HOW DOES THE BRAIN TRACK MULTIPLE OBJECTS?

By

Amanda Skoranski

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Bachelor of Science in Psychology with Distinction.

Spring 2009

HOW DOES THE BRAIN TRACK MULTIPLE OBJECTS?

By

Amanda Skoranski

Approved: ____
James Hoffman, PhD.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: ____
Steven Most, PhD.
Committee member from the Department of Psychology

Approved: ____
Susan Hall, PhD.
Committee member from the Board of Senior Thesis Readers

Approved: ____
Ismat Shah, Ph.D.
Chair of the University Committee on Student and Faculty Honors

ACKNOWLEDGMENTS

I wish to thank everyone involved with the Visual Cognition Lab at University of Delaware, especially Matt Doran and Dave McCoy, who helped pick up my slack.

I also wish to thank my thesis committee, Steve Most and Susan Hall. I would like to extend a special thanks to my thesis advisor, Jim Hoffman, whose support was crucial to this achievement and immensely appreciated.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

How does the brain allocate attention to a visual display? Do they pay attention more to the visual space or to the objects within the display? Questions like these can be answered by using the multiple object tracking (MOT) paradigm. This task requires observers to track target objects in an array as they move amongst identical distracter objects. In addition, we can use psychophysiological measures while participants complete this task, such as event related potentials (ERP). These potentials represent brain activity in response to a particular stimulus (in this case, tracking). Drew & Vogel (2008) used this technique and found that as the participants tracked more targets, ERP activation increased. We replicated this experiment and added another variable - spacing; objects could either be spaced far apart or close together. We found that while left hemisphere activation is sensitive to increasing targets (as in Drew & Vogel, 2008), the right hemisphere activation is sensitive to the spacing manipulation. Therefore, our results support a dual-system hypothesis; observers use both spatial and object-based attention to complete the task.

Chapter 1

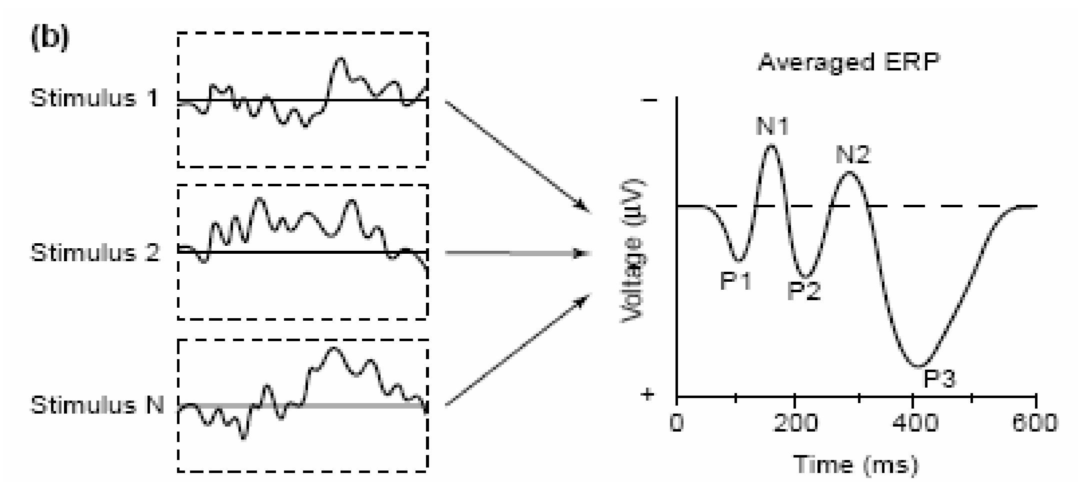
INTRODUCTION

1.1 Event-Related Potentials

As we advance through the 21st century, cognitive science has adapted new efficient methods of data collection. Brain-imaging methods, such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG), supply important insight into the workings of our brains. These different imaging methodologies have complementary strengths and weaknesses. fMRI, which is based on blood flow in the brain, provides precise information about the location of activation in the brain but has poor temporal resolution. In contrast, measures of electrical activity such as EEG provide millisecond temporal resolution but can only provide coarse information about brain location.

EEG is a measure of the combined activity of billions of neurons in the brain and this makes it difficult to discern activity arising from specific areas, such as the visual system. The area of interest is generating a signal that is buried in the general background noise of the brain. A solution to this problem comes from the calculation of event related potentials (ERP). ERPs are changes in electric potential measured at the scalp that are triggered by particular events such as presentation of a stimulus. In order to calculate ERPs, the stimulus must be presented many times while EEG readings are being taken. The EEG waves following the presentation of the stimulus are then averaged together to eliminate the noise, thereby enhancing the signal of interest. ERPs provide a measure of brain electrical activity that is uniquely associated with a particular stimulus (see figure 1.1).

Figure 1.1 **Analyzing ERPs:** From Luck, Woodman, and Vogel (2000); shows an example of how a number of rough EEG waves can be averaged to form a clean ERP curve.

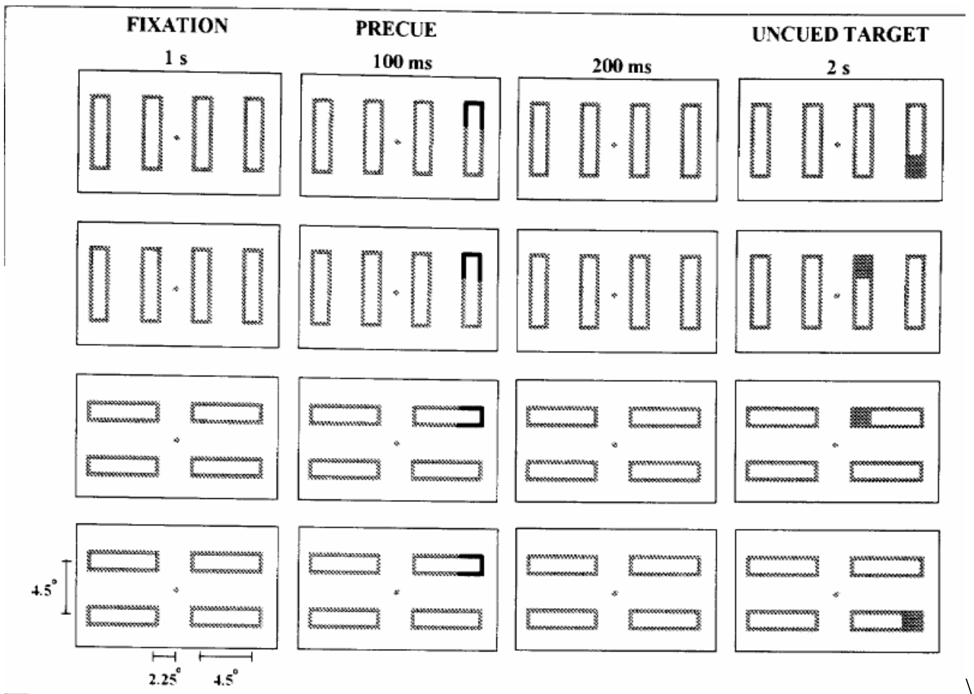


1.2 Theories of Attention Allocation: Object vs. Spatial

In the present experiment, we used ERPs to examine the process of visual attention. The traditional view of allocating attention has been the ‘spotlight’ or ‘zoom lens’ theory which assume that attention is limited to a certain spatial area. Recently, however, there has been a debate over whether attention to objects may require different attention mechanisms than attention to spatial locations. For example, Scholl (2008) argues that, in some tasks (including the MOT task to be discussed later), attention is limited by the number of objects that need to be attended rather than spatial area.

It might also be that both spatial and object attention are at work, and these systems operate exclusively. This idea is supported by the work of Egly, Rafal, Driver, and Starrveveld (1994) who examined attentional allocation in split-brain patients. Observers in this study viewed a display containing two rectangles, one on each side of the fixation point. Their task was to rapidly press a key as soon as they detected a signal which was a monentary brightening of one end of a rectangle. Prior to this event, they were cued as to the likely location of the signal. On valid trials, they light occurred in the cued location and on invalid trials, it occurred in a different location than indicated by the cue. There were two kinds of invalid trials. On same-object invalid trials, the brightening occured in the cued object but in the opposite end from the cue. On different-object invalid trials, the brightening occurred in the uncued object. Importantly, in both of these invalid trials, the signal was the same distance from the cue and thus any differences in the speed of detection could not be attrubuted to spatial distance (see figure 1.2).

Figure 1.2 **Egly, et al. (1994) Paradigm:** This figure is a depiction of the possible invalid trials presented in the Egly, et al. (1994) experiment. The target could be flashed either in the same box as the invalid cue or in an equidistant second box.



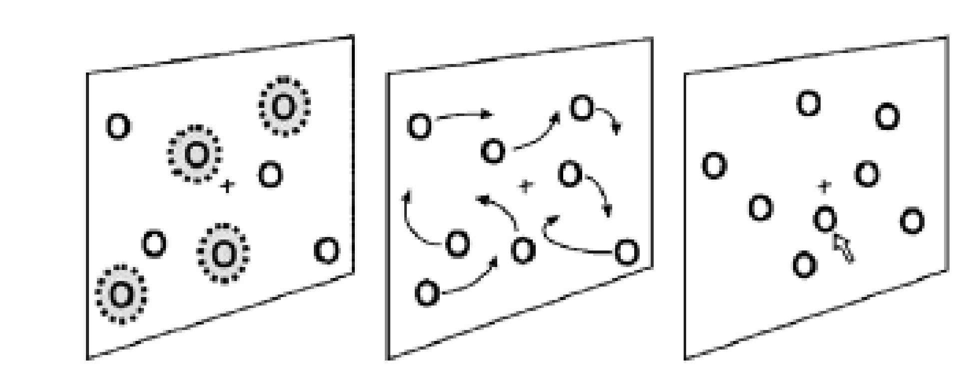
When the task was presented to the right visual field (and thus processed by the left hemisphere), the split-brain patients had similar reaction time differences as a normal participant – same-object invalid trials were faster than different-object invalid trials showing that it is easier to switch to a location within an object than to a location on a different object. However, when stimuli were presented in the left visual field (and processed by the right hemisphere), there was no difference in the RTs to same and different object conditions. Therefore, it appears that not only are the spatial and object-based attention systems different, but that the object-based attention is localized in the left hemisphere. (Egly, et al. 1994)

1.3 Multiple Object Tracking Paradigm

One paradigm that has proven useful in the past for evaluating object-based attention (and the one which was utilized in the present experiment) is the multiple object tracking (MOT) technique (Scholl, 2008). In this task, participants are required to track a number of objects marked as targets as they move among other identical objects (distracters) (see figure 1.3 for more complete explanation). MOT is important because it requires attention to some objects that are interspersed with other objects that are to be ignored. This feature makes it unlikely that observers can use a single spotlight of spatial attention to separate targets and distracters and thus appears to require object-based attention. This task reveals that object-based attention is limited in capacity as people can only about four objects before becoming confused and making errors. For most people, this limit is around 4, but can be as low as 2 or as high as 7, depending on the difficulty of the MOT task (i.e. target complexity, speed, etc.), as well as individual differences between individuals.

There are many theories to explain this limitation in tracking ability. Pylyshyn (2003) proposed an object-based theory in which observers track object using cognitive indexes. Pylyshyn believes that there are four separate indexes which can be allocated to different objects during tracking. Since there are only four, tracking more than this number exhausts observers' tracking capacity (Pylyshyn, 2003). Another theory presented by Alvarez & Franconeri (2007) focuses less on objects and more on a central supply of attention. They found that the tracking limit is not fixed, and that manipulating other aspects of task difficulty, such as speed, can allow observers to track up to 8 objects accurately. This theory then supports the idea that it is a limited capacity than prohibits observers from tracking more than 4 objects (at normal speeds) and not four discrete pointers (Alvarez & Franconeri, 2007).

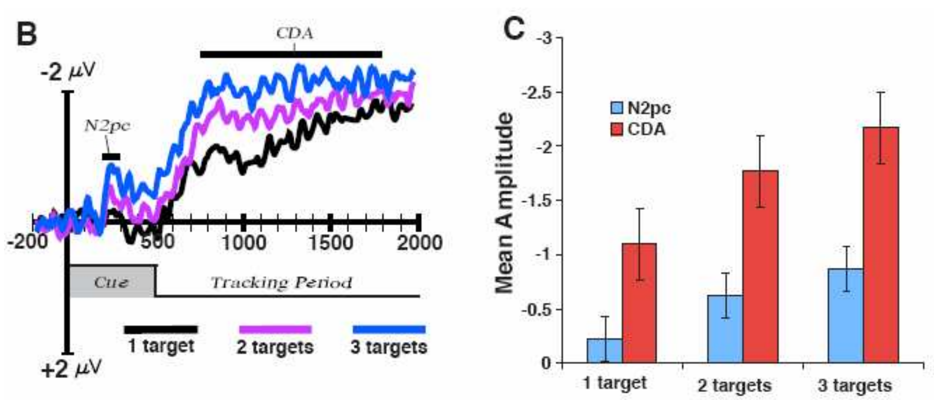
Figure 1.3 **MOT task:** From Keane & Pylyshyn (2005); shows the MOT task in three phases. First, the targets are identified amongst the distracters. Then, the identical objects move around the screen while the observer keeps track of where the targets are. Finally, the test phase has the observer either select the targets or decides whether an identified object is a target or a distracter.



1.4 Contralateral Delay Activity

The MOT task requires attention, and behavioral measures can give a fairly accurate idea of the restraints and limits of attention. In addition, psychophysiological measures such as ERP can be used in conjunction during tracking to give a more complete picture of how attention is allocated during the task. Drew and Vogel (2008) conducted an experiment doing just that. They found that tracking evoked a sustained negative potential in the opposite hemisphere to the visual field containing the tracked objects. For example, when objects were being tracked in the left visual field, this potential was observed over the right hemisphere. They called this potential the contralateral delay activity (CDA). They also reported that the amplitude of the CDA increased as the number of objects being tracked increased but only up to the number of objects that the observer could accurately track (see figure 1.4). Thus, the CDA would be directly related to an object-based attention system. (Drew & Vogel, 2008

Figure 1.4 **CDA:** Results from Drew & Vogel (2008)



Although it is possible that the CDA reflects a strictly object-based system, it is also possible that it represents a more general attentional effort. If this were the case, then any variable which increases tracking difficulty would also increase the amplitude of the CDA. Decreasing inter-object spacing is one such variable. Shim, Alvarez, and Jiang (2008) showed that decreasing proximity between targets decreased subjects' tracking accuracy. In the present experiment, we chose to alter the spacing of all the objects in the display to manipulate difficulty. Decreasing spacing is a potent variable affecting tracking difficulty and may underlie the effects of other variables, such a object speed, that also appear to affect tracking difficulty (Franconeri, Lin, Pylyshyn, Fisher, & Emms, 2008)

1.5 Current Work

In the present experiment, we had participants perform a MOT task in which objects were either in close proximity to each other (near condition) or further away (far condition). We also varied the number of objects to be tracked from one to two in order to provide a replication of Drew & Vogel (2008). We expected to find an increase in CDA amplitude with increases in set size, which would replicate previous work (Drew & Vogel, 2008). If this increase reflected a general increase in the difficulty level of the task, we should observe a similar increase with decreasing spacing between objects.

Chapter 2

METHODS.

2.1 Participants

Twenty eight right handed, neurologically normal volunteers (ages 18-32 years) participated in this experiment. All were naïve to the purpose of the experiment and were paid $10/hour for their participation. All participants reported normal or corrected-to-normal acuity and provided informed consent.

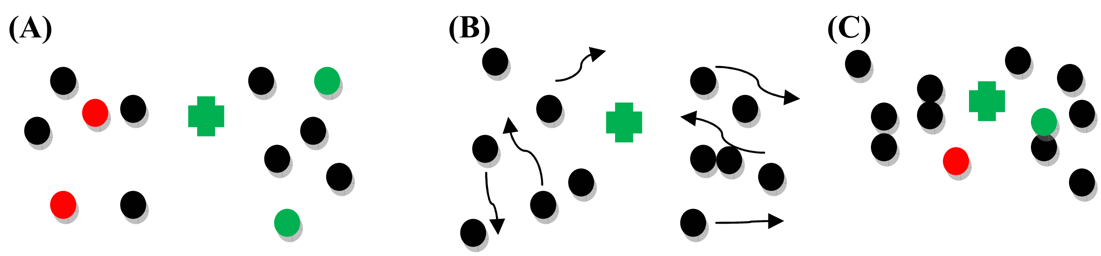
2.2 Stimuli & Procedure.

Stimuli were displayed on a Dell 2.99 GHz computer running custom software written with Blitz3D (Sibly, 2005) and presented on a 17 inch Mitsubishi CRT (1024 X 768 pixel resolution; 75 Hz frame rate). Eye fixation was monitored using a Tobii x50 50-Hz eye tracker (Tobii Technology, Stockholm Sweden) controlled by a Sony 2.86 GHz computer. Testing was conducted in a dimly-lit, electrically shielded room with a chinrest maintaining a 70 cm viewing distance. The monitor screen subtended approximately 27.5º by 21.1º.

On each trial, participants were presented with 6 solid black circles (1.08^{o} in diameter) on each side of the fixation marker which was a green or red "+" in the center of the screen (see figure 2.1). The background of the display was white (82.5 cd/m^2). Circles were randomly distributed in a rectangular area that could have two sizes. In the Close spacing condition, objects were confined to an area 7.8^{o} in height and 6. 7^{o} in width and the minimal spacing between objects was 1.2^{o}. The center of this area was located 7.7^{o} from the fixation marker and was horizontally aligned with it. In the Far spacing condition, the rectangular area was enlarged to 14.3^{o} by 11.0^{o} and the minimal spacing between objects was increased to 2.2^{o}. The center of this area was located 9.0^{o} from fixation.

Each trial consisted of cueing, movement, and response phases which are illustrated in Figure 2.1. During the cueing phase, one (set size one) or two (set size two) circles on one side of the screen were red (19.1 cd/m$^{2)}$ and the same number of circles on the opposite side were green (52.1 cd/m$^{2)}$. The number of colored circles and the assignment of colors to left and right visual fields varied randomly across trials. The starting locations of the circles were random within constraints detailed below. Individual observers were instructed to attend to either red or green and their task was to track the cued objects over the course of the trial. Five hundred ms after cue onset, the red and green disks changed color to black and the movement phase began 100 ms later. All 12 objects moved in randomly determined directions with a velocity of 3.4 ^{o}/sec. During movement, objects bounced off of the edges of the area boundaries as well as each other (when inter-object distance was less than the minimal distance for that spacing condition). Immediately after cessation of movement, one of objects on each side was colored red or green (the test probe). The observer had to click on one of two onscreen response buttons with their mouse to indicate whether the object in the attended color was one of the cued objects. Matching and mismatching probes were presented equally often. Observers received immediate feedback regarding the accuracy of their response.

Figure 2.1 **Our MOT task:** An example trial of the MOT task used in the current experiment. (a) The targets are identified by changing color (the green cross indicates the subject should track the green objects). (b) The targets turn back to black and the move among the identical distracters while the participant tracks them. (c) The objects stop moving and one object on either side changes color. The participant must then choose whether (in this case) the green object is a target or a distracter by clicking 'same' or 'different.'



There were 16 types of trials: two Set Sizes (1 or 2 objects to be tracked) by two Spacings (Close vs. Far) by two attended visual fields (left or right) by two test probe types (matching/mismatching). Each of these 16 trial types was repeated 50 times for a total of 800 trials. All 800 trials were presented in a random order with a short break after each set of 160 trials. Observers were instructed to maintain fixation throughout the trial. If the eye tracker indicated more than a 0.5 degree deviation from fixation, observers were presented with a visual playback of their eye movements and warned to remain fixated. Observers received practice trials at the beginning of the experiment until it was clear they understood the task and were proficient in maintaining fixation.

2.3 Electrophysiological Recording.

The electroencephalogram (EEG) was recorded with an Electrical Geodesics Inc. (EGI; Eugene, OR) 128-channel Geodesic Sensor Net with individual electrode impedances kept below 50-75 kΩ, as recommended by the manufacturer. The data were referenced online to the vertex, bandpass filtered between .01-80 Hz, and were digitized at 200 Hz. Subsequent processing was performed offline using EGI Net Station 4.1.2. The data were bandpass filtered between 1-55 Hz and were segmented using an epoch that began 100 ms prior to the onset of the cue display flash and ended 2000 ms after. Net Station's artifact detection routines were then applied. Individual channels were marked as bad if there was zero variance, the fast average amplitude exceeded 200 μV, or the differential average amplitude exceeded 200 μV. Individual segments were rejected if they contained eye movements or blinks (threshold = 70 μV) or if more than 10 channels were marked as bad. For the remaining segments, bad channels were replaced using interpolation from surrounding channels. Finally the segments were averaged, rereferenced to the average reference, and baseline corrected using a 100 ms prestimulus interval.

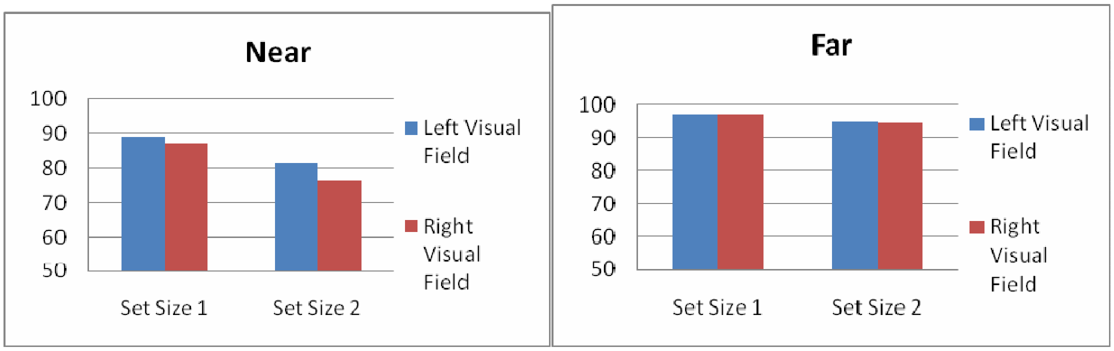
Chapter 3

RESULTS

3.1 Behavioral Results

The percent correct scores for each participant were entered into a repeated measures factorial analysis of variance with factors of spacing (close vs. far), set size (1 or 2), and visual field (left vs. right). All main effects were involved in interactions and we only report the significant interactions. There was a spacing X visual field interaction ($F(1, 22) = 78.81$, $p < .001$), reflecting greater effect of spacing for tracking in the right visual field compared to the left. These results can be seen in figure 3.1. Participants were more accurate with far spacing compared to near particularly when tracked objects were in the right visual field. Therefore, it can be inferred that the spacing manipulation was sufficient to make the task harder for participants. There was also a significant set size X visual field interaction ($F(1,22) = 7.25$, $p = 0.013$) reflecting a greater set size effect in the right visual field versus the left. Again, participants were more accurate at tracking 1 object than 2, especially when tracking in the right visual field. Therefore, the set size manipulation was also successful.

Figure 3.1 **Behavioral Results:** Behavioral results for the near versus far conditions.



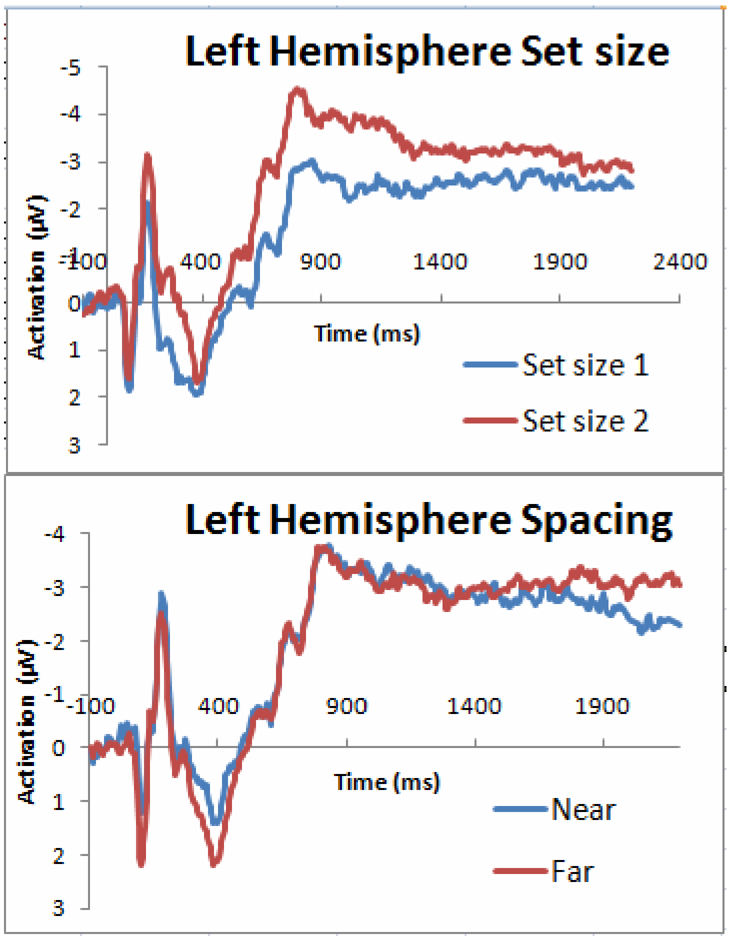
3.2 ERP Results

The psychophysiological data was analyzed by averaging together ERPs from electrodes near the back of the head over each hemisphere; a technique similar to the one used by Drew & Vogel (2008). However, we decided to look at the data from each hemisphere separately instead of collapsing over the two sides and looking at an overall CDA effect (see Drew & Vogel, 2008 for complete description of their experimental techniques).

3.2.1 Left Hemisphere Activation

The results for left hemisphere are shown in figure 3.2. The average ERP voltage for each participant was entered into a repeated measures factorial analysis of variance with factors of set size (1 and 2), spacing (near and far), and visual field (left or right). There was a significant effect for visual field ($F(1,21) = 25.09$, $p < 0.001$) in that tracking in the right (contralateral) elicited more activation than the left (consistent with Drew & Vogel, 2008). There was also a significant visual field X set size interaction ($F(1,21) = 7.53$, $p = 0.012$) such that when looking exclusively at the right (contralateral) visual field there was a significant effect for set size ($F(1,21) = 8.1$, $p < 0.001$) where observers showed more activation when tracking 2 targets versus 1. However, when looking at the left (ipsilateral) visual field, no significant effect for set size was obtained. In addition, there was no significant main effect or any significant interactions involving the spacing manipulation.

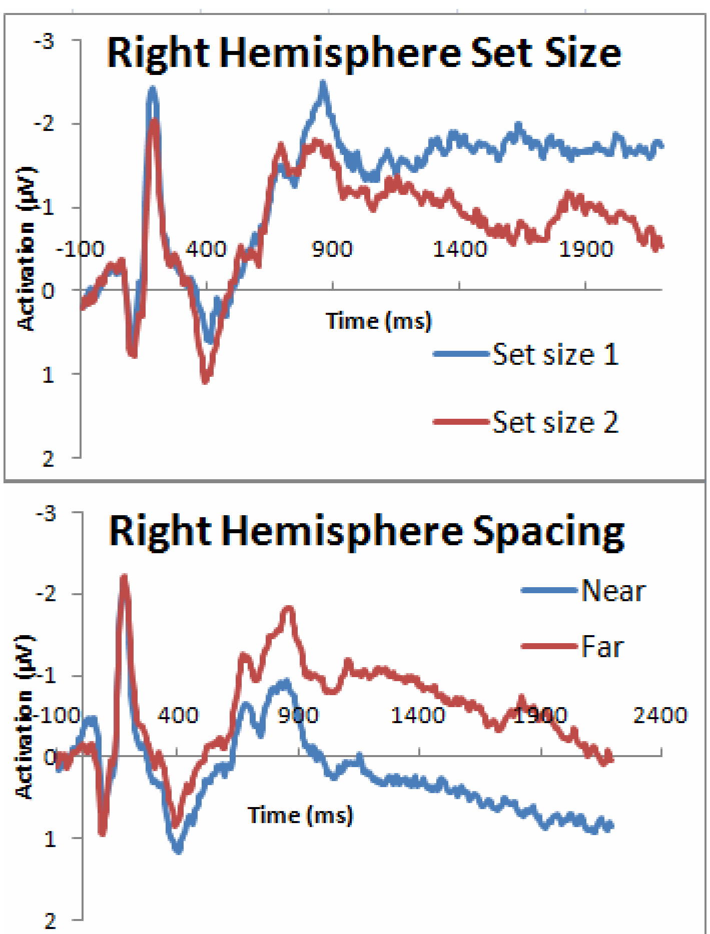
Figure 3.2 **Left Hemisphere Results:** CDA for the left hemisphere when tracking in the right (contralateral) visual field. There is a significant difference in activation due to set size where 2 targets caused more activation than 1 target (top). However, no significant effect for the spacing variable was found in this hemisphere (bottom).



3.2.2 Right Hemisphere Activation

The results for right hemisphere activation are shown in figure 3.3. The average ERP voltage for each participant was entered into a repeated measures factorial analysis of variance with factors of set size (1 target or 2 targets), spacing (near or far), and visual field (left or right). Again, we saw a main effect for visual field ($F(1,21) = 15.49$, $p = 0.001$) such that tracking in the left (contralateral) visual field elicited more activation than tracking in the right (ipsilateral), also consistent with Drew & Vogel (2008). There was no significant effect for set size in either visual field. However, there was a significant main effect for spacing ($F(1,21) = 10.11$, $p = 0.005$) such that observers showed more activation in the far spacing condition than the near (see figure 3.3).

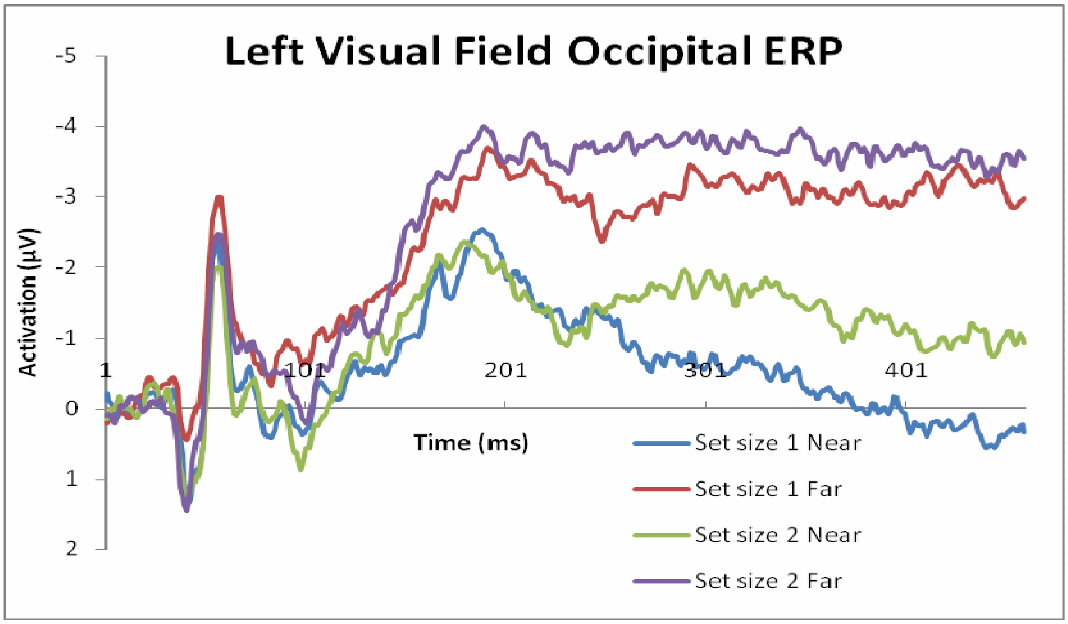
Figure 3.3 **Right Hemisphere Results:** CDA results for the right hemisphere when tracking in the left (contralateral) visual field. There were no significant effects for the set size manipulation in this hemisphere (top). However, there were significant effects for the spacing manipulation in that the far spacing caused more activation than the near (bottom).



3.2.3 Occipital Results

We also averaged data from electrodes over the occipital component of the brain and the resulting waveforms are shown in figure 3.4. The average ERP activation for each participant was entered into a repeated measures factorial analysis of variance with factors set size (1 target or 2 targets), spacing (near or far), and the results were grouped by visual field (left or right). We found that when participants were tracking in the right visual field, neither of the manipulations had significant effects. However, when tracking in the left visual field, there was a main effect of spacing such that observers showed more activation in the far condition than the near condition ($F(1,21) = 7.53$, $p = 0.012$).

Figure 3.4 **Occipital Area Results:** ERP results for the Occipital area when tracking in the left visual field. The results for this area show a trend similar to the right hemisphere results. Again, there is an effect for spacing such that the far condition caused more activation than the near. In addition, there was no effect for set size and no effects for either condition when tracking in the right visual field.


Left Visual Field Occipital ERP
Activation (µV)
-5
-4
-3
-2
-1
0
1
2
1
101
201
301
401
Time (ms)
Set size 1 Near
Set size 1 Far
Set size 2 Near
Set size 2 Far

Chapter 4

DISCUSSION

We found that increasing the set size does increase the amplitude of the CDA component; a finding that replicates the results reported by Drew & Vogel (2008). There was a significant increase in activation when participants tracked 2 objects versus when they were tracking only one. However, this set size effect was only significant with activation in the left hemisphere. When looking at right hemisphere activation, there wasn't a significant difference between set size 1 and 2. Therefore, an object-based increase in activation seems to occur only in the left hemisphere and not the right. This difference may not have been reported in the Drew & Vogel (2008) results as their CDA was collapsed over the hemispheres and presented as overall data. A hemispheric difference could only be observed if data from the two hemispheres was dealt with exclusively.

We also found that although there was an effect for spacing, it was contrary to our original hypothesis. If an underlying capacity theory were true, we should see that near spacing induced more activation than far. Yet, in fact, we see the complete opposite. Therefore, our results do not represent a general increase in effort as the task gets harder. The fact that this spacing effect was only seen in the right hemisphere, and this area was not sensitive to number of objects, is also interesting

This may lend support to previous findings, such as the Egly, et al. (1994) experiment, that imply the right hemisphere may only deal with spatial information.

Furthermore, our findings may imply that the right hemispheric CDA represents a separate mechanism from the left hemispheric CDA. The left hemisphere was the only site where a set size effect was observed. Again, this finding goes along the lines of work such as the Egly, et al (1994) experiment that has isolated the left hemisphere as the site of a cognitive system representing object-based attention. The CDA seen over the right hemisphere is sensitive to space while the left hemisphere CDA is sensitive to number of objects. Thus, the results may support a dual-system hypothesis of attention; that object-based attention requires a neurological process that is distinct from the spatial attention system.

It is not known why the far condition elicited more activation than the near in the spacing manipulation. However, one plausible explanation could simply be that the far condition required the viewer to attend to more visual space than the near and the increased activation represented this difference. This thought could be supported by the fact that a similar pattern in activation was seen over the occipital area of the brain. It is possible that the increased activation of these electrodes reflects an increase in activity in the occipital cortex. It could then be that this increase in attended area also increased the area which needed to be represented in an early V1 system located in this neural region. It is also important to note that the occipital effects were only observed when the objects were being tracked in the left visual field;

information that would be processed by the right hemisphere. Again, this supports the idea that the right hemisphere is specialized for space-based processes.

Whatever the reason for the spacing effect, it is clear from these results that the set size and spacing manipulations caused changes to different systems in the brain. This finding strongly supports the hypothesis that the brain has two separate cognitive systems that deal with visual attention. Object-based attention may involve a system which is distinct from spatial attention, and this system may be represented by the CDA component in the left hemisphere. In addition, spatial attention may involve processes distinct from object-based attention in early visual processing, which may be represented by both the right hemispheric CDA component and the occipital component observed here. Whether or not these two systems are mutually exclusive or if they have some overlap is not evident through the present research, but may be an area of interest for future studies. Continuing to use psychophysiological methods such as ERP can help untangle the questions and bring us closer to knowing the answers.

WORKS CITED

Alvarez, G. & Franconeri, S. (2007) How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision,* 1-10.

Drew, T., & Vogel, E. (2008). Neural measures of individual differences in selecting and tracking multiple objects. *Journal of Neuroscience*, 4183-4191.

Egly, R., Rafal, R., Driver, J., & Starrveveld, Y. (1994). Covert orienting in the split-brain reveals hemispheric specialization for object-based attention. *Psychological Science* , 380-383.

Franconeri, S., Lin, J., Pylyshyn, Z., Fisher, b., & Enns, J. (2008). Evidence against a speed limit in multiple object tracking. *Psychonomic Bullentin & Review*, 802-808.

Keane, b., & Pylyshyn, Z. (2006). Is motion extrapolation involved in multiple object tracking? Tracking as a low-level non-predictive function. *Cognitive Psychology* , 346-368.

Luck, S., Woodman, G., & Vogel, E. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences* , 432-440.

Pylyshyn, Z. (2003). *Seeing and visualizing: It's not what you think.* Cambridge, MA: MIT Press.

Scholl, B. J. (2008). What have we learned about attention from multiple object tracking (and vice versa)? In D. Dedrick, & L. Trick, *Computation, Cognition, and Pylyshyn.* Cambridge, MA: MIT Press.

Shim, W., Alvarez, G., & Jiang, Y. (2008). Spatial separation between targets constrians maintenance of attention on multiple objects. *Psychonomic Bulletin & Review* , 390-397.