Impact of Subtle Gaze Direction on Short-Term Spatial Information Recall

Reynold Bailey[*] Rochester Institute of Technology

Ann McNamara[†] Texas A&M

Aaron Costello[‡] Rochester Institute of Technology

Srinivas Sridharan[§] Rochester Institute of Technology

Cindy Grimm[¶] Washington University in St. Louis

Abstract

Contents of Visual Short-Term Memory depend highly on viewer attention. It is possible to influence where attention is allocated using a technique called Subtle Gaze Direction (SGD). SGD combines eye tracking with subtle image-space modulations to guide viewer gaze about a scene. Modulations are terminated before the viewer can scrutinize them with high acuity foveal vision. This approach is preferred to overt techniques that require permanent alterations to images to highlight areas of interest. In our study, participants were asked to recall the location of objects or regions in images. We investigated if using SGD to guide attention to these regions would improve recall. Results showed that the influence of SGD significantly improved accuracy of target count and spatial location recall. This has implications for a wide range of applications including spatial learning in virtual environments as well as image search applications, virtual training and perceptually based rendering.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms;

Keywords: Gaze direction, short-term memory, eye-tracking

Links: DL PDF

Figure 1: *Photograph of memory-recall experiment setup. Participants viewed a sequence of images and were asked to recall the location of specific objects or regions. Participants were guided using a novel gaze manipulation technique called Subtle Gaze Direction which attracts the viewer's attention to the target regions without permanently altering the image. The Subtle Gaze Direction technique uses real-time eye-tracking. In the photograph, the eye-tracking hardware is fixed to the bottom of the screen.*

1 Introduction and Background

Everyday tasks such as retrieving items from specific locations or navigating a familiar route rely on accurate spatial understanding of the environment. Spatial memory enables a variety of species, including humans, to maintain stored information about the location of objects in their environment [Loomis et al. 1993; O'Keefe and Nadel 1978; Wang and Spelke 2002]. Spatial memories of the location, shape, and number of items in a scene become reinforced and more accurate in direct proportion to the amount of time spent in an environment or the frequency with which tasks are completed. Spatial memory becomes critical for tasks that occur in the absence of clear visual information such as navigating through a building in the dark. Humans also rely on spatial memory for simpler tasks such as interacting with images or videos viewed on 2D displays. Over time, for example, people learn menu options or icon locations, the path an object takes through a video sequence, or how to navigate through various levels of video games. This type of spatial (also spatial-temporal for video) learning contributes to the overall understanding of a scene.

A key characteristic of accurate scene interpretation is the amount of information that can be extracted from a scene and retained in memory. Visual memory can be categorized into Visual Short-Term Memory (VSTM), Visual Long-Term Memory (VLTM), and Iconic (sensory) Memory [Palmer 1999]. VSTM is limited in terms of storage capacity but creates representations very rapidly - representations which can then be used to inform ongoing cognitive tasks. On the other hand, VLTM boasts virtually unlimited storage capacity and over time forms detailed representations. For this project we are interested in VSTM, in particular whether/how we can use gaze manipulation to help prioritize the information it acquires in order to better perform a spatial recall task.

It is well known that the content of VSTM is highly dependent on *where the viewer attends* in the scene [Awh et al. 2006]. Jonides [1981] explored the differences between voluntary and involuntary attention shifts and referred to cues which trigger involuntary eye-movements as *pull cues*. In this paper, we utilize a technique called Subtle Gaze Direction (SGD) [Bailey et al. 2009] to provide the pull cues necessary to guide the viewer's focus to specific regions of the display.

Previous studies have shown that guiding attention to relevant regions of a display aids problem solving [Grant and Spivey 2003; Thomas and Lleras 2007; Groen and Noyes 2010]. However, these

[*]e-mail: rjb@cs.rit.edu
[†]e-mail: ann@viz.tamu.edu
[‡]e-mail: atc2436@rit.edu
[§]e-mail: sxs9716@rit.edu
[¶]e-mail: cmg@cse.wustl.edu

studies typically used overt techniques such as visible pulses or permanent alterations to the images to highlight areas of interest. In this paper we explore whether it is possible to improve short-term spatial information recall by subtly manipulating the viewer's gaze using SGD. The SGD technique, which combines real-time eye-tracking with subtle image-space modulation, has minimal impact on viewing experience as it does not change the overall appearance of the scene being viewed. Subtlety is achieved by presenting the modulations only to the low-acuity peripheral regions of the field of view so the viewer is never allowed to scrutinize the modulations. The technique has been shown to be quite fast and accurate: viewers typically attend to target regions within 0.5 seconds of the onset of the modulation and the resulting fixations are typically within a single perceptual span of the target. While this shows that the technique is successful in directing gaze it does not necessarily mean that the viewer fully processed the visual details of those regions or remembered them. To gain a better understanding of the level of visual processing involved, our study explores the impact of SGD on short-term spatial information recall. Our results show that the influence of SGD significantly improved accuracy of target count and spatial location recall for a variety of images.

The remainder of this paper is organized as follows: the design of the experiment to test the impact of SGD on spatial information recall is presented in section 2. Analysis and discussion of the experimental results are presented in section 3. In section 4, the paper concludes with a summary of the contributions and potential avenues of future research.

2 Experimental Design

This section describes an experiment conducted to investigate the impact of SGD on short-term spatial information recall. Participants viewed a randomized sequence of images. Following each image, they were presented with a blank screen and asked to recall the location of specific objects or regions. They were instructed to use the mouse to draw the smallest rectangles that bounded each target object or region. Their input was later analyzed to determine how accurate their short-term spatial recollection was in terms of number of targets, location, and shape.

2.1 Stimuli

Stimuli were presented on a 22 inch widescreen monitor, operating at 60 Hz with a resolution of 1680 x 1050. The stimuli consisted of 28 images (3 training images and 25 test images) compiled from various sources. The images ranged from simple scenes with a few objects to complex scenes with many objects. The complete set of images used in the experiment is shown in Figure 10. The number of objects or regions that the participants were asked to recall for each image ranged from 1 to 9 with each number in the range represented. We used Miller's observation [Miller 1956] that the average human can only hold 7 ± 2 items in working memory to establish the upper limit of 9 for the experiment.

2.2 Participants

30 participants (4 females, 26 males), between the ages of 18 and 35 volunteered to participate in this study. All participants reported normal or corrected-to-normal vision with no color vision abnormalities. Participants were randomly assigned to one of two groups:

- **Static group:** 10 participants were presented with a randomized sequence of the 25 test images without the use of gaze direction. This group served as the control group for the experiment.

- **Gaze-directed group:** 20 participants were presented with a randomized sequence of the 25 test images with gaze-direction turned on. For each image presented to the participants in this group, gaze-direction was randomly selected to either:

  - Direct the viewer's gaze to randomly selected regions of the image away from the correct targets.

    or

  - Direct the viewer's gaze to the correct target regions of the image.

  Counterbalancing was used to ensure that every image appeared equally often in both gaze-directed conditions (i.e. modulation at correct targets and modulation away from correct targets).

This approach facilitates the following comparisons for each image:

- 10 viewers from the static group versus 10 viewers from the gaze-directed group who were correctly guided.

  and

- 10 viewers from the static group versus 10 viewers from the gaze-directed group who were incorrectly guided. i.e. gaze was directed away from the information most pertinent to accurately responding to the subsequent recall task.

We hypothesized that using SGD to guide the viewer's focus to the correct target regions of the image would improve short-term spatial recall accuracy compared to static viewing. Similarly, we hypothesized that using SGD to guide the viewer's focus to incorrect target regions would lower short-term spatial recall accuracy compared to static viewing.

2.3 Procedure

Since we are interested in exploring the impact of SGD on VSTM, each image was shown for only 5 seconds and then replaced with a blank screen. While the blank screen was being displayed, an audio question about the spatial content of the image was played. The questions were recorded in a normal voice by a male native English speaker. Sampling rate for the audio questions was 44.1 kHz. This approach is preferred to displaying the question as text on-screen as this may disrupt the participant's short-term memory of the image [Altmann 2004]. The complete list of questions asked during the experiment is given in Appendix B.

The participants responded to the questions by using the mouse to draw rectangular regions on the blank screen. Three images from the complete set were used in a brief training session to ensure that the participants understood the procedure for completing the experiment. The training images are highlighted in Figure 10. During the training session, participants were encouraged to ask questions and were able to view their solution and the correct solution after each image (see Figure 2). The complete set of instructions read verbatim to each participant is given in Appendix A.

SGD was implemented as described in [Bailey et al. 2009], i.e. viewer gaze was monitored in real time to ensure that modulations were only presented to the peripheral regions of the field of view

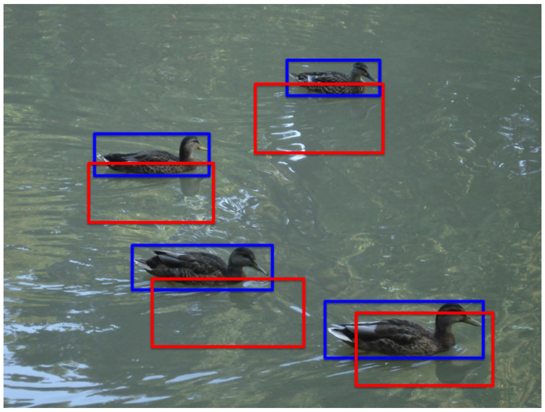

Figure 2: *One participant's solution (red) and correct solution (blue) for an image from the training set. Participants were shown the image for 5 seconds then presented with a blank screen and asked to recall the location of each duck in the image by drawing rectangular regions on the blank screen.*

and modulations were immediately terminated as the viewer's focus approached the modulated regions. The eye-tracker used in this study is a SensoMotoric Instruments iView X Remote Eye-Tracking Device operating at 60 Hz with gaze position accuracy $< 0.5°$. A photograph of the experiment setup is shown in Figure 1.

Assuming that there were n correct solutions and m participant responses for a given image, we define the following measures for spatial recall accuracy:

- **Counting error** is the difference between the number of correct regions in an image and the number of regions submitted by the participants. There is no penalty for incorrect location or shape of the rectangular regions that the participants submit. Counting error is defined as follows:

$$|n - m| \tag{1}$$

- **Location error** is a measure of how close the participant's responses were to the actual targets. There is no penalty for incorrect count or shape of the rectangular regions that the participants submit. Location error is defined as follows:

$$\frac{\sum_1^i \left(\sqrt{(x_{ai} - x_{bi})^2 + (y_{ai} - y_{bi})^2}\right)}{i} \tag{2}$$

  where i is the smaller of n and m and (x_{ai}, y_{ai}) and (x_{bi}, y_{bi}) represent the centroids of the ith closest pair of rectangles chosen from the set of actual solutions and participant solutions.

- **Shape error** is a measure of how different the widths and heights are of the actual solutions and the rectangular regions that the participants submit. There is no penalty for incorrect count or location. Shape error is defined as follows:

$$\frac{\sum_1^i (|width_{ai} - width_{bi}| + |height_{ai} - height_{bi}|)}{i} \tag{3}$$

  where i is the smaller of n and m and $width_{ai}$ and $width_{bi}$ and $height_{ai}$ and $height_{bi}$ are the widths and heights of the ith closest pair of rectangles chosen from the set of actual solutions and participant solutions.

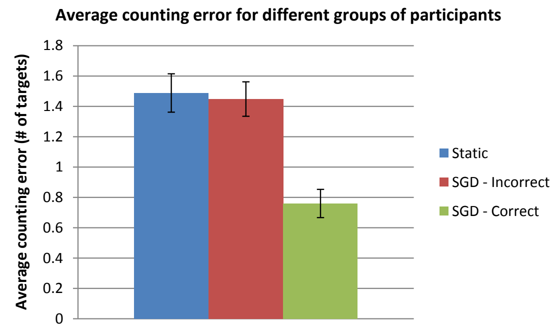

Figure 3: *Average counting error for participants from the static group and gaze directed group.*

Location error and shape error are undefined in cases where the participants did not submit any solutions. This occurred for 15 of the 250 data points for the static condition, 13 or the 250 data points for the SGD at incorrect locations condition and 6 of the 250 data points for the SGD at correct locations condition. These data points were not included in the computation of location error and shape error.

3 Results and Discussion

We measured the impact of SGD on short-term spatial information recall by computing the participants' counting error, location error, and shape error. In summary, we observed the following effects:

- SGD to correct targets results in a significantly lower counting error compared to static viewing
- SGD to correct targets results in a significantly lower location error compared to static viewing
- SGD to incorrect targets results in a significantly higher location error compared to static viewing
- SGD has no significant impact on shape error

Following the experiment, participants were asked informally if they noticed anything strange in the images. Several participants noted that they were able to detect subtle modulations in their peripheral field of view but were never able to focus on them. A few participants even realized that the modulations in some cases guided them to the correct regions and in other cases to incorrect regions. Their strategies ranged from trying to learn if the modulations were correct or incorrect to trying to ignore them altogether. In future studies we will consider formalizing this process in the form of an exit interview to fully determine the noticeability of the modulations.

3.1 Counting Error

Figure 3 summarizes the average counting error for the participants from each group. Counting error for the static group averaged 1.488 targets while the averages for the gaze-directed group were 1.448 targets (modulations at incorrect locations) and 0.76 targets (modulations at correct locations). These values were obtained by averaging the counting error for all participants in a group over all images.

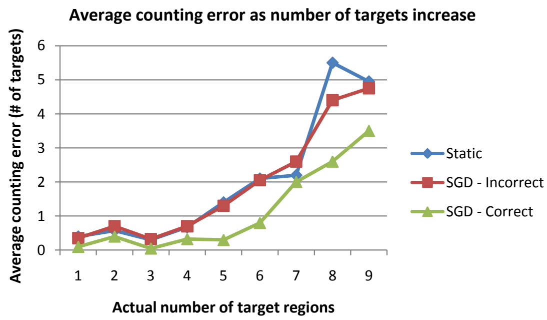

Figure 4: *Average counting error for participants from the static group and gaze directed group as number of target regions increase.*

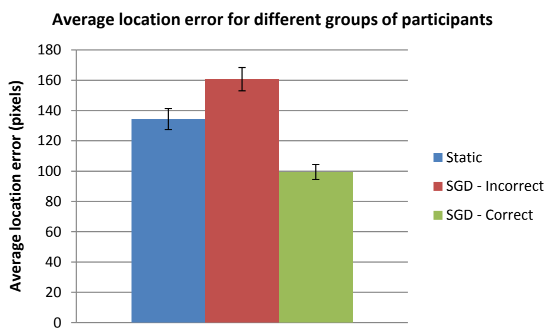

Figure 5: *Average location error in pixels for participants from the static group and gaze directed group.*

The differences in the averages show that SGD to correct locations results in a lower counting error compared to static viewing. An independent-samples t-test revealed that this effect was significant and not due to chance:

$$t(498) = 4.640; p < 0.05$$

Figure 4 shows how the average counting error varies as the number of regions the participants were asked to recall increases. As expected, the counting error increases as the recall task becomes more difficult. Note, however, that the counting error for SGD to correct targets is consistently lower than that of static viewing and that of SGD to incorrect targets.

3.2 Location Error

Figure 5 summarizes the average location error for the participants from each group. Location error for the static group averaged 134 pixels while the averages for the gaze directed group were 161 pixels (modulations at incorrect locations) and 99 pixels (modulations at correct locations). These values were obtained by averaging the location error for all participants in a group over all images.

The differences in the averages show that SGD to correct locations results in a lower location error compared to static viewing. An independent-samples t-test revealed that this effect was significant and not due to chance:

$$t(477) = 4.123; p < 0.05$$

The differences in the averages also show that SGD to incorrect locations results in a higher location error compared to static viewing.

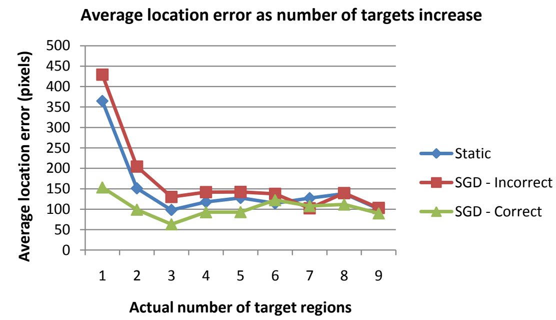

Figure 6: *Average location error for participants from the static group and gaze directed group as number of target regions increase.*

Figure 7: *An image from the test set that was potentially ambiguous. Several participants confused squashes and peppers possibly due to lack of familiarity. This resulted in a spike in location error for this image.*

An independent-samples t-test revealed that this effect was also significant and not due to chance:

$$t(470) = -2.526; p < 0.05$$

Figure 6 shows how the average location error varies as the number of regions the participants were asked to recall increases. Ignoring the outlier (actual number of targets = 1), the plot shows that location error remains fairly constant as the number of targets increase with the location error for SGD to correct locations generally being the lowest and location error for SGD to incorrect locations generally being the highest.

Closer examination of the outlier revealed that the irregularity was due to a potentially ambiguous trial in our test set. Figure 7 shows the image in question. Participants were asked to recall the location of each pile of yellow peppers. We used the word “each” for every question in the experiment to avoid implying the correct count. We observed that several participants selected the yellow squashes instead of the yellow peppers possibly due to lack of familiarity. This led to a spike in the location error associated with this image. It is interesting to note however, that SGD to the correct location seems to have helped to resolve this ambiguity for the gaze-directed group.

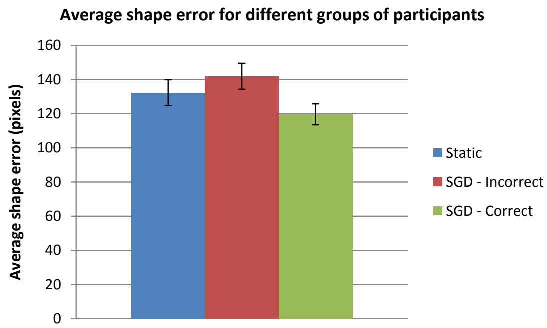

Figure 8: *Average shape error in pixels for participants from the static group and gaze directed group.*

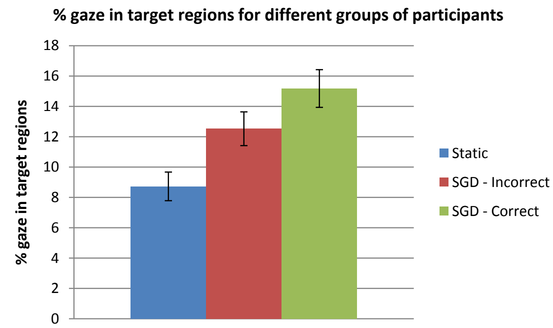

Figure 9: *Percentage gaze time spent within the target regions for participants from the static group and gaze-directed group.*

3.3 Shape Error

Figure 8 summarizes the average shape error for the participants from each group. Shape error for the static group averaged 132 pixels while the averages for the gaze directed group were 142 pixels (modulations at incorrect locations) and 120 pixels (modulations at correct locations). These values were obtained by averaging the shape error for all participants in a group over all images.

The differences in the averages suggest that SGD to correct locations results in a lower shape error compared to static viewing. However, an independent-samples t-test revealed that this effect was not significant:

$$t(477) = 1.310; p > 0.05$$

In this experiment, SGD was used to direct gaze to locations that were clearly within the boundaries of the target regions or objects. We speculate that using SGD to also guide viewer gaze along the bounding box of the target may further improve short-term shape recall.

No clear trend was observed in the average shape error as the number of target regions increased.

3.4 Percentage Gaze Time

We have observed that SGD to correct image locations significantly improves accuracy of target count and spatial location recall. To gain a better understanding of why these effects occur we examine how the percentage gaze time spent within the target regions is affected by SGD. Figure 9 shows the percentage of total gaze time spent within the target regions for the different groups of participants. For the static group, 8.7% of gaze time was spent within the target regions. For the gaze-directed group, 12.5% (which represents a 43.7% increase) and 15.2% (which represents a 74.7% increase) of the gaze time was spent within the target regions for modulations at incorrect locations and modulations at correct locations respectively.

The differences in the averages show that SGD to correct locations results in more gaze time spent within the target regions compared to static viewing. This observation was expected and an independent-samples t-test confirms that the difference in gaze time in the target regions was significant and not due to chance:

$$t(498) = -4.135; p < 0.05$$

Interestingly however, the differences in the averages also show that SGD to incorrect locations results in more gaze time spent within the target regions compared to static viewing. An independent-samples t-test also confirms that the difference in gaze time in the target regions was significant and not due to chance:

$$t(498) = -2.605; p < 0.05$$

This observation was somewhat unexpected. We speculate that the increase in gaze time in the target regions is due to the fact that SGD to various incorrect locations helps to distribute the viewer's gaze more evenly across the image. This causes more of the viewer's scan-paths to intersect with the target regions. A similar observation was also made in another study involving SGD [McNamara et al. 2009]. The researchers noted that the presence of "distractors" (i.e. modulations at incorrect regions) helped to spread the viewer's gaze across images and led to improved performance on a search task compared to static viewing. We plan to explore this further in future experiments.

For this study, the fact that both correct and incorrect modulations resulted in increased gaze time in the target regions but only correct modulations led to improved accuracy of target count and spatial location recall seems to suggest that the modulations themselves serve as triggers for retaining information in our short-term memory.

4 Conclusions and Future Work

We presented an experiment designed to test the impact of the SGD technique on short-term spatial information recall. The results indicate that SGD significantly improves the accuracy of target count and target location recall. The effect was observed on a wide variety of images ranging from simple scenes with a few target regions to complex scenes with many target regions. This has implications for a wide range of applications including:

- **Training Simulations and Virtual Environments:** Training simulations often involve complex, cluttered environments. Key to the success of these simulations is the ability to rapidly attend to targets and extract relevant information from the abundance of information available. Pairing SGD with simulations could facilitate faster responses to those spatial locations deemed more significant for status understanding. For example in Air Traffic Control (ATC) systems SGD could help improve short-term spatial recall of aircraft imminent for arrival.

- **Education and Learning:** Often the amount of information presented to students in educational settings, be it in the classroom on over distance education, is overwhelming.

Using SGD to localize the important information first could help the student build a better cognitive map of the scene they are exploring, thus opening potential for an enhanced learning experience. This could also benefit those that have difficulty paying attention during learning.

- **Gaming:** When environments change rapidly and successful game play depends on successful navigation, which in turn depends on constructing an accurate spatial map of the environment, SGD could be employed to help game players to rapidly build a spatial map prioritizing the most important game features including enemy or friendly targets, and portals to new levels.

In addition to these applications, there are several other avenues for future work. There is significant evidence in psychological literature which suggests that eye-movements triggered by spoken expressions tend to result in fixations on the locations of previously viewed items referred to in the expression [Cooper 1974; Tanenhaus et al. 1995; Kamide et al. 2003]. This is true whether or not the item is still present in the scene [Brandt and Stark 1997; Spivey and Geng 2001; Altmann 2004; Johansson 2006]. These observations have led some researchers to suggest that oucculomotor position serves as an index into spatial memory [Laeng and Teodorescu 2002; Spivey et al. 2004; Ferreira et al. 2008]. We plan to conduct further analysis of the eye-tracking data recorded during this experiment to reveal if the fixations immediately following the question corresponded to the correct target regions and moreover if these fixations correlate well with the actual solutions submitted using the mouse. Additional analysis of the data may also reveal if SGD has an impact on viewer response time i.e. how quickly the task was completed.

Other natural extensions of this study include exploring the impact of SGD on spatial information recall for dynamic environments or videos and investigating the effects of SGD on Visual Long-Term Memory (VLTM).

Acknowledgements

This material is based on work supported be the National Science Foundation under Award No. IIS-0952631. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

A. Instructions to participants (read verbatim)

In addition to the following instructions which were read to the participants at the start of the study, participants were also provided with documentation showing that the study was reviewed and approved by the Institutional Review Board (IRB) at the institution where this study was conducted.

> The purpose of this study is to explore the use of various display techniques to improve spatial memory recall. You will be shown a sequence of images. Each image will be displayed for 5 seconds followed by a blank screen. While the blank screen is being displayed, you will hear a question asking you to recall the location of specific objects or regions of the image.
>
> Please wait for the cursor to change to a crosshair before entering your solutions. Use the left mouse button to click and drag a box to specify where you recall seeing the object. You may repeat this process to select multiple objects or regions. Right click when you are finished. You should use the smallest box possible to specify an object's location.
>
> The images that you will be viewing consist of photographs and computer generated scenes and range from very simple scenes with a few objects to very complex scenes with many objects. If you do not recall seeing a particular object you should not select any regions. Simply right click to move on to the next question. If an object was not entirely visible, you should select only the region that corresponds to the visible part.
>
> The first three images you see will be training images. Feel free to ask any questions while training is in progress. For the training images, you will be able to see your solution (in red) and the correct solution (in blue) after each image. Once training is complete and the experiment begins you will not be able to view your solutions until the end of the entire experiment. When the experiment is complete, your solution (in red) and the correct solution (in blue) for each image will be displayed as a slide show.
>
> During the course of the experiment, a non-invasive camera will be used to record your eye-movements. Please try to minimize your head movements as this may adversely affect the quality of the results. A short calibration process is necessary to ensure that your eyes are being accurately tracked. This will occur at the start of the experiment. Calibration simply involves looking at the targets on the screen until they disappear. The entire experiment should take no longer than 15 minutes to complete.
>
> The results of this study may be published in scientific research journals or presented at professional conferences. However, your name and identity will not be revealed and your record will remain anonymous. Your name will not be used in any data collection, so it will be impossible to tell your answers from others.
>
> The potential benefits of this study to society include improvements in display techniques for digital images and the advancement of scientific knowledge of human visual perception.
>
> Participation is entirely voluntary. Additionally, you may choose to withdraw from this study at any time. If you decide not to participate or to withdraw from this study, there will not be a penalty to you. Do you have any questions before we begin?

B. List of audio questions for each image

The following list contains the spatial recall questions that were asked for each image used in the experiment. For convenience the order of the list follows the layout of the images shown in Figure 10:

- Where was each occurrence of the word "happy" located?
- Where was each dark chocolate cookie located?
- Where was each fish with white stripes located?
- Where was each green can located?
- Where was each player holding a bat located?
- Where was each triangle located?
- Where was the yellow center region of each flower located?
- Where was each female located?
- Where was each person located?

- Where was the head of each child located?
- Where was each napkin located?
- Where was each orange shape located?
- Where was each pile of yellow peppers located?
- Where was each red block located?
- Where was each ball on the pool table located?
- Where was each shoe located?
- Where was each up arrow located?
- Where was each person located?
- Where was each kayak located?
- Where was each player wearing a green shirt located?
- Where was each drum located?
- Where was each rock located?
- Where was each image that contained water located?
- Where was each of the man's fingertips located?
- Where was each diamond shape located?
- Where was each duck located?
- Where was each player wearing a white shirt located?
- Where was the mouth of each child located?

References

ALTMANN, G. T. M. 2004. Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition 93*, B79–B87.

AWH, E., VOGEL, E., AND OH, S. 2006. Interactions between attention and working memory. *Neuroscience 139*, 1, 201–208.

BAILEY, R., MCNAMARA, A., SUDARSANAM, N., AND GRIMM, C. 2009. Subtle gaze direction. *ACM Trans. Graph. 28*, 4, 1–14.

BRANDT, S. A., AND STARK, L. W. 1997. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J. Cognitive Neuroscience 9*, 1, 27–38.

COOPER, R. M. 1974. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology 6*, 84–107.

FERREIRA, F., APEL, J., AND HENDERSON, J. M. 2008. Taking a new look at looking at nothing. *Trends in cognitive sciences 12*, 11 (November), 405–410.

GOOGLE INC. Google™ Image Search. http://images.google.com/.

GRANT, E., AND SPIVEY, M. J. 2003. Eye movements and problem solving: guiding attention guides thought. *Psychological Science 14*, 5, 462–466.

GROEN, M., AND NOYES, J. 2010. Solving problems: How can guidance concerning task-relevancy be provided? *Comput. Hum. Behav. 26* (November), 1318–1326.

JOHANSSON, R. 2006. Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science 30*, 1053–1079.

JONIDES, J. 1981. *Voluntary versus automatic control over the mind?s eye's movement*, vol. 9. Erlbaum, 187203.

KAMIDE, Y., ALTMANN, G. T., AND HAYWOOD, S. L. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language 49*, 1 (July), 133–156.

LAENG, B., AND TEODORESCU, D.-S. 2002. Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science 26*, 2, 207–231.

LOOMIS, J. M., KLATZKY, R. L., GOLLEDGE, R. G., CICINELLI, J. G., PELLEGRINO, J. W., AND FRY, P. A. 1993. Nonvisual navigation by blind and sighted: assessment of path integration ability. *J Exp Psychol Gen 122*, 1, 73–91.

MCNAMARA, A., BAILEY, R., AND GRIMM, C. 2009. Search task performance using subtle gaze direction with the presence of distractions. *ACM Trans. Appl. Percept. 6*, 3, 1–19.

MILLER, G. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review 63*, 81–97.

O'KEEFE, J., AND NADEL, L. 1978. *The hippocampus as a cognitive map*. Clarendon Press ; Oxford University Press, Oxford : New York :.

PALMER, S. E. 1999. *Vision science : photons to phenomenology*. MIT Press, Cambridge, Mass. Stephen E. Palmer.; "A Bradford book."; Bibliography: Includes bibliographical references (p. [737]-769) and indexes.

SPIVEY, M. J., AND GENG, J. J. 2001. Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research 65*, 4 (November), 235–241.

SPIVEY, M. J., RICHARDSON, D. C., AND FITNEVA, S. A. 2004. *Thinking outside the brain: Spatial indices to visual and linguistic information*. Psychology Press, 161–189.

TANENHAUS, M., SPIVEY-KNOWLTON, M., EBERHARD, K., AND SEDIVY, J. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science 268*, 5217, 1632–1634.

THOMAS, L., AND LLERAS, A. 2007. Moving eyes and moving thought: on the spatial compatibility between eye movements and cognition. *Psychonomic bulletin and review 14*, 4 (August), 663–668.

WANG, R. F., AND SPELKE, E. S. 2002. Human spatial representation: insights from animals. *Trends in Cognitive Sciences 6*, 9, 376 – 382.

YAHOO! INC. flickr™. http://www.flickr.com/.

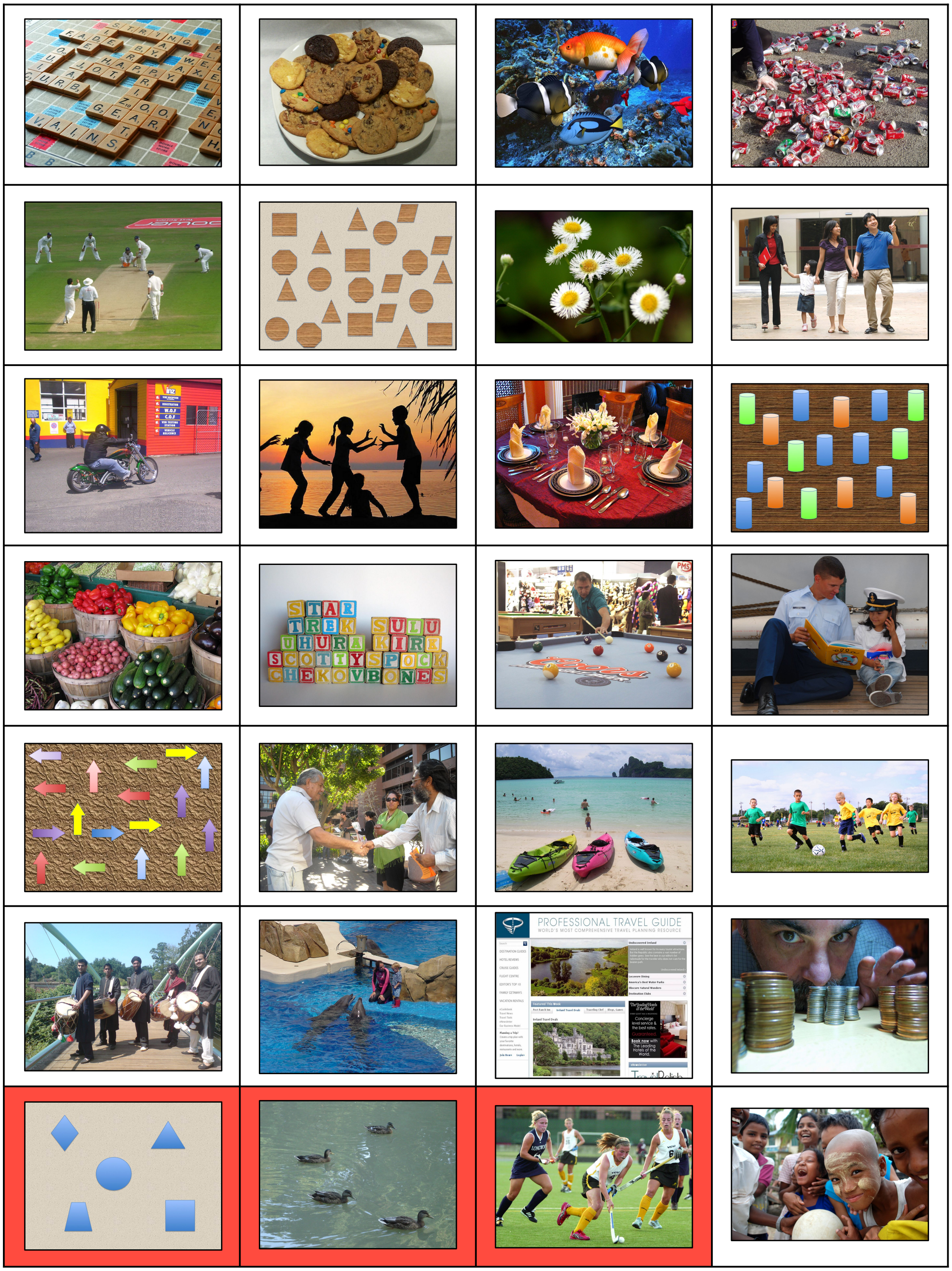

Figure 10: *The complete set of images used in this study. The images depict a wide variety of scenes and were gathered from various sources on the web [Yahoo! Inc.] [Google Inc.] or created by the researchers. The three images that were used in the tutorial are bordered in red.*