Improving Recovery from Multi-task Interruptions Using an Intelligent Change Awareness Tool

Harvey S. Smallman & Mark St. John
Pacific Science & Engineering Group
9180 Brown Deer Road
San Diego, CA 92121
smallman@pacific-science.com
stjohn@pacific-science.com

Abstract

Detecting changes to complex monitoring tasks is important for situation awareness, yet it is surprisingly difficult. Interruptions due to multi-tasking or other distractions exacerbate this problem. Current display technologies do not provide much support for detecting and identifying significant situation changes and therefore do not provide much support for the recovery of situation awareness following interruptions. Instead, situation displays typically only represent the current situation, which forces users to rely on their own ability to extract changes by cognitively integrating events over time. Sustained situation awareness can be greatly improved by augmenting users' abilities with automated situation change detection. However, the design of the presentation of change information turns out to be crucial and hinges on the issue of how to alert and inform the user effectively without distracting from other important on-going tasks. Here, we review two empirical studies that compare different design approaches to this challenge including visual alerts and an intuitive Instant Replay tool. The experiments use a naval warfare task in which users monitor a busy airspace. The results, however, strongly favor a new set of display concepts that we developed called CHEX (Change History Explicit). CHEX augments the human attentional system with a set of intelligent change detectors whose output is logged in a re-configurable table format that is linked back to the situation display. CHEX is extremely effective both for maintaining situation awareness when monitoring a situation as well as when recovering situation awareness following an interruption.

1 Introduction

A wide range of operations tasks, from airspace management to industrial plant control and disaster relief, involve monitoring a situation display over time and addressing issues that arise. This monitoring process involves detecting significant changes and then responding to them in a timely manner. Current situation display technologies often fail to support change detection and change awareness by representing only the current state of the evolving situation. By representing only the current state, users are required to rely on their own abilities to extract changes by remembering and integrating states of the situation across time.

A burgeoning cognitive science literature is documenting the difficulty of unassisted change detection. Humans are unable to spot significant changes to simple scenes. This poor performance obtains both when users are momentarily distracted by glancing away from the display, and even when they are actively looking at the display, but are attending elsewhere. This change blindness is surprisingly severe and has been documented in many contexts (see Rensink, 2002 for a review), including air warfare (DiVita, Obermayer, Nugent, & Linville, 2004; Smallman & St. John, 2003) and is only just being recognized as a significant problem for human computer interaction (Varakin et al., 2004).

Longer distractions, due to multi-tasking for example, can exacerbate change blindness because memory for prior situation states will decay and the situation will evolve over the course of the distraction, making integration and change detection more difficult. Our contention is that recovering from interruptions and detecting significant situation changes requires computer-assisted augmentation because unassisted recovery and change detection is so poor.

Designing the augmentation involves addressing at least two separate issues. First is the issue of detecting significant changes to information on the display, such as changes in aircraft kinematics. The second issue is the re-presentation of those changes to users in a useful, effective fashion. We do not address the detection issue in this paper, although one straight-forward method would be to develop change detection agents, or "sentinels," that would monitor the displayed information. The criteria for defining an information change that was important enough to bring to a user's attention could be set through consideration of the task domain and through user interviews (for example, "I only care when aircraft altitude drops below 10,000 feet"). More sophisticated criteria could be based on changes in the assessed threat levels of aircraft (see St. John, Smallman, & Manes, 2005 for an example). Some difficulties arise concerning the rate of changes and sizes of changes that should elicit an alert. The human attentional system, of course, possesses transient detectors in abundance for detecting changes to a scene. Rather than define changes based on task analysis, artificial change sentinels could imitate the well-characterized spatio-temporal filtering properties of the Magno-cellular retino-cortical pathways that are thought to subserve visual transient detection in the brain (e.g., Shapley, 1990). It is intriguing to reflect on the findings of the Change Blindness literature which show that the central bottlenecks in human attentional processing lead to the loss of so much of the signals of these transient detectors. In this sense, the goal of our research has been to augment the attentional system by usefully maintaining signals from artificial transient detectors and re-presenting their signals for later exploitation by the user.

The second issue, how to re-present the change information, is the focus here. We review two experiment that first document poor change detection and interruption recovery using conventional displays and then go on to show that augmenting the user's natural abilities with automatic change detection can lead to significantly better performance (Smallman & St. John, 2003; St. John, Smallman & Manes, 2005). In a simplified version of naval air warfare, undergraduate participants monitored a busy airspace to detect significant changes to aircraft such as course, speed, and electronic emissions. Response times and percent correct detections were measured. Augmented change detection, however, is not sufficient for good performance. The design of the re-presentation method is crucial. A poorly designed augmentation may provide no value or even degrade performance below baseline.

2 Experiment One

In the first experiment, a baseline display that showed only the current situation was compared with three methods for automatically detecting and displaying aircraft change information (see Smallman & St. John, 2003, for details). In the first alternative, changes were logged into a static text table next of the map display. The chronologically sorted table listed the time and nature of each significant change. The second alternative also included the table, but it added red "circle alerts" around each changed aircraft on the map. One circle for each change listed in the table. The third alternative also included the table, but the table was linked to the map so that selecting a change entry in the table would highlight the changed aircraft on the map, and vice versa. This table could also be sorted by type of change and type by aircraft as well as chronologically, at the discretion of the user in order to facilitate different tasks. This interactive, linked table of automatically detected changes was called the CHEX tool (Change History EXplicit) because it explicitly and automatically identified changes to the situation rather than leaving users to their own abilities. Aircraft moved around the display at realistic rates over time, occasionally changing their behavior in ways that could be threatening to "own ship", the blue dot at the of the display in Figure 1. The participants' task was to detect and identify the most threatening changes as quickly as possible. They had to report those changes both while monitoring the display and when returning after a minute processing a secondary, mental arithmetic task. Aircraft density was also varied between subjects as high (40 aircraft) and low (13 aircraft).

Our hypothesis was that the CHEX table would 1) facilitate change detection and interruption recovery relative to the baseline display, and 2) it would do so more effectively than either the chronological table or the alert circles. We predicted that the problem with alert circle was that they could quickly clutter the map with distracting and relatively uninformative alerts (all alerts look the same). The CHEX tool, on the other hand, provided a less distracting solution that did not clutter the map, but was effectively linked to it, and that provided more descriptive and better organized information.

Participants in the baseline condition, who used comparable display tools to those available to operators in Navy Combat Information Centers, correctly identified only 34% of critical aircraft changes. Interestingly, these same participants exhibited an over-confidence in their ability to do the task since their confidence ratings dropped a full

26% after having performed the task. This overconfidence in unaided change detection replicates and extends the meta-cognitive underestimation of change detection ability found by Levin et al. (2000).

Augmenting the display by adding an explicit log of automatically detected changes improved performance. However, the static, separated table provided no alerts, and it left to the user the difficult problems of aggregating changes and correlating information in the table to the appropriate symbols on the map. The addition of change alert circles to the map did not further improve performance.

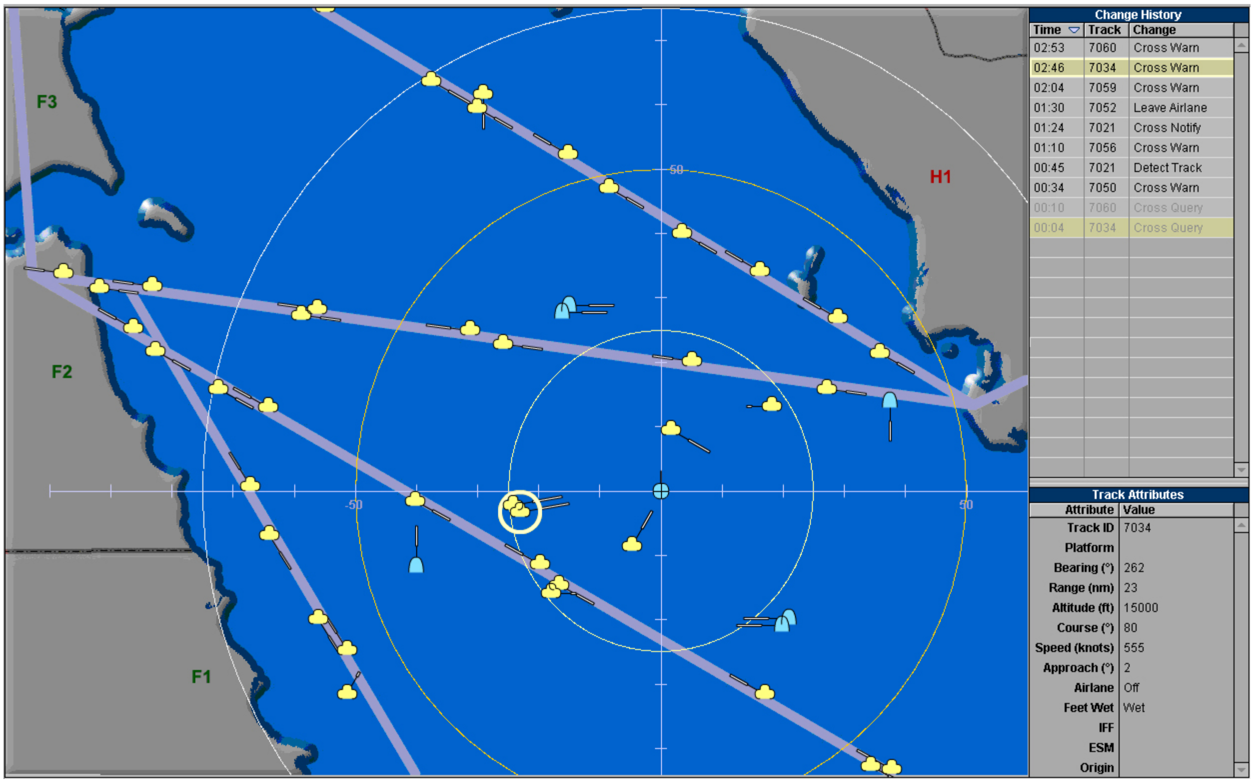

Figure 1. Screenshot of map and CHEX table (upper right) from experiment two (experiment one display was similar). Each row of the table describes one significant change to an aircraft (time, aircraft identification number, and change type). Selecting a row causes the aircraft to be selected on the map and its detailed kinematics information to be displayed in the Track Attribute list (lower right).

CHEX did away with the need for alerts on the map by dynamically linking the output of the (now flexibly sortable) Change History Table. Critical aircraft could be quickly found in the Table, selecting one automatically highlighted each of its changes, and the automatic linking to the map removed the need to search the map to find the location of the relevant aircraft. In the dense display condition, these benefits resulted in an 80% improvement in change identification speed compared with the baseline condition and a 40% improvement in speed over alert circles. Further, whereas participants lowered their confidence in performing the task after using the baseline tools, participants raised their confidence after using the CHEX tools. These benefits were found both for monitoring and recovering changes and across display densities, although there was an interaction between density and tool-type, with CHEX providing impressively density-independent support for change awareness.

Experiment one demonstrated that this serious deficit in change detection and identification can be dramatically improved by augmenting the display with intelligent change awareness tools, but the degree of improvement differed depending on the specific design of the tools.

3 Experiment Two

The first experiment demonstrated that one apparently sensible alternative to CHEX, alert circles, turned out to provide substantially inferior support, especially for longer interruptions. Another intuitively appealing alternative to CHEX is Instant Replay. With Instant Replay, users could replay interrupted periods at high speed to quickly perceive changes. Instant Replay's appeal seems to rest on its familiarity and realistic re-presentation of the temporal sequence of the interrupted situation. We have been developing a theory about user preference for realistic displays and HCI that is maintained in the face of poor performance called *Naïve Realism* (Smallman & St. John, 2005). Naïve Realism predicts that the preference for temporal realism in the Instant Replay tool is misguided.

In a naval air warfare task similar to the one used in Experiment one, we compared two versions of replay against three alternative tools, including CHEX and a baseline condition similar to the baseline used in experiment one (See St. John, Smallman, & Manes, 2005a, for details). Following interruptions of variable length, participants could replay the interrupted period of the simulated task. The map "re-wound" to the beginning of the interruption and then played forward at 20x speed until it caught up with current time. Participants could initiate a replay whenever they desired. The results again supported CHEX over the alternative designs.

In its basic form, Instant Replay simply replays the interrupted sequence and changes remain implicit. With this form of Instant Replay, changes remain difficult to detect, as difficult, in fact, as during real-time monitoring. In both the baseline condition and in the basic Instant Replay conditions, only 60% of the significant changes were detected and reported by participants.

Augmenting the display by adding explicit detection again helped significantly. However, providing the explicit change detection information within the context of a replay tool proved substantially inferior to providing that information in the form of the CHEX tool. Change detection approached 80% using the augmented Instant Replay tool, but participants achieved 100% detection using CHEX. In the augmented replay condition, changes to aircraft were marked on the map as they occurred during the replay sequence. They were removed when the map returned to real-time speed at the end of the replay sequence in order to avoid cluttering the map during real-time monitoring. Again participants could initiate a replay whenever they chose. In the CHEX condition, changes were again logged in an interactive table that was linked to the map.

Relying solely on the Baseline display to detect changes is insupportable, in spite of the commonly held belief that changes are easy to detect. Rather, the Baseline display produced high miss rates, high error rates, and generally slow response times. Replay offered little support for detecting or identifying changes over the Baseline condition. In fact, the Basic Replay tool may even be worse than nothing since it can add a delay as it replays the temporal sequence. The CHEX tool does not offer the same level of realism as the Instant Replay tools, but it offers superior functionality: it provides easily accessible information for both the detection and identification of changes and without cluttering the map.

4 Conclusions

Change detection is difficult enough that augmenting human cognition with automatic detection of significant events is warranted, in spite of the potential reliability and trust issues that accompany automation (but see St. John, Smallman, & Manes, 2005b; St. John, Smallman, Manes, Feher, & Morrison, in press). The method for presenting change detection information, however, can have profound implications for the ultimate success of augmented tool. Furthermore, intuitions about the effectiveness of simple alerts on a map or Instant Replay turn out, on close inspection, to be misguided. The explicit representation of situation changes within an interactive table display proved to be the best design.

5 References

DiVita, J., Obermayer, R., Nugent, W., & Linville, J. M. (2004). Verification of the change blindness phenomenon while managing critical events on a combat information display. *Human Factors, 46,* 205-218.

Levin, D.T., Momen, N., Drivdahl, S.B., & Simons, D.J. (2000) The metacognitive error of overestimating change-detection ability. *Visual Cognition*, *7,* 397-412.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53,* 245-277.

Shapley, R. (1990). Visual sensitivity and parallel retinocortical channels. *Annual Review of Psychology, 41*, 635-658.

Smallman, H. S. & St. John, M. (2003). CHEX (Change History EXplicit): New HCI concepts for change awareness. In *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 528-532). Santa Monica, CA: Human Factors and Ergonomics Society.

Smallman, H. S. & St. John, M. (2005). Naïve Realism: Misplaced faith in the utility of realistic displays. *Ergonomics in Design*, submitted.

St. John, M., Smallman, H. S., & Manes, D. I. (2005a). Recovery from interruptions to a dynamic monitoring task: The beguiling utility of instant replay. Paper submitted for presentation to the *49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, Sept 26-30, 2005.
.

St. John, M., Smallman, H. S., & Manes, D. I. (2005b). Assisted focus: Heuristic automation for guiding users' attention toward critical information. This meeting.

St. John, M., Smallman, H. S., Manes, D. I., Feher, B. A., & Morrison, J. G. (2005). Heuristic automation for decluttering tactical displays. *Human Factors*, in press.

Varakin, D. A., Levin, D. T., & Fidler, R. (2004). Unseen and unaware: implications of recent research on failures of visual awareness for human-computer interface design. *Human-Computer Interaction, 19,* 389-422.