A Pilot Study Using Tobii Studio 1.3.22
to Recreate the Burke et al. (2004) Banner Ad Study

Kyle Vessey and Anthony J. Hornof
September 1, 2009
University of Oregon
Department of Computer and Information Science
Technical Report CIS-TR-2010-05

Introduction

Advertisers often use flashy animated banners to grab a web users attention. The empirical evidence however seems to contradict this commonly used tactic. In 2002 Bayles found no evidence of a correlation between animation and memory recall. In 2004 Burke, Hornof, Nilsen, and Gorman confirmed this by finding in a series of two experiments that animated banner advertisements do interfere with common web tasks, such as in casual browsing; however, the animated banners are no more memorable then the static ones. The goal of this study is to reenact part of the second experiment done by Burke, Hornof, Nilsen, and Gorman in order to check the validity of their findings by examining where visual attention is being drawn during the task. In other words, do animated banner ads in a top-down visual search task draw more attention than other types of banner ads.

Background and Similar Work

Memory and eye movements are closely connected. In order to recall an object, some amount time gathering visual information must have occurred for that object, or in other words, that object must have had some amount of visual attention given to it. Eye movements can be classified into two sets of movements: fixations, and saccades. A fixation is a period of time when the eyes are relatively still, and gazing at one point lasting 150-600 milliseconds. Saccades on the other hand are much faster eye movements that occur between fixations [4]. For this reason eye tracking is employed because visual information is generally only collected during times when the eyes are fixated [3]. This is important because fixation count and length can be used to quantitatively determine how much visual attention an object receives.

In 2009 Buscher, Cutrell, and Morris preformed a similar eye tracking study to the one reported in this paper. In their experiment Buscher, Cutrell, and Morris conducted a series of web based search tasks in order to understand how people allocate their visual attention to web pages. Their experiment consisted of four tasks broken into two different categories: foraging tasks, and free query tasks. In the foraging tasks, participants would be confined to a set of nine web pages, they would be given a certain topic, and then given a search task on that topic. For example, let the topic be "cars" and the search task could be "which car gets the best gas milage?". The participant must then search through the provided set of web sites to find the answer to the given search task. Free query tasks however did not confine the participants to a defined set of web sites, and each participant was freely able to search the web in any way that he or she wanted to in order find the answer to some search task. Participants were each given two foraging tasks and two free query tasks. The websites that the participants viewed were recorded and used next in the final recognition task. The recognition task involved each participant to state the familiarity with each page, and with twelve previously unseen pages taken from popular web sites such as google.com or amazon.com. From the results of this experiment it can be concluded that where visual attention is being drawn during web page viewing depends on the task. A person will not give the same visual attention to one web page if he or she encounters it during a different task. This result is also noted by Diaper and Waeland (2002), who also observed that the nature of the task influences the strength of distractors within the web page.

Another interesting and similar experiment was preformed in 2009 by Teevan et. al. [5]. This experiment is similar to the previously mentioned study as it too has a search task followed by a recognition task. During the search task participants were presented with a portal containing three different types of links: text snippets, visual snippets, and thumbnails. Text snippets are a simple text link followed by some brief site overview text, visual snippets are small 120x90 squared pixels pictures that contain an image representing the site, and a very short piece of text, and finally thumbnails are again small 120x90 squared pixels pictures that are simply a shrunken picture of the entire site. In the search task the participants were asked to find a website containing some desired information, for example “diabetes medication,” using either the text snippets, visual snippets, or thumbnails to navigate. The site the participant ended up at was recorded and then reused a day later for the second task, the recognition task. In this task the participants would have to relocate the same sites they ended up at the day before. The results of this experiment conclude that searching is faster with the text snippets, but recognition was faster with the visual snippets.

These two experiments are very similar to the experiment conducted by Burke, Hornof, Nilsen, and Gorman in 2004 . They contain a search task followed by a recognition test based off of what the participant encountered during his or her searching. They also all contained a top-down, or goal orientated, visual search task. In a top-down visual search task, the participant is searching for a certain bit of information based off of some information he or she received earlier. This is also very similar to the experiment conducted in this paper. The main difference is that there was no recognition test carried out at the end of the this experiment however. The main similarity is that they all use a top-down visual search task. The experiment conducted in this paper only consists of this search task where in it the amount of visual attention given to each banner ad is measured, and then this experiment will try to prove the validity of the argument that states that animated banners are no more memorable then other types of banner ads in a top-down visual search task.

Methodology

Apparatus

A 60 Hz Tobii T60 eye tracker will be used and all gaze data will recorded and analyzed using Tobii Studio version 1.3.22. Gaze data will be filtered into fixations by Tobii Studio using the ClearView built in dispersion based fixation detection method.

Stimulus

A total of 24 stimulus webpages were built containing two banner ads with nine distractor headlines, and one target headline. The participant was first presented with a precue page which contained the precue at the top of the page written in black 14pt Arial font and a link to the searching page labeled “GO”. On the searching page the target and distractors were randomly located in a 2x5 grid. All target and distractor headlines were colored blue and in 14pt Arial font. Each searching page contained two banners ads of one of the following types: static, animated, or solid gray. Static banners were simple ads only containing non-moving images, animated banners were ads that contained images that moved or flashed, and solid gray banners were plan gray rectangles with no images. Both banner ads on a searching page were of the same type; one banner was always located at the top of the page, and the other was located randomly between two of the five rows of headlines.

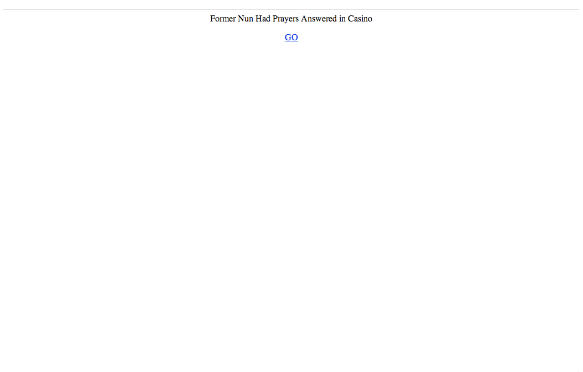

Fig 1. Precue Page.

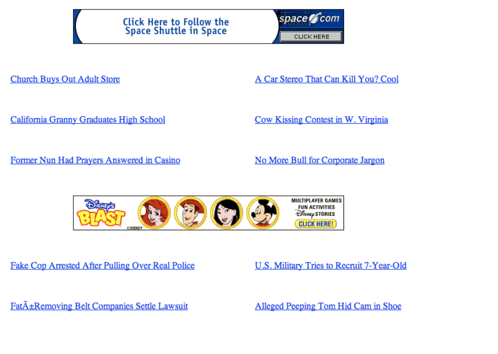

Fig 2. Searching Page.

Subjects

Five students enrolled in CIS 410/510 Eye Tracking methodology and Applications were recruited to participate.

Experiment Design

The experiment consisted of a total of 36 trials. Each trial consisted of a precue page followed by a searching page. One of two types of precues were used: exact precues, and semantic precues. Exact precues matched the headline exactly word-for-word, and semantic precues consisted of a few sentences that appeared to be from the the target headlines story. For example, the semantic precue for “Drop-outs doing just fine, thanks” is:

> New research debunks the common belief that leaving school before completing year 12 diminishes a teenager’s chance of a successful career. [2]

After clicking a link in the search page the participant navigated to a pass or fail page which informed the participant if he or she clicked on the correct headline. When navigating away from the pass or fail page the participant was presented with the precue for the next trial. The trials were separated into two different blocks. Each block consisted of an explanation of the search task, one example trial, and 12 experimental trials. Only exact precues were used in the first block of trials, and only semantic precues were used in the second block of trials.

For each block, each type of banner showed up in four trials. The order of the 12 trials were randomly assigned during the creation of the webpages; this was so that the order of encountering banner ads was randomized. Each participant viewed the same set of web pages in the same order. Due to time constraints, there was not enough time to build different sets of web pages for each participant.

Results

Fixation count and the total time in seconds that gaze was on a region (time on region) were the two measures that were found to be most useful in answering the question of whether or not animated banners in a top-down visual search task draw the users attention more than other types of banner ads. A 2x7 grid of areas of interest (AOIs) were overlaid on each webpage. One AOI represented either the target headline, a distractor headline, or one half of a banner ad. This resulted in having two banner ad AOIs per one banner ad. Having two AOIs per banner ad did not cause problems because AOIs with the same name get grouped together to represent one AOI in Tobii Studio, so banner ad AOIs were named either gray, static, or animated.

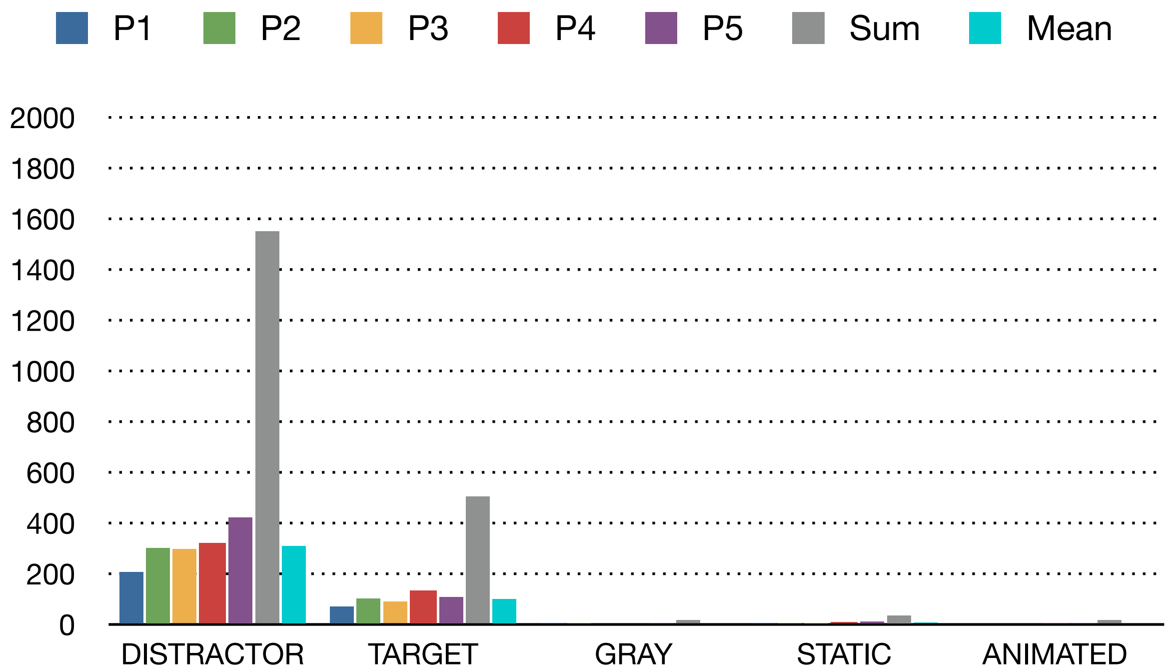

Fig 3. Fixation count as a function of screen region for all trials. P1, P2, P3, P4, and P5 show the individual participant's fixation count. Sum is the sum of all the participants fixation counts, and Mean is the average fixation count for all participants.

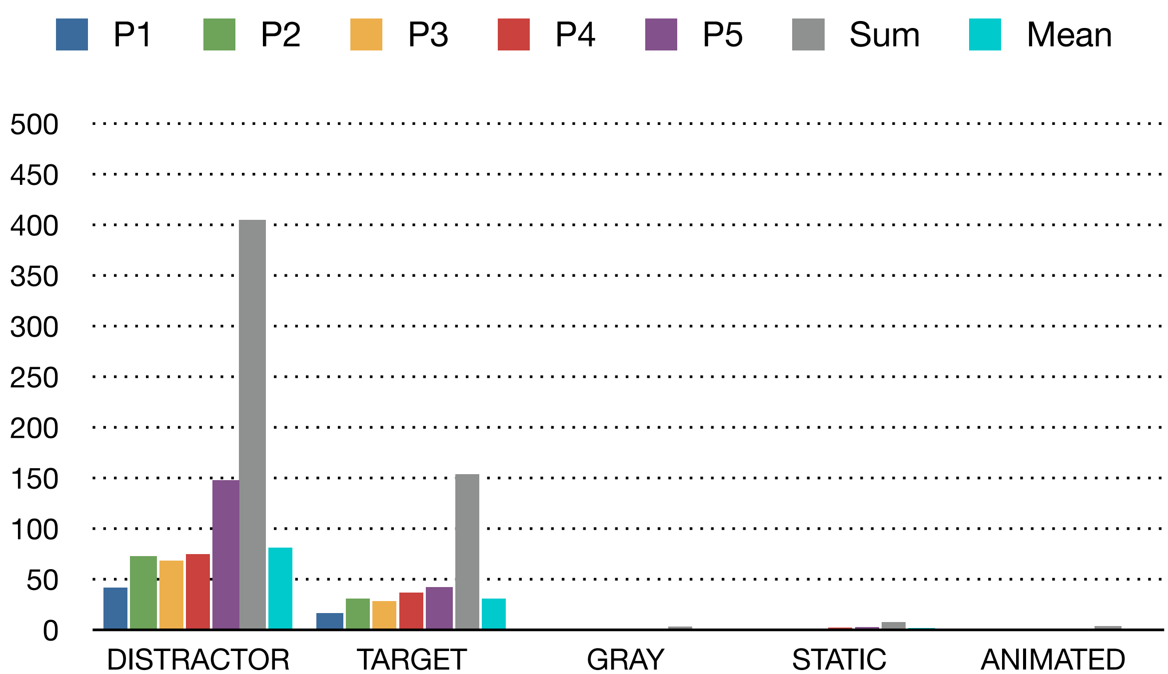

Fig 4. Time on region in seconds as a function of screen region for all trials. P1, P2, P3, P4, and P5 show the individual participant's time on region. Sum is the sum of all the participants times on region, and Mean is the average time on region for all participants.

Assuming that fixation count and time on region is a measure for determining visual attention, the first observation is that participants clearly devoted most of their visual attention to the distractor and target headlines. This is seen from figure 3 and 4, where the banner ad AOIs fixation count and time on region cannot even be observed from the graph; though they are all artificially depressed since they only appear one third of the time. This is not surprising or unexpected since the task was a top-down visual search of the webpage. In order to find the target headline the participant had to search through all the headlines, and so more visual attention was clearly spent on the task. The interesting results come from figures 5 and 6, where the headline AOIs are ignored, and visual attention is only compared between the different banner ads. Here we can clearly see that the animated banner ads did not attract any more visual attention than the other types of banners. (Each type of ad appeared equally as often across the experiment.)

Discussion

Looking at figure 5 and 6, we see overall that some visual attention was given to the banner ads, but there was no evidence that more visual attention was given to the animated banners. In fact, the data seems to show that overall less visual attention was given to animated banners. This may be due to the participants knowing that there will be banner ads in the study that may impede their search task, and thus they could have made a conscious decision to avoid looking at movement picked up in their periphery. Without the movement of an animated banner the participant may not be able to see the ads in their periphery as well, and so they may not be able to make this conscious decision to avoid looking at the ads.

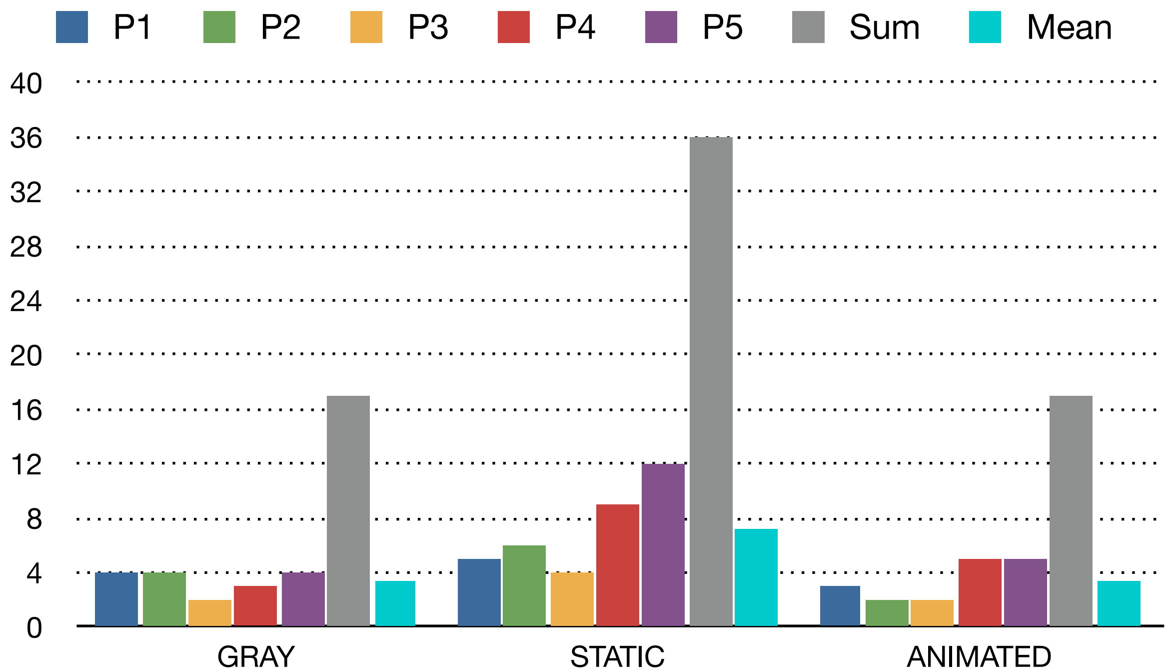

Fig 5. Fixation count as a function of screen region for all trials. P1, P2, P3, P4, and P5 show the individual participant's fixation count. Sum is the sum of all the participants fixation counts, and Mean is the average fixation count for all participants.

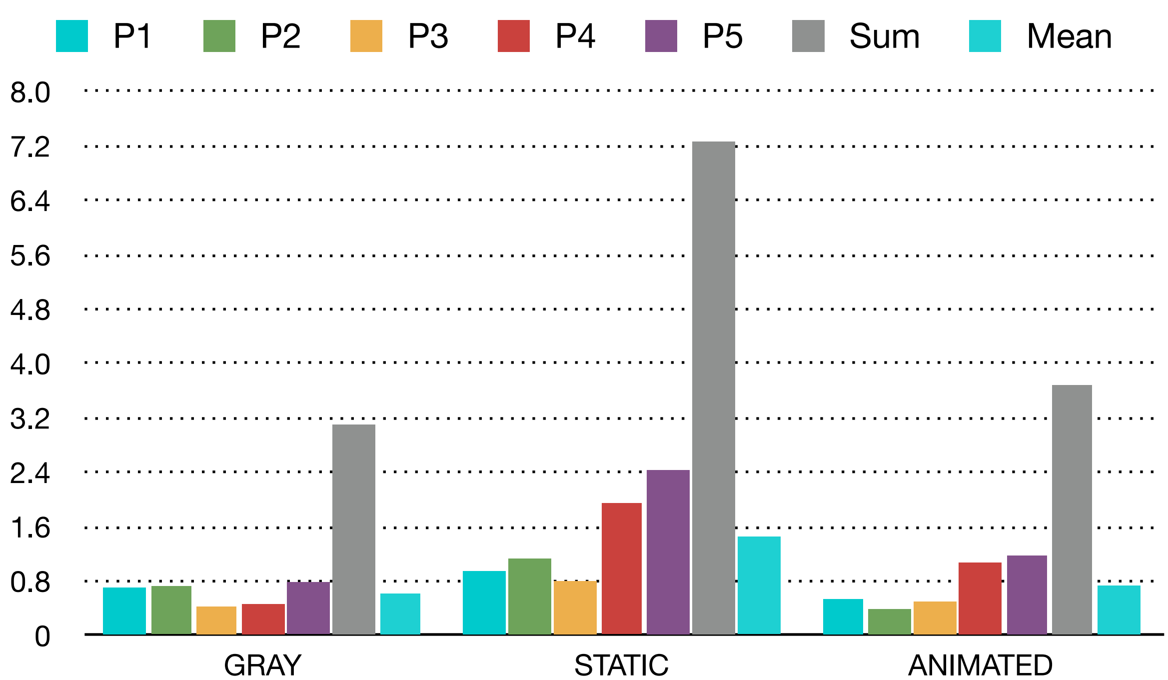

Fig 6. Time on region in seconds as a function of screen region for all trials. P1, P2, P3, P4, and P5 show the individual participant's time on region. Sum is the sum of all the participants times on region, and Mean is the average time on region for all participants.

There was also no evidence to show that more visual attention was given to the animated banners when comparing visual attention in the exact precue condition versus the semantic precue condition. Figure 7 to 10 show again fixation count and time on region per banner ad AOI, but these only show data from either exclusively the exact precue condition, or the semantic precue condition. The number of fixations and total duration of gaze on animated banners are consistently lower than on the static banners. We do see that the gray banners less time and fewer fixations when looking at the data from the semantic condition however. Since the semantic precues occur during the last block, this may be due to a learning effect. That is, the participants may have learned to avoid looking at the gray banner ads. This brings up an interesting question: Why did the participants learn to avoid looking at the gray banners, but not the static banners? This may be from the fact that the static banners have some information on them, and may be attracting the participants visual attention to it. This can also be said about the animated banners as well, but in the animated case the participant may be actively trying to avoid looking at them. This shows that banner ads do in fact attract a users visual attention, but there is no evidence that the animated banners attract more visual attention then static. In fact the data from this experiment seems to show the opposite. It appears from these results that animated banners attract less visual attention because the user can pick up the movement from their periphery vision, and then decide not to look there.

For Exact Precues:

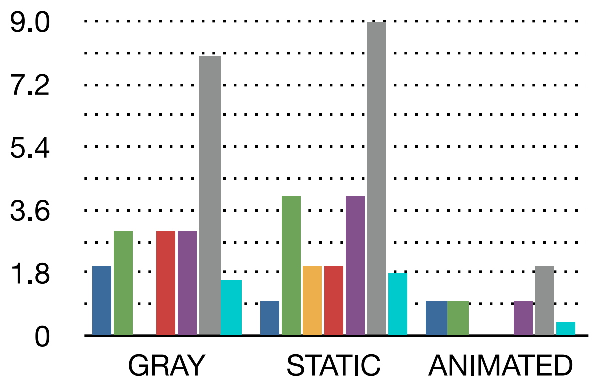

Fig 7. Fixation count as a function of screen region.
P1, P2, P3, P4, and P5 show the
individual participant's fixation count.
Sum is the sum of all the participants fixation counts, and
Mean is the average fixation count of all participants.

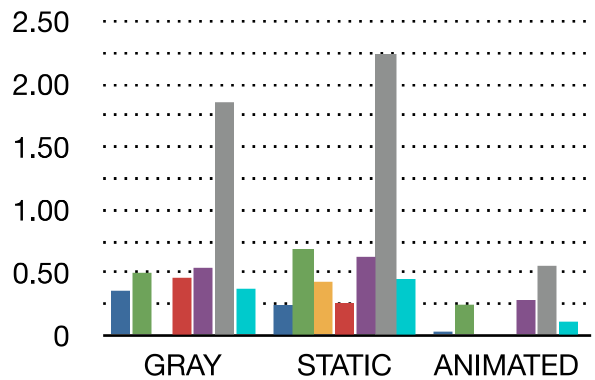

Fig 8. Time on region in seconds as a function of screen region.
P1, P2, P3, P4, and P5 show the
individual participant's time on region.
Sum is the sum of all the participants times on region, and
Mean is the average time on region for all participants.

For Semantic Precues:

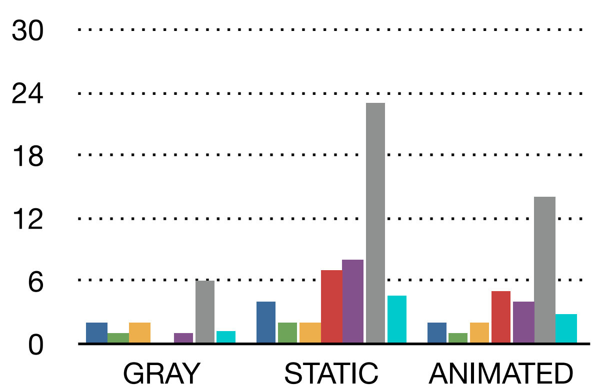

Fig 9. Fixation count as a function of screen region.
P1, P2, P3, P4, and P5 show the
individual participant's fixation count.
Sum is the sum of all the participants fixation counts, and
Mean is the average fixation count of all participants.

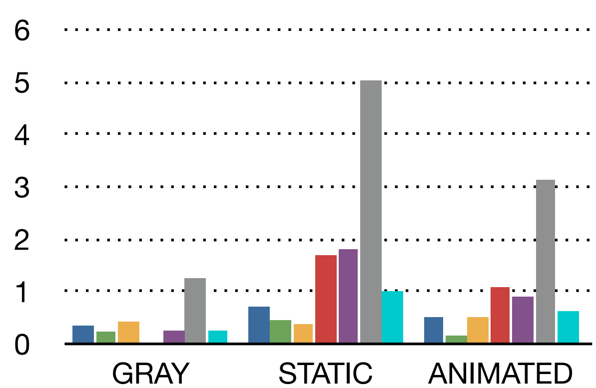

Fig 10. Time on region as a function of screen region.
P1, P2, P3, P4, and P5 show the
individual participant's time on region.
Sum is the sum of all the participants times on region, and
Mean is the average time on region for all participants.

Conclusion

The results of this experiment show that some visual attention is given to animated banner ads, but they do not attract any more attention than the other banner ads. In fact, it appears that participants are actively trying to avoid looking at the animated banners by sensing the movement from the animated banners in their periphery and then trying to avoid looking at that location. This strategy may only work for animated banners because the static banners have no movement, and so they cannot be picked up as well from peripheral vision. Lastly participants may have learned to avoid looking at the gray banner ads since they are dull and boring, but may have still been attracted to the images on the static banner ads.

This conclusion comes from a top-down visual search task. An interesting follow up experiment would be to try this with a bottom-up task. For example the participant could be given some amount of time to browse through the headlines on the webpage, and then he or she would just click the headline that he or she found the most interesting. The would be no precue, and no target or distractor headlines. This would be a very easy to do with the current experiment design, and would make a nice extension to the conclusion of this paper.

Tobii Studio Discussion

As mentioned earlier Tobii Studio version 1.3.22 was used to record the eye tracking data. One of the primary motivators for this project was to determine and evaluate how Tobii Studio is used for collecting eye movement data for a web study. The process for this experiment can be broken into three steps: stimulus creation, running participants, and data analysis. What follows is a brief discussion on my experiences with using Tobii Studio during these three steps.

Stimulus Creation

Creating the webpages was straight forward. Templates were created for a precue page, search page, pass and fail pages, all banners, links, and precues were randomly entered into the templates. Tobii Studio has nothing to do with the generation of these web pages and thus they can be created in any way the experimenter wants. The web pages for this experiment may have been hacked together, but a more sophisticated method could be employed to make the experiment more like the original. For example, a simple script could be written to randomly generate the webpages. This way each participant will have a randomly generated series of webpages, and they will not all experience the exact same stimuli like how it was carried out here. The only limitation that Tobii Studio imposes on the web pages is that they must open up in Internet Explorer 7.

Running Participants

Tobii Studio provides a very easy way to start up and run a participant. The process of doing this merely consisted of creating a new Tobii Studio experiment, dragging a web page element into the experiment time line, entering the path name to the starting page of the web study, and then hitting run. This was very quick and easy to do. In fact an experimenter would only have to know how to create web pages for running a web study on Tobii Studio because no other background knowledge is needed in order to run a web studio on Studio.

Data Analysis

Tobii Studio does do an effective job of allowing an experimenter to create and run experiments rather quickly and easily. Keeping in pace with the other two steps, the data analysis part of Tobii Studio is fast and easy as well. An experimenter could very quickly show some results from his or her study and move on to report them. In my opinion however, Tobii Studio fails in the data analysis part of the process because two key questions are not addressed properly by Studio: How accurate is the data I am reporting, and how can I report it?

From my experience and knowledge with Tobii Studio, there is no way to properly adjust and fix the eye movement data in order to make it more accurate. For example, say you have a required fixation location at the top of your screen, and the participant most look there at regular intervals in the study. From these events an analyst would be able to judge how accurate the eye data was, and would probably be able to determine some amount of regular correction to apply to the data in order to fix it. This is not possible with Studio however. No matter how obvious the error is, or how simple the correction would be, there is no way to adjust the data to correct this. This leads to two options: Run more participants to try to get better data, or ignore the errors and report results based off of data that is not as accurate as it should be. With Tobii Studio, what you record is what you get, and so you better hope that it is good.

The second downfall of Tobii Studio for me was its capability to show results. In Studio's favor, it is very easy to define areas of interests (AOIs), every webpage that is visited by the participant is logged, and Studio allows AOIs to be placed onto those pages. Another point for Tobii Studio goes to its ability to keep track of AOIs. For example if you define twenty AOIs and name them all *AOI1*, then Studio does a good job by grouping them together and reporting the data for them as one AOI. However, studio does not fully group these AOIs together. Although they are reported as one, they are all still separate AOIs, and if you want to disregard *AOI1* during a showing of the results, like I did with the distractor and target AOIs in my study, then you will have to disengage every single separate AOI named "AOI1" from the AOI list, which is tedious and annoying.

Another tedious and annoying part of AOI generation is the AOI copy and paste feature. The shape and location of the AOI is copied, but the name is not. When pasted, the AOI will have a new name. This does not seem to be a big issue, but as mentioned earlier AOIs with the same name are grouped together. Let's say that you had AOIs defined as the four quadrants of the stimulus image, and you had say 50 stimulus images. If you wanted to see how many times participants looked in the different quadrants for all of the images, you could copy and past an initial AOI overlay onto each image, but then you would have to go back through all the AOIs and rename them in order to get the grouping. That would be renaming four AOIs on 50 images, so a total of 200 renaming operations. This is all very tedious and annoying to do.

Another issue that does not help with AOI generation, is that there is no *Undo* command in Tobii Studio. I would often hit a wrong button when doing the repeated steps for AOI generation, and this would usually result in me going back and re-doing what I was working on because there was no way to undo my previous mistake. This makes the AOI generation even more tedious and annoying then it already is.

One last point that hurts Studio's analysis capabilities is the chart generation. It does make quick and very nice-looking heat maps and gaze plots that are very well done and will allow you to select which participants' data to include in the graph, but these are more often used for getting a general understanding of the data. General chart generation and display for Tobii Studio is very limited. It can make pretty-looking charts, but you do not have much control over how that chart looks and how that chart presents the data. It was far easier to export the data into a spreadsheet and generate the charts from there.

Overall Impression

Tobii Studio is a very fast and easy program to make and conduct eye tracking experiments. It also provides a quick and easy way to visualize the data and report some results. Tobii Studio is not a bad program, but it may not be worth the cost of upwards of $10,000. I am not an expert with Tobii Studio, so the limitations and annoyances I encountered with Tobii Studio may have solutions which I was unable to discover, but they cannot be ignored. In my opinion, Tobii Studio is not a data analysis tool, but a data collection and display tool. I can see that Tobii Studio has the potential to do great things, and maybe if it did do these things properly then it would be worth the cost, but for me it has not yet reached that potential.

References

[1] Bayles, M. (2002). Designing Online Banner Advertisements: Should We Animate? *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves,* 363-366.

[2] Burke, M. , Hornof, A., Nilsen, E. & Gorman, N. (2004). High-Cost Banner Blindness: Ads Increase Perceived Workload, Hinder Visual Search, and Are Forgotten. *Transactions on Computer-Human Interaction, 12*(4), 423-445.

[3] Buscher, G., Cutrell, E. & Morris, M. (2009). What Do You See When You're Surfing? Using Eye Tracking to Predict Salient Regions of Web Pages. *Proceedings of the 27th international conference on Human factors in computing systems,* 21-30.

[4] Diaper D. & Waelend, P. (2000). World Wide Web working whilst ignoring graphics: good news for Web page designers. *Interacting with Computers 13*, 163-181.

[5] Jacob, R. K. (1990). What you look at is what you get: eye movement-based interaction techniques. *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people,* 11-18.

[6] Teecan, J., Cutrell, E., Fisher, D., Drucker, S., Ramos, G., Andre, P. & Hu, Chang. (2009). Visual Snippets: Summarizing Web Pages for Search and Revisitation. *Proceedings of the 27th international conference on Human factors in computing systems,* 2023-2032.