longer it takes to encode the names, the fewer of them observers are able to store in memory, and the worse they do on the task. If we were able to store the objects in memory as images, rather than as words, we would not expect the number of syllables in the name to matter. Since our data suggest that observers do encode pictures the same way they encode words, one would think that slow readers would also be slow at encoding pictures. Indeed, studies of dyslexic children show that their memory for and ability to name objects is also impaired (Denckla & Rudel, 1976; Jorm et al., 1986).

CONCLUSION

Remembering a phone number greatly impairs visual memory. This shows that the verbal and visual memory tasks compete for a common resource. It seems very likely that both are competing for space in the language system's verbal store. Furthermore, we find an effect of the number of syllables in the names of the objects to be remembered on observers' performance in a visual memory task. Observers can remember fewer objects when they have longer names. This shows that observers are verbally encoding the images. All of this evidence points toward the same conclusion: visual memory is verbal.

ACKNOWLEDGEMENTS

My deepest thanks go out to Denis Pelli for allowing me to use his lab to conduct my experiments, for providing me with such enriching knowledge and for guiding me through this dynamic project. I would also like to thank Katharine Tillman for helping me with my paper and answering countless questions. My appreciation goes out to Jordan Suchow for writing the MATLAB program for my project. I am thankful for everyone who has attended lab meetings for making them interesting and fun. Sincere gratitude goes out to Mr. Alan Stack for working with me on my paper and for providing help during this rewarding experience. Lastly, I would like to thank all the observers who took the time out of their busy schedules to participate in my experiment.

REFERENCES

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436. http://psychtoolbox.org/

Denckla, M.B. & Rudel, R.G. (1976). Naming of object drawings by dyslexic and other learning disabled children. *Brain and Language,* **3,** 1-15.

Johnston, R. S., Rugg, M. D., & Scott, T. (1987). Phonological similarity effects, memory span and developmental reading disorders: the nature of the relationship. *Br J Psychol*, 78 (Pt 2), 205-211.

Jorm, A.F., Share, D.L., Maclean, R. & Matthews, R. (1986). Cognitive factors at school entry predictive of specific reading retardation and general reading backwardness: A research note. *Journal of Child Psychology and Child Psychiatry,* **27,** 45-54.

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev*, 63(2), 81-97.

Pelli, D.G., Burns, C.W., Farrell, B., & Moore, D.C. (2006), Identifying letters. *Vision Research*, In press.

Pelli, D.G., & Farrell, B. (1992). Visual Memory. Talk presented at ECVP meeting in Pisa, Italy.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychol Sci*, 2(2 Suppl), 31-74.

Torgeson, J, K. (1978). Performance of reading disabled children on serial memory tasks. *Reading Research Quarterly*, 19, 57-87

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychon Bull Rev*, 8(1), 44-57.

Wolfe, J. M. (1998). Visual memory: what do you know about what you saw? *Curr Biol*, 8(9), R303-304.