Visual Change Detection in Digital Military Displays

Paula J. Durlach
Army Research Institute
Orlando, Florida
Paula_Durlach@peostri.army.mil

Jessie Y.C. Chen
Army Research Lab/HRED
Orlando, Florida
Jessie_Chen@peostri.army.mil

ABSTRACT

As the army undergoes its planned transformation, it will rely more and more on humans interacting with complex visual displays. As the complexity of those systems increases, so also does the possibility that important changes in visually presented information will be missed. Research on change detection suggests that people often fail to notice changes in visual displays when they occur at the same time as various forms of visual transients, such as eye blinks or screen flashes, or scene relocation. In this experiment, we investigated change detection using the Force XXI Battle Command, Brigade, and Below display, with no more than 2 map icons ever depicted at one time. Participants were instructed to monitor the display and report changes they noticed as quickly as possible. They were also periodically instructed to conduct certain tasks with the system, such as performing a circular-line-of-sight analysis (CLOS). Changes scheduled for observation included icon appearance and disappearance, changes in icon platform, color/affiliation, or position. We analyzed the observer's ability to detect these different types of changes, as well as the effect of performing concurrent tasks (such as the CLOS) on change detection. Our dependent measures were speed and probability of change detection. Observers were very good at detecting icon appearance, disappearance and color changes; however, they missed more than 20% of the platform changes. Detection of icon movement was dependent on the length of movement, with detection being particularly poor (more than 25% misses) when the icon was in the periphery of the map and the distance moved was small. There was a dramatic effect of performing a concurrent task. When an icon changed simultaneously with the conclusion of a concurrent task, detection dropped to less than 50%. Possible training or design solutions to overcome this type of detection failure need to be considered as the army transforms to the Future Combat System. For example, inclusion of a dedicated change detection tool could explicitly provide operators with changes that occurred during an interruption, and would greatly assist recovery of situation awareness. Instead of an operator needing to figure out what changed, the system should provide that information.

ABOUT THE AUTHORS

Paula J. Durlach has been a research psychologist at the Army Research Institute for the Behavioral and Social Sciences since October 2001. She is based at ARI's Simulator Systems Research Unit located in Orlando, Florida. Having done academic research in the area of associative learning, in 1996 she transitioned to industrial research at a global consumer goods company where she conducted research on the development of food preferences. Her current work concerns training issues relevant to the Future Combat System.

Jessie Y. C. Chen was a post-doctoral research fellow at the Army Research Institute- Simulation Systems Research Unit, Orlando, Florida, when this research was conducted. She earned her Ph.D. in Applied Experimental and Human Factors Psychology from University of Central Florida in 2000. Her research interests include computer game instructions and human change detection performance. She is currently a research psychologist at Army Research Lab/Human Research and Engineering Directorate.

Visual Change Detection in Digital Military Displays

Paula J. Durlach
Army Research Institute
Orlando, Florida
Paula_Durlach@peostri.army.mil

Jessie Y.C. Chen
Army Research Lab/HRED
Orlando, Florida
Jessie_Chen@peostri.army.mil

INTRODUCTION

As the army undergoes its planned transformation, it will come to rely more and more on humans interacting with complex visual displays. Recent research on change detection suggests that people often fail to notice changes in visual displays when they occur at the same time as various forms of visual transients, such as eye blinks, screen flashes, or window relocation (O'Regan, Deubel, Clarke, & Rensink 2000; Rensink, 2002; Rensink, O'Regan, & Clarke, 2000). Therefore, as the complexity of army systems increases, so too does the possibility that important changes in visually presented information will be missed. The purpose of this research was to investigate the extent to which failure of change detection is an issue, using an actual army digital system, the Force XXI Battle Command, Brigade, and Below (FBCB2). FBCB2 is a fielded digital battle command information system that is used across echelons from vehicle commanders up through battle command staff. It provides on-the-move, near real-time command and control information relevant to each of the battlefield functional areas, supports situational awareness, and a common operating picture.

Participants in the present experiment were instructed to monitor the display and report changes they noticed as quickly as possible. They were also periodically instructed to conduct certain tasks with the system, such as performing a circular-line-of-sight analysis (CLOS). Changes scheduled for observation included an increase in Flash Immediate Priority Routine (FIPR) message count, icon appearance and disappearance, and changes in icon platform, affiliation, or position. We analyzed the observer's ability to detect these different types of changes, as well as the effect of performing concurrent tasks (such as the CLOS) on change detection. Our dependent measures were speed and probability of change detection.

Being our first experiment with the system, we used relatively simple displays, and unrealistic changes (i.e., not all the changes examined would be expected to occur in a tactical situation, such as the change of a friendly icon to an enemy icon). Our participants were university students, unfamiliar with the system prior to the experiment. Thus, the results reflect purely the perceptual, attentional, and memorial cognitive processes of the participants, unaffected by any meaningful interpretation of the changes, per se.

METHOD

Participants and Apparatus

The participants were 12 female and 6 male students from the University of Central Florida, who received extra credit in a psychology course for taking part. Average age was 21 years, and all students said they used a computer daily, or a few times a week. They all had normal color vision.

Two computers were used for the experiment, a laptop, which displayed an introductory PowerPoint presentation, and a PC with a 15" color monitor, which was used for the experiment proper. The PC enabled the use of a modified version of FBCB2 (v3.5.4). The modifications included a superimposed program that allowed for the collection of change detection data. An image similar to the display is shown in Figure 1. The size of the map portion of the display was 7 by 5.75 inches, and the map icons were approximately .25 by .25 inches. The Yakima map, provided with FBCB2 was used at 1x resolution (The Figure 1 is not to scale due to processes of reproduction). A "self" icon always appeared in the center of the map; but was never changed. In addition to this self icon there could be 0 or 1 other icon on the map (target icon). When present, that icon was selected from a set of 12, formed by

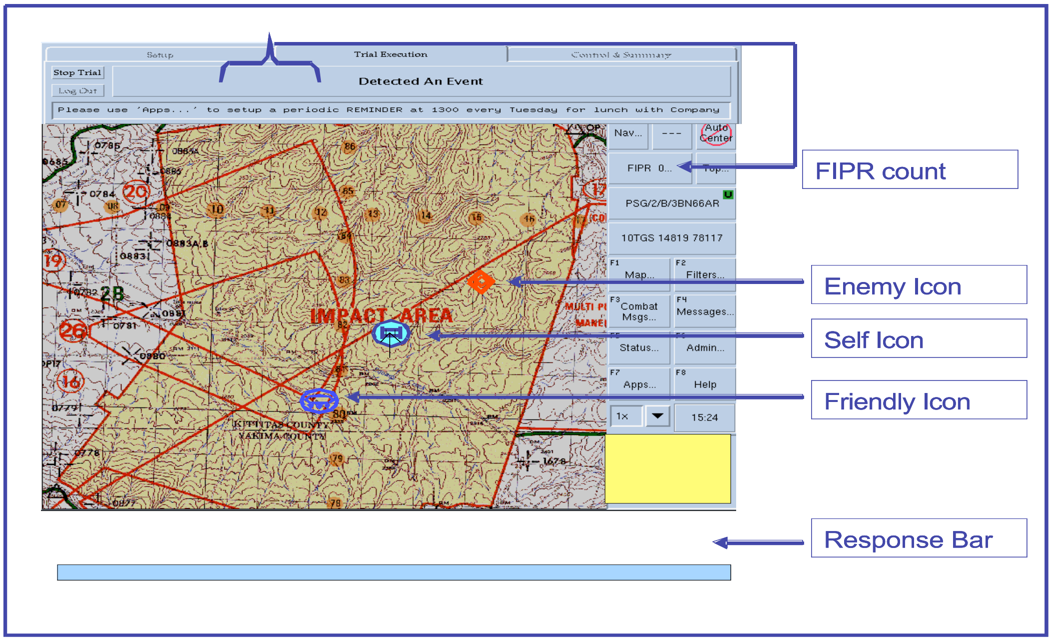


Figure 1. A rendering of the FBCB2 display, as seen by participants, and explanation of certain elements.

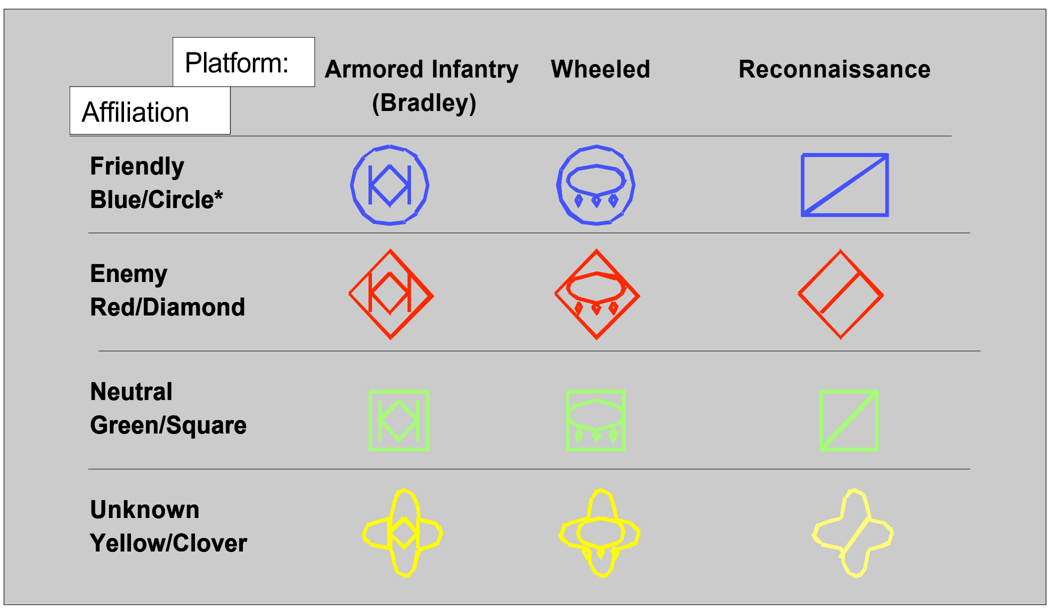


Figure 2. The 12 icons used in the experiment. *Note that while Friendly icons typically appear as circles, in certain cases, such as the Reconnaissance icon, this was not always the case.

combining 4 icon affiliations and 3 platform types, as shown in Figure 2. These icons are native to FBCB2.

In addition, periodically, pop-up windows would appear on the screen requesting the participant to perform one of 3 tasks using the FBCB2 system. These were: (1) perform a circular line of sight analysis, (2) set a periodic reminder, or (3) send a free text message. These tasks will be referred to as distractor tasks (DTs). The participant responded to the DT by clicking the OK button in the pop-up window, and performing the DT. This involved clicking on one of the buttons on the right of the FBCB2 display, and responding with information (either through button clicks or text entry) requested from a series of windows the FBCB2 system uses for these tasks.

When open, the windows involved in the DTs covered various portions of the map display.

Experimental Design

Participants were instructed to look for 6 general change types: the appearance of a new icon on the map, the disappearance of an icon from the map, a position change of an icon on the map, a color change of an icon on the map, a platform change of an icon on the map, and an increase in the FIPR count (see Figure 1). Only one change ever occurred at a time, and changes never occurred less than 5 seconds apart. The actual time between changes varied, depending on the time required by the FBCB2 to implement the change and the actions of the participant.

For purposes of analysis, we conceived of the map display as consisting of a central area (3.75 by 3 inches) and a peripheral area (any location outside the center). Thus, combining position with the 12 icons produced 24 potential icon stimuli. 24 blocks of changes were composed. Each block began with an icon appearance and ended with an icon disappearance. Between these events, 9-14 other changes were scheduled; these could be changes in icon position, color, or platform, or an increase in FIPR count. On average, each block contained 3 color changes, 2 platform changes, and 2 position changes. Every possible kind of specific icon change occurred at least once over the 24 blocks. About 80% of the blocks also contained a FIPR count change.

Each block also contained one DT. For 13 DTs, a minimum of 5 seconds elapsed between the participant's completion of the DT and the next scheduled change, whereas for 11 of the DTs, an icon change was visible either during performance of the DT or simultaneously with the conclusion of the DT (closing of the last window of the task). These will be referred to as distractor trials. Distractor trials included 4 color, 4 position, and 3 platform changes.

In order to keep session length approximately an hour, each participant received only 12 of the 24 blocks. Four subgroups observed different block combinations, such that each of the 24 blocks was observed by half of the participants. Order of blocks was randomized for each participant.

Procedure

A brief taped explanation of the study was given to participants, and then they read and signed the informed consent form. This was followed by a period of training, during which the participant viewed a PowerPoint presentation that reviewed the purpose of the experiment, the FBCB2 display and icons, the changes they needed to watch for, and how to respond when they observed a change. They were also given instruction and practice at performing the DTs. After this tutorial, the participants performed 24 practice trials, which included 3 appearances, 4 color changes (1 trial per color), 3 platform changes (1 trial per platform), 5 position changes, 3 FIPR count increases, 3 distractor tasks, and 3 disappearances. During practice the experimenter corrected any misunderstandings; and the participant was allowed to refer back to the tutorial slides as they wished.

After training, the experimenter answered any questions and collected demographic information such as date of birth and computer experience. The experiment proper then began.

When the participant observed a change they were to use the mouse to click on the response bar as quickly as possible (see Figure 1). This opened a choice window (and ceased all experimental timing). The participant chose from a list of 32 possible changes (grouped by type), with changes defined by color. For example, in the Move grouping, the list was Blue unit moved, Red unit moved, etc., and in the Platform grouping the list was Blue unit changed, Red unit changed, etc. Thus, participants did not have to memorize symbol meaning; nevertheless, they were allowed to have a figure available to refer to, similar to Figure 2. The choice window also gave participants the option not to make any choice, in case they had clicked the response bar by mistake. After making their selection, they clicked on a response bar when they were ready to resume the experiment. This resumed experimental timing and closed the choice window.

The program recorded the time from when a change appeared on the screen, until the participant hit the response bar (in the case of correct detections). False alarms (reporting changes that did not actually occur) were also recorded. When a change occurred, the participant could report the change up until the point at which a change of the same type was scheduled again, at which point, that change was considered missed.

Each participant completed 12 blocks of trials, after which they were debriefed and thanked for their participation.

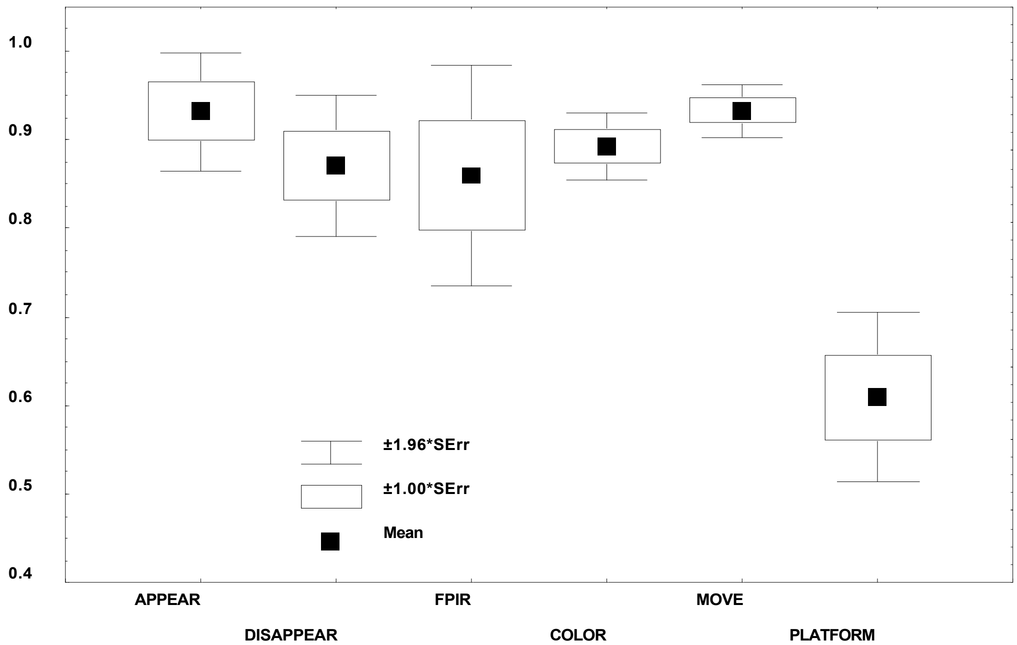


Figure 3. Mean sensitivity for the 6 changes used in the study. SErr=Standard error of the mean.

$$\text{Sensitivity} = \frac{\#\text{ correct detections} - \#\text{ false alarms}}{\#\text{ actual changes} + \#\text{ false alarms}} \qquad (1)$$

RESULTS

Figure 3 shows detection sensitivity for each kind of change participants were asked to report, exclusive of distractor trials. Sensitivity was calculated as outlined in Equation 1. A maximum score of 1 indicates that every change was detected and no false alarms were committed. A score of less than one can result if changes were undetected, if changes were reported that did not occur (false alarms), or both. For example, icon appearance and disappearance were detected 100% of the time; but because some false alarms were also made, sensitivity was less than 1. Detection of color changes was also quite high. Correct detections for position changes and FIPR count changes were around 90%, but with greater variation across participants. On average, participants missed about 25% of the platform changes, and for this type, there was also considerable spread in the distribution.

Analysis of the sensitivity measure indicated a significant difference across types [$F(5,85) = 9.2$; $p < .001$]. A Tukey HSD analysis indicated that a significantly lower sensitivity for platform changes, compared with all the other change types; these failed to differ among one another.

Figure 4 shows mean reaction time for correct detections, for each kind of change participants were asked to report. As can be seen in the figure, reaction times hovered between 1.5 and 3.5 s, except for FIPR count increases. Mean time to detect changes in FIPR count was just below 8 s, and there was a lot of variation across participants.

Position and platform changes were further analyzed, to determine if particular stimulus conditions associated with detection failures. For position changes, we examined the potential effects of icon position before the move, and distance moved. Distances were either near or far. Near moves were ones in which the icon stayed within the same screen area and within a 4.5° visual angle. Far moves always moved an icon from one screen area to another (i.e., central to peripheral or peripheral to central), and were all greater than a 5° visual angle. As shown in Figure 5, percent of correct detections varied as a function of distance moved and starting position.

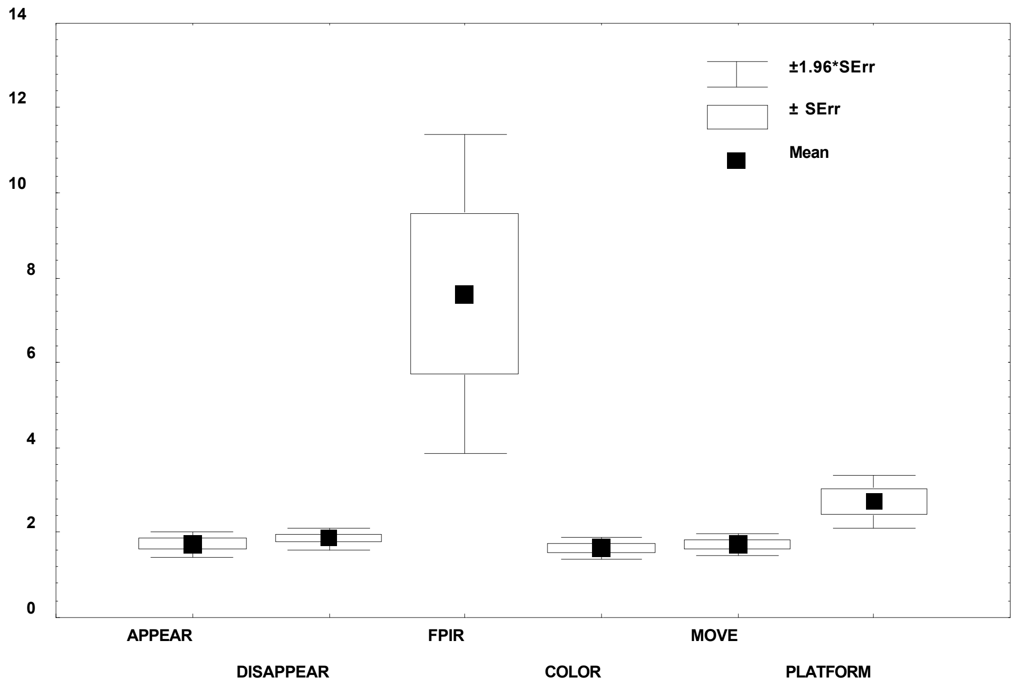


Figure 4. Mean reaction time for correct change detections. SErr=Standard Error of the mean.

There was a main effect of distance moved [$F(1,17) = 33.03$; $p < .001$], and a distance by start position interaction [$F(1,17) = 4.64$; $p < .05$]. Detection was poorer when the new position was near vs. far; and detection was especially poor when the move was near, and the start position was peripheral.

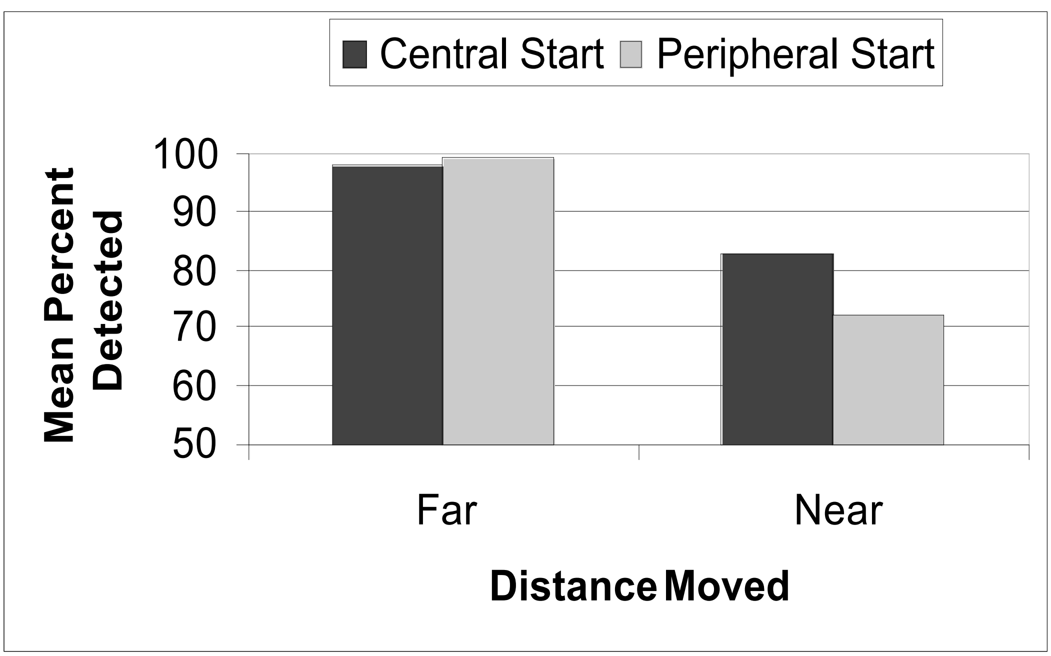


Figure 5. Mean percent of position changes detected, depending on starting location and distance moved.

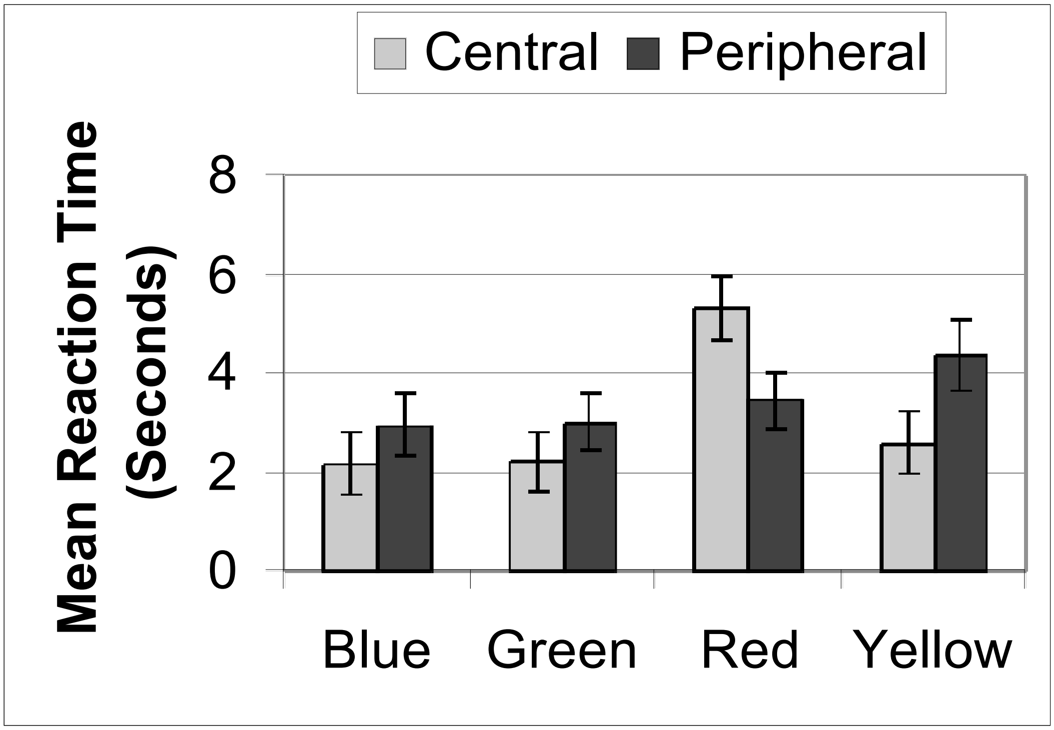


Figure 6. The interaction of icon position and icon color on speed to detect platform changes.

The percent of platform changes correctly detected depended on the pair of platforms involved in the change. As shown in Table 1, detection rate was better than 85% if one of the platforms involved in the change was the reconnaissance platform; however, if the change was between Bradley and wheeled, detection was only 53.4%. Detection rate was analyzed according to platform pair and icon location. The effect of platform pair was highly significant [$F(2,34) = 20.52$; $p < .001$]. There was also an effect of platform pair on reaction time, reaction time being significantly slower for the wheeled/Bradley pair [contrast analysis $F(1,318) = 13.11$; $p<.001$]. The RT analysis was carried further, by adding color as a factor. This analysis revealed a significant color by location interaction [$F(3, 318) = 3.157$; $p = .025$]. The nature of this interaction is depicted in Figure 6. Platform changes for red icons took longer to detect when they were in the center than when they were in the periphery, whereas yellow icons took longer to detect in the periphery than in the center.

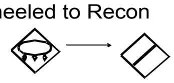
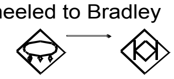
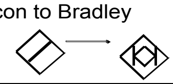
Platform Change	Mean Percent Detected	Mean Reaction Time (ms)
Wheeled to Recon 	87.6	2407.5
Wheeled to Bradley 	53.4	3544.4
Recon to Bradley 	85.9	2411.0

Table 1. Results for the 3 different kinds of platform changes (ignoring direction of change).

Performance on distractor trials vs. nondistractor trials is shown in Figure 7. There were 11 distractor trials, each involving a unique change, and each experienced by 8 or 9 of the participants. The data shown for the nondistractor trials were taken from trials with exactly the same kinds of changes, but without a distractor. For example, if one of the distractor trials involved the change of centrally located reconnaissance platform from yellow to red, the corresponding data from trials with a change of a centrally located reconnaissance platform from yellow to red, but with out a DT, were taken as the comparison data for the nondistractor trials. This produced 11 pairs of detection rates, which were analyzed using a matched-pairs Wilcoxon T-test.

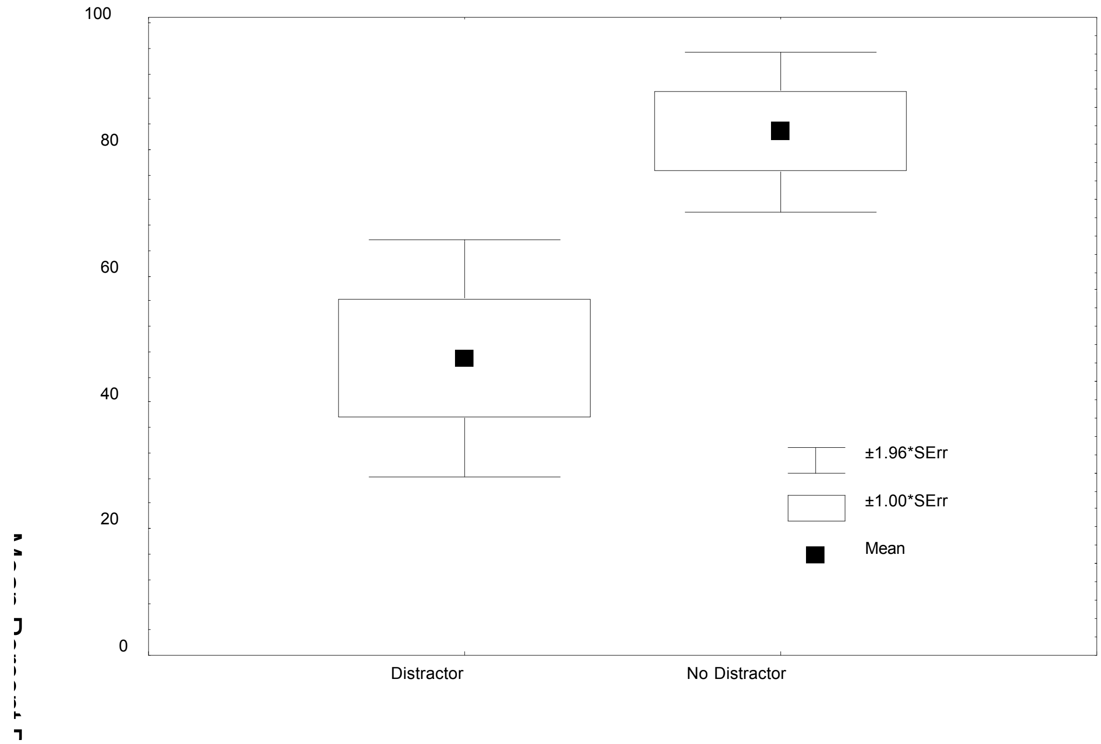


Figure 7. Mean percent of changes detected for trials with and without a distractor. SErr = Standard error of the mean.

The analysis confirmed that which is obvious in Figure 7. There was a significant negative effect of the distractor on detection rates [T=0; $p < .01$]. Because detection rates were so poor on distraction trials, there was insufficient data to analyze reaction times.

DISCUSSION

The purpose of this research was to examine people's ability to detect changes in the FBCB2 display. With only one target icon on the map at a time, detection rate for icon appearance and disappearance was 100%; though it should be noted that occasional false alarms also occurred. For increases in FIPR count, icon color changes, position changes, and platform changes, there was greater variation across participants, and detection rates varied between 100 and 39%. For the case of position changes, the distance moved and the initial position of the icon on the map were both important factors. The data indicate that people missed more than 25% of the near position changes in the peripheral area of the map. With respect to tactical employment, such moves are potentially very important, as icon moves occurring in near real-time are likely to be relatively small (depending on the map resolution being used), and in the periphery, if they are not friendly units.

For the case of platform changes, it appears that the similarity of the platform symbols plays an important role. The platform representations are small. The more complex they are, the harder they are to distinguish. Consequently, it is more difficult to detect a change between two complex platform representations than between a complex one (e.g., Bradley) and a simple one (e.g., Reconnaissance). Moreover, distinguishing platform changes also depends upon an interaction of the icon color and map characteristics. Recall that red platform changes in the central map area were more slowly detected than in the peripheral map area. Inspection of the map used (see Figure 1) suggests the explanation: map features included the text, "impact area" written in red letters, several red intersecting lines, and several black numbers within orange circles in the central area of the map. These features more than likely made discerning the platform details of red icons more difficult. Were a different map background used, the results likely would have been different. The icon colors and platform representations used by FBCB2 are determined by Military Standard 2525B (1999). Standards also exist for military maps (e.g., Army Field Manual 3-25.26, 2001). These standards were not necessarily chosen in order to maximize discriminability, or to stand out clearly from standard map colors and features. It is suggested that some consideration be given to this factor in the future development of digitized systems, maps, and terrain databases.

When a color, move, or platform change occurred co-incidentally with the conclusion of a DT, detection rates were on average only about 50%. These are exactly the conditions under which psychological research suggests that people are most vulnerable to missing changes; and our results confirm this in the context of an actual fielded military system. Given that the army is committed to future use of complex visual information displays, in which users will be multi-tasking, map-resizing, and opening and closing several windows, can anything be done to lessen the user's vulnerability to detection failure? Durlach (in press), discussed potential interface design features and training solutions to ameliorate visual change detection failure. Among the design innovations she most favored was a change detection tool, which the user could query, and which would highlight all the changes that had occurred within a specified time period (see also, Smallman and St. John, 2003). Training to illustrate the benefits of such a tool would be required, as people tend to overestimate their ability to detect changes (Levin, Momen, & Drivdahl, 2002), and might otherwise fail to utilize it; but once operators become aware of the utility of a change detection tool, they could get information about changes explicitly from the system, instead of having to figure them out. This should greatly aid recovery of situation awareness after an interruption.

Durlach (in press) also suggested that practice at change detection itself, with the military systems to be used, could ameliorate change detection failure with those systems. Some psychological research indicates that change detection can improve with practice and feedback (Austen and Enns, 2000; Williams and Simons, 2000). Other forms of training might also improve change detection; Green and Bavelier (2003) found that action video game playing led to better performance on a range of visual skills. We suggest that more consideration be given to the perceptual and attentional factors that influence change detection in the design of future military systems, and that further research be conducted into the means that can be employed to train the user to better detect changes.

ACKNOWLEDGEMENTS

We would like to acknowledge Calvin Lombard and Paul Fleetwood for their assistance in software development and Laticia Bowens for help with data collection. We would also like to thank Michael Green for advice on sensitivity analysis.

REFERENCES

Durlach, P.J. (in press). Change detection and its implications for complex monitoring and control systems design and operator training. *Human-Computer Interaction*.

Field Manual 3-25.26 Map Reading and Land Navigation.(2001). Retrieved May 28, 2003, from the General Dennis J. Reimer Training and Doctrine Digital Library website: http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/3-25.26/toc.htm

Green, C.S & Bevelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537.

Levin, D. T., Momen, N., & Drivdahl, S. B. (2002). Change blindness: the metacognitive error of overestimation change-detection ability. *Visual Cognition, 7,* 397-412.

Mil-Std 2525B: Department of Defense Interface Standard: Common Warfighting Symbology (1999). Retrieved May 28, 2003, from the General Dennis J. Reimer Training and Doctrine Digital Library website: http://www.adtdl.army.mil/cgi-bin/atdl.dll/fm/3-25.26/toc.htm

O'Regan, J. K., Deubel, H., Clarke, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7,* 191-211.

Rensink, R.A. (2002). Change Detection. *Annual Review of Psychology*, *53*, 245-77.

Rensink, R. A., O'Regan, J. K., & Clarke, J. J. (2000). On the Failure to detect changes in scenes across brief interruptions. *Visual Cognition, 7,* 127-145.

Smallman, H.S. & St. John, M. (2003). CHEX (Change history explicit): New HCI concepts for change awareness. *47th Annual Meeting of the Human Factors and Ergonomics Society.* , October, Denver, Co.