Using Eye-Tracking Data and Mouse Cursor Location To Examine Visual Alerting in a Multi-Display Environment

Jacquelyn Crebolder\textsuperscript{1} and Joshua Salmon\textsuperscript{2}

\textsuperscript{1}Defence Research & Development Canada Atlantic Research Center
Dartmouth, NS, B2Y 3Z7, Canada

\textsuperscript{2}Dalhousie University
Halifax, NS, B3H 4R2, Canada

ABSTRACT

A study was conducted to assess the effectiveness of visual alerting during a task that required full attention and that used multiple displays. Alert detection time was collected, and eye-tracking data was recorded to determine where participants were looking, particularly when an alert appeared. Results showed that a full border around the display was detected faster than a short bar at the top of the display. This finding is in contrast to previous work in our lab, where the bar alert has always been superior for detection. Previous findings had inferred that the bar alert can be included in a spotlight of attention created during the task. The new findings suggest that the spotlight was expanded in the current experiment as a consequence of limitations in head movement imposed by wearing the eye-tracking equipment. As a result the bar was not captured and detection time was slower. The eye-tracking data was also used to validate mouse cursor location as a reasonable indication of where eyes are looking. The data showed a relatively strong correlation between eye and cursor and indicated that for the task used the cursor is a suitable tool for collecting data on where an individual is looking.

Keywords: alerting, workload, attention, visual, cursor, eye-tracking

INTRODUCTION

Alerts or alarms come in many forms and can be delivered through any of the sensory modalities. For example, the smell of smoke is conveyed through the olfactory system, a ring of a fire alarm arrives through the auditory modality, the visual system conveys the flashing light of the arriving fire truck, and tactile input of the cell phone vibrating in your pocket as your office mate checks that you are out of the flame engulfed building. These are all examples of ways we are cued to take notice of something our brain deems as relevant. Alerts are designed to make us take notice, to stop what we are doing and pay attention. Alerts give us important information about the world around us, sometimes by their mere presence (e.g., the smell of smoke) and at other times through communicating more detail, like an auditory beep indicating battery power on an electronic device is getting low.

In complex, high workload environments, where individuals use multiple information displays to conduct several tasks concurrently, alerting can be critical. Under pressure of high workload, and in the absence of effective automated alerting, operators may become focused on a task and miss important peripheral information. One example of this kind of environment is the operations room of a navy frigate. The frigate operations room is the processing hub for all the sensory information pulled in from the world outside. Incoming information is collated here and formed
into a global picture that provides the command team with situation awareness to support operational decision making. Operators in the operations room use multiple displays to perform their jobs and they are heavily tasked, being required to quickly read and interpret incoming information while monitoring for new information and changes to existing data. Speed and accuracy are fundamental to timely decision making and the operations room is at times a noisy, demanding, and intense arena.

Because of the concentrated effort and decision making pressure within the operations room operators could be prone to attentional tunneling (Wickens, 2000; 2005) where the focus is so concentrated on a particular task or area of interest on the display that critical information is missed. This is one reason effective modes of automatically alerting operators to specific conditions, states, and points of interest is critical. Most of the alerting in the operations room is currently provided through the auditory modality which is heavily taxed even when the alerting component is excluded. Thus, the visual realm may be an alternate channel for supplementing automated alerting in this complex environment.

**Visual automated alerting**

In a multi-year research program we have been investigating visual forms of alerting to supplement the fully loaded auditory modality in the navy frigate operations room. Two forms of visual alerting have been examined, i) a short red bar 2 centimeters (cm) wide appearing at the side, top or bottom of the operator’s display, and ii) a red border 2 cm wide around the perimeter of the display. Using a three display workstation (see Figure 1), location of the alerts has also been studied, by presenting the alert on a single display (either the left, middle, or right display), or on all three displays simultaneously (Crebolder, Salmon & Klein, 2010; Nakashima & Crebolder, 2010). We have also investigated whether detecting alerts was affected by their appearance as static or flashing (Crebolder & Beardsall, 2008; 2009).

These studies were conducted using a task in which participants were required to detect the visual alerts while performing a secondary task that required their full attention. The results have consistently shown that the bar alert was responded to more quickly than the border form of alert, and, in contrast to other research (Goldstein, 1967; Li et al., 2014) that flashing alerts did not improve response time over static presentation. We have also found that responses to bar alerts were fastest when alerts appeared simultaneously on all three displays, but for border alerts left and middle displays produced response times that were equal to the all display condition.

The consistent result that bar alerts were detected faster than border is a somewhat non-intuitive finding since one might expect that the larger surface area of the border surrounding the entire perimeter of the display would be easier to detect that a smaller short bar specifically positioned on the display. One hypothesis for the result is that the bar alert is detected more rapidly because of its more compact, concise form and its consequent ability to fall into a defined attentional radius or spotlight, as compared to the larger spread of the border. If this is the case, bar alerts should be detected faster when participants are attending to the same display the alert appears on because the alert is being captured in the radius of attention.

The next step in the research program then was to capture in the data the delay the alert appeared on and the display the participant was looking at when the alert appeared. These data could verify whether eyes on the same display as the alert affected detection time. We used the location of the mouse cursor (the display the cursor was on) as a basic estimation of where the participant was looking (Crebolder, 2011). Findings were varied and they did not provide solid evidence for or against the spotlight of attention theory. Results showed that, for the border alert response times were in fact slower when the alert appeared on the left display and the participant was attending to that display. This finding suggests an effect of attentional tunneling, or tunnel vision, whereby the user is so immersed in a task and is focusing on a particular area of the display to the point that they miss other critical information (Wickens, 2000), which in this case was the alert. This is not an uncommon phenomenon in environments where an individual is required to perform multiple simultaneous tasks under time pressure and where the consequences of inaccuracy are severe, such as for air traffic controllers, maritime helicopter flight deck operators, and nuclear power plant operators (Rubinstein & Mason, 1979). The abrupt onset of a visual stimulus has often been used successfully to capture attention, but in some cases, where attentional resources are allocated to other information and tasks the effect of abrupt onset can be significantly reduced (Yantis & Jonides, 1990).

The results further showed that eyes on the right or middle displays resulted in no difference in response time when the alert was on either of those displays. In fact response time was faster when eyes were on the left display. On the other hand, for the bar alert, response to an alert was considerably slower when it appeared on the right display and
eyes were on the left as compared to when eyes were on the right display. Generally the findings showed that alerts presented on all displays were attended to fastest but that in some cases attentional tunneling may have been evident whereby alert detection was hurt by having the alert on the same display that the participant was looking at.

Using the cursor position as an estimate of where participants were looking is a relatively elementary method but one that others have used and regarded as valid (Bieg, Chuang, Fleming, Reiterer, & Bülhoff, 2010; Guo & Agich-tein, 2010; Huang, White & Dumais, 2011). Much of the work using cursor position as an assessment of where people are looking and attending is in the web-based applications research where the interest is in where users are focusing on a webpage or how the cursor is used to help a user read a webpage. For example, looking at eye movement and cursor movement, Rodden and Fu (2007) found a strong relationship, as did Chen, Anderson, and Sohn (2001), in web browsing tasks. There are a number of advantages if the cursor can be used as an estimator of where users are looking as compared to using an eye-tracker. Using the cursor is of no cost as compared to investing in an eye-tracking system, there is no set-up required, a cursor is non-intrusive, and perhaps of most importance in the web-based research, users do not have to be physically present in order to track their web-based behaviour. For our needs, being able to use the cursor as a means of estimating where a participant is looking would be advantageous for all those reasons, but particularly because the cursor is non-intrusive. Eye-tracking equipment attached to a participant’s head may impede their ability to easily turn and attend to all information displays without affecting the sensitive calibration of the system and without discomfort. Crebolder et al. (2010) have examined the relationship between cursor and eye-tracking and found the cursor to be a reasonable assumption of where the eyes are looking (Crebolder, Salmon & Klein, 2010). The present study is an opportunity to provide further validation to that assumption for the kind of detection task used in our studies.

Thus, one objective of the study reported here was to validate use of the cursor against eye-tracking data in a multiple display detection task. The other primary objective was to gather in-depth information about where the participant is looking when an alert appears in order to delve more deeply into the spotlight of attention theory and why a bar alert has consistently proven to be more quickly detected than a border. As in previous experiments, data were collected on human performance as a function of alert Type (border, bar) and alert Location (left, middle, right1).

METHOD

Task

As in previous studies in this series, the participant’s primary task was to detect alerts that appeared randomly on a three-display workstation while performing a secondary task that required detecting and categorizing targets as they appeared on one of the display. The task was designed to emulate in simple form the workstation and the kind of task a sensor operator might do in the frigate operations room.

Apparatus

The task was presented on a workstation consisting of three 20.1” liquid crystal display (LCD) computer monitors, running Windows XP Professional (Service Pack 2), with a single keyboard and mouse input device. The displays were configured with the middle display directly in front of the participant, and one display on either side (see Figure 1).

Participants were fitted with eye-tracker apparatus, Viewpoint® EyeTracker PC-60 SceneCamera System by Arrington Research, which was used to monitor where the eyes were looking while completing the task.

Tasks and Alerts

Primary task - The primary task was to detect alerts that appeared randomly on the displays. Responses to detection were made by pressing the spacebar on the keyboard as quickly as possible.

1 Alert Location level ‘all’ was removed in this study because eyes and cursor could never be on all displays.
Alerts were presented as static (i.e., not flashing\(^2\)) and were either in the form of a Border (a red, 2 centimeter (cm) continuous band around the display perimeter); or a Bar (a red, 2 cm x 10 cm strip placed at the top of the display\(^3\).

Alerts could appear on any single display. The display the alert appeared on (left, middle, right, all) was randomized with the condition that alerts appeared at all possible display locations an equal number of times.

Figure 1: Bar alerts on all three displays. Border alerts were similar in width and colour but surrounded the entire perimeter of the display.

Secondary task

The secondary task was a categorization task that required using the three displays to categorize targets as neutral or hostile. Failure to correctly categorize a hostile target within a limited time period resulted in destruction of the participant’s ship, and a subsequent restart of the task. Details of the task and display set-up were as follow, beginning with the middle display which describes the fundamental task:

**Middle display - Tactical display** - depicted the participant’s ship (ownship), represented as a grey filled circle (60 pixel radius), that remained stationary in the center of the display, as well as other vessels (contacts) represented as yellow triangles that originated in the periphery of the tactical display and advanced toward the ownship in incremental steps at two second intervals. Thus, there were a number of contacts on the screen at once, all moving incrementally toward the participant’s ownship in the center. The task was to categorize contacts on the display as hostile or neutral by moving the mouse cursor over a chosen contact. This action generated attribute information about that specific contact and the information appeared on the left display.

**Left display - Status display** - showed attribute information about each contact after it had been moused-over on the middle, tactical, display. The information on the status display was required to classify contacts as hostile or neutral. Three categories of information were provided:

- **Speed**: Fast = hostile, Slow = neutral
- **Size**: Small = hostile, Large = neutral
- **Weapons on board**: Yes = hostile, No = neutral

Based on the above, a score of >2 attributes in one of the hostile or neutral categories resulted in the contact being classified as such. The participant was required to use the cursor to select and click on one of two text boxes representing hostile and neutral.

---

\(^2\) In some previous experiments a flashing component of 3.333Hz was used (Crebolder & Beardsall, 2008; 2009).

\(^3\) Previous work has examined placing the bar at the side, and on the bottom of the display (Crebolder & Beardsall, 2008; 2009).
Right display - Reporting display - participants entered their response of neutral or hostile in a text box on the reporting display located to the right of the tactical display. The cursor was used to click on the text box before a response could be entered. A correct response resulted in the contact of interest disappearing from the tactical display. An incorrect response required repeating the mouse-over contact process and reviewing the status display attribute information once again, subsequently going through the action of reporting the choice on the reporting display using the cursor to highlight the reporting box before entering a response.

The entire task was very interactive, employing all three displays relatively equally and ensuring that use of the cursor was required on all the displays.

Participants were instructed that there was a time limit to categorize incoming targets and that contacts coming within a pre-determined radius of the ownship would result in the ship being destroyed (accompanied by an audio file sound effect “kaboom” with a JPEG picture of an exploding ship displayed on the tactical display). If the ownship was destroyed the task was paused for 3 seconds after which it was automatically restarted, with contacts once again originating in the periphery of the tactical display and moving incrementally toward the ownship. Participants were instructed that detecting alerts was their primary task and that they were to make their response as quickly as possible after by pressing the spacebar.

Participants

Twenty-four volunteers participated in the study. Because of the eye-tracker headwear participants who typically wore glasses could not participate. Contact lenses were acceptable. All participants reported normal or corrected to normal vision without glasses. The study took approximately 120 minutes to complete.

Procedure

After explanation of the task the participant put on the eye-tracking headwear and the eye-tracker was calibrated. Four blocks of practice trials and 18 experimental blocks (6 alerts per display per block) followed. The alert condition of Type was held constant throughout each block of trials and order of blocks was counterbalanced across participants.

Participants were instructed to reduce head movement as much as possible so that the eye-tracker calibration would remain stable. The restriction was somewhat artificial considering the task involved several displays but it was a necessary request based on limitations of the head-mounted eye-tracking equipment. A key variable of interest was the comparison of performance between bar and border alerts with the addition of the eye-tracker allowing for more in-depth analysis. Thus it was deemed reasonable to accept the limitations of the system. The effect of performance on the task itself would remain to be seen.

Performance measures

Performance measures on several factors were collected but those of most relevance to this study were response time to alerts and position of alert, cursor, and eyes with respect to the display.

Response time was automatically collected via keystroke. For the eye-tracking and cursor data the initial plan was to divide each display into a 3x3 grid and record the grid in which the eyes were focused or the cursor was placed, but this approach proved to be problematic. First, the border and bar would not naturally fit in the same space. The bar would occupy a single grid, while the border would fill eight grid locations on the display with the middle grid unoccupied. We chose to calculate the eye distance as the distance of the eye from the lines making up the border (EyeXdist, EyeYdist) as the best way to measure distance from the eye location to the border location. Secondly, if frame of reference moved (e.g. head movement) then the association with the correct grid location might be compromised. Finally, how best to represent these grid locations in a parsimonious way for analysis was an issue. Rather than using discrete eye-locations, a program was written to capture eye-space as a continuous spatial value in the X and Y domains (independently). The advantage of this approach was that the data could be represented as a function of distance from the alert over time (similar to the way event-related potential waveforms are represented during electroencephalography analysis).
An additional constraint was that eye-location could only be sampled every so often since pure continuous tracking would result in a data file infinitely long. Consequently a manageable sampling period pre- and post- alert was determined. A resolution of 30 millisecond (ms) was used, being shorter than average fixation length (50 ms, cf. Nuthmann, Smith, Engbert, & Henderson, 2010) and about equivalent to two screen refreshes on a 60 Hz display. Eye-location data were recorded continuously 300 ms before alert onset, while the alert was visible to the participant, and continued for 300 ms after the alert disappeared.

Data preparation

Comparison of eye to cursor

The comparison between eye points to cursor points would be easier to interpret if the two were in the same space (i.e. the range of the values were equivalent). By default, eyes were tracked from values 0.000 to 1.000, which, when multiplied by 1000 equalled values of 0 to 1000 (in both the X and Y dimensions). The cursor, on the other hand, was measured by pixel coordinates, which on three 1280 x 1024 displays, were 3840 pixels wide by 1024 pixels tall. In order to compare eye position to cursor position these numbers were translated so that the same values corresponded to the same locations (i.e. a cursor position of 100 equals an eye position of 100). This space-translation was computed by calculating the range of the spaces and calculating how much the data should be shifted to overlap each other. Within excel, a method of calculating both the slope/scalar value and intercept was required in both X and Y dimensions to perform a linear transform \( y = mx + b \) on the data. These values were then applied to generate a position of where the eyes were in pixel space.

Data extraction

For the eye-tracking data a ‘parser’ file was used to manipulate and retrieve relevant data points. The parser was embedded in a Microsoft Excel® file with underlying VBA code/macros that could be accessed through pull-down menus. Because the measure for the eye tracking data was a continuous one the amount of data per trial was much larger than earlier experiments. For example, previous experiments generated approximately 3,000 lines of data while the current experiment, with a sampler of 30 ms, generated about 17,000 lines of data for a 50 minute period. As a result, the cursor versus eye data was limited to the X-dimension since distance measures were not of value in the Y-dimension. Thus the data analysed for the eye cursor was cursor X dimension, eye X dimension, for border and bar.

DATA ANALYSIS

Response time

Response times to alerts detected correctly were analyzed. Cell means for alert Type and alert Location, for each participant, were entered into a repeated measures Analysis of Variance (ANOVA). Significant effects of alert Type \( F(1, 23) = 25.205, p < 0.001, \text{MS}_e = 16597.283 \), and alert Location \( F(2, 46) = 10.889, p < 0.001, \text{MS}_e = 5293.428 \) were present, with no interaction. Overall, responses were significantly faster to the border alert (M=806.38 ms) than to the bar alert (M=874.33 ms). Furthermore, responses were fastest when the alert appeared on the middle display and slowest on the right display. The results are graphed in Figure 2.
The alert Type result is contrary to previous work in the series where the bar alert has been consistently detected faster than the border. As such, further analysis of alert Type is not extended here, but this contradictory finding is discussed in the Discussion section.

**Eye-tracking/Cursor**

A primary objective of the study was to validate the use of cursor location as a means of inferring eye location. In the eye-tracking/cursor analysis the data were collapsed across the variable alert Location. The mean for the eye-tracking and cursor data was calculated over all participants and plotted as a function of alert Type (Border/Bar) and distance from the alert (in pixels) over time, in the X dimension (see Figures 3 and 4).
Not surprisingly, the average distance from the alert for participants’ eyes varied because of differences in eye-calibrations and head positions. To correct for this variance a baseline correction was applied by subtracting each participants’ average distance pre-alert from all their data points. The correction had the effect of re-centering each individual’s average distance at 0, with any significant deviations from 0 representing a significant deviation from the average. Figures 5 and 6 show the data plotted after baseline correction.

![Figure 4: Distribution of mean eye location (eX) and cursor location (mX) for the border alert in the X dimension, over time (ms).](image)

**Figure 4:** Distribution of mean eye location (eX) and cursor location (mX) for the border alert in the X dimension, over time (ms).

![Figure 5: Distribution of baseline corrected mean eye location (eX) and cursor location (mX) for the bar alert in the X dimension, over time (ms).](image)

**Figure 5:** Distribution of baseline corrected mean eye location (eX) and cursor location (mX) for the bar alert in the X dimension, over time (ms).
The Y axis represents pixels, 0 being the alert. The X axis represents time, 0 being alert onset. Data points below 0 on the Y axis indicate that the eyes are moving toward the alert, and above 0 that they are moving away from the alert. For the bar and the border the eyes moved toward the alert and then moved away. Participants looked toward the alert approximately 200 ms after onset, and subsequently the eyes looked away at about 450 ms. Border data was a little more variable showing an initial move away at about 100 ms before changing toward the alert.

Generally, the cursor and eyes were well-aligned, the Pearson product-moment correlation coefficient for the bar alert was $r = .55 \ [p < .0011]$ and for the border $r = .67 \ [p < .0001]$, with a combined correlation of $r = .62 \ [p < .001]$ (see Figure 7). The values represent positive correlations of eye and cursor location, showing that where the cursor moved so did the eyes. Thus, the data support the assumption that the location of the cursor is a reasonable estimate of where the eyes are looking in the task that was used in this experiment and in previous studies in the series.
DISCUSSION

The objectives of this study were to replicate previous findings in a series of studies where one form of alert has been found to be more quickly detected than another, and to validate the use of cursor location against the location of the eyes in a computer-based task. Alerts were presented randomly while participants performed a categorization task that required their full concentration.

The bar alert, which had been detected faster than border in all previous studies in the series, was in fact slower to detect in this experiment. The initial rationale for that result might be that wearing the eye-tracking equipment impeded performance. As such, the attentional spotlight, that may have served to capture the bar in previous studies, was not as effective in this particular experiment. One theory is that the instruction to limit head movement, brought on by the head-mounted eye-tracker, forced participants to expand their attentional beam. Consequently the advantage of the smaller concise bar falling into an attentional spotlight was reduced, resulting in the border being detected faster than the bar. The theory sounds plausible, although the same eye-tracking equipment had been used as part of a previous experiment in which the tracker, being new to the lab, was tested. No eye-tracking data was collected for formal analysis in that experiment but the bar-border comparison was consistent with all previous ones, with the bar being responded to faster than the border (Crebolder, 2011). In that study, participants were not instructed to limit head movement, because the eye-tracker was an add-on for initial testing purposes. As such, expansion of the attentional beam remains a plausible theory and worthy of further investigation.

We had hoped to use the eye-tracking data to delve more deeply into the previously shown superiority of the bar alert to shed light on why such an alert would prove to be faster to respond to than a full border. However, because of the unexpected result it is not possible to analyze the data as hoped. Overall, responses were fastest when the alert appeared on the middle display and slowest on the right display, which are findings in line with previous work.

Nevertheless, it is of interest to use the eye-tracking data to examine behaviour with respect to alert onset. Participants looked toward the alert on average 200 ms after onset and subsequently looked away at approximately the 400 - 450 ms mark. The cursor and eyes were quite well-aligned except during the short toward-away alert behaviour of the eyes. This pattern is to be expected since the eyes are capable of quick saccadic movement and the cursor was not required to contact with the alert. The speed at which the eyes moved toward the alert is consistent with typical reflexive saccade movement (Purves, Augustine, & Fitzpatrick, 2001). Saccades to an unexpected stimulus are usually about 200 ms and last up to about the same amount of time. So the time frame is as expected. Wood (1995) has noted that attention to alerts of any kind does not come from a neutral impartial state but requires actually shifting attention from an existing event to a new and relevant one. Participants looked away from the alert relatively quickly which was most likely a consequence of the task, where lingering on the alert after detection would be detrimental to performance in the secondary task. Note that overall mean response time was 840 ms showing that participants continued to process the alert into the motor action of response (pressing the spacebar) after their focus had moved away from the visual appearance of the alert on the screen. This pattern is also in accordance with the literature in which eye-fixation is followed by processing toward response (Just & Carpenter, 1976; Carpenter, 1998).

CONCLUSIONS

Overall this study demonstrated that cursor location is an accurate estimate of where an individual is looking, at least for the task used here. More investigation is required to understand the implications of wearing a head-mounted eye-tracking system, which may shed light on why the border alert proved to be easier to detect in this particular study.

REFERENCES


Crebolder, J., Salmon, J., & Klein, R. (2010). The cost of location switching during visual alerting: effects of experience and


3468