Pip and Pop: When Auditory Alarms Facilitate Visual Change Detection in Dynamic Settings

Cindy Chamberland[1], Helen M. Hodgetts[1, 2], Benoît R. Vallières[1], François Vachon[1], & Sébastien Tremblay[1]

[1] Université Laval, Québec, Canada [2] Cardiff Metropolitan University, UK

Dynamic and complex command and control situations often require the timely recognition of changes in the environment in order to detect potentially malicious actions. Change detection can be challenging within a continually evolving scene, and particularly under multitasking conditions whereby attention is necessarily divided between several subtasks. On-screen tools can assist with detection (e.g., providing a visual record of changes, ensuring that none are overlooked), however, in a high workload environment, this may result in information overload to the detriment of the primary task. One alternative is to exploit the auditory modality as a means to support visual change detection. In the current study, we use a naval air-warfare simulation, and introduce an auditory alarm to coincide with critical visual changes (in aircraft speed/direction) on the radar. We found that participants detected a greater percentage of visual changes and were significantly quicker to detect these changes when they were accompanied by an auditory alarm than when they were not. Furthermore, participants reported that mental demand was lower in the auditory alarm condition, and this was reflected in reduced classification omissions on the primary task. Results are discussed in relation to Wickens' multiple resource theory of attention and indicate the potential for using the auditory modality to facilitate visual change detection.

INTRODUCTION

Within complex, dynamic, and safety-critical work environments, the detection of changes is an important cognitive function since change might signal a vulnerability or potential threat that requires a timely response. This change might be a fairly obvious indication that action is needed (e.g., a warning light turning to red), or it might be seemingly innocuous (e.g., within surveillance, a slight change in the speed of an individual/vehicle), but that nonetheless may require further investigation. Because such environments are often also characterized by time pressure, uncertainty, distraction, and high workload, the occurrence of important changes can sometimes be overlooked – a phenomenon known as change blindness (Rensink, O'Regan, & Clark, 1997) or inattentional blindness (Mack & Rock, 1998). This failure to notice relevant change can occur within a static visual scene, but is compounded in dynamic and multitasking situations. When critical changes are missed, situation awareness is compromised which can impact upon decision-making quality (e.g., Varakin, Levin, & Fidler, 2004) and may leave operators prone to human error.

Graphical user interface add-ons have been developed to support change detection, although sometimes they can come with their own challenges. Detecting a change relies on memory and attentional processes, whereby the visual system must capture visual transients associated with the difference between pre- and post-change states. It can be difficult to capture these transients when a situation is changing frequently and unpredictably, or if attention is diverted away by an interruption. One tool developed to improve situation awareness within an air-warfare context, the Change History Explicit (CHEX; see Smallman & St. John, 2003), aims to ease the burden on the human operator by automatically detecting changes in the environment and logging them within an interactive table. This external aid supplements the operator's limited memory and attention by providing a permanent record of changes that can be consulted should an event have been missed. However, recent research suggests that within a multitasking environment that requires the completion of many different subtasks (rather than just the detection of changes), such a tool may actually increase workload and detrimentally impact the primary task due to the sheer amount of information displayed on the interface (Vallières, Hodgetts, Vachon, & Tremblay, 2012). Thus, when introducing technology it is important that it fits with the available attentional resources of the operator.

Given the high burden already placed upon the visual modality in detecting visual changes, one alternative might be to exploit the auditory modality as a means to support change detection, for example by introducing an auditory tone to coincide with a visual change. Although this would not specifically indicate what the change was, within a multitasking environment in which an operator has several subtasks to complete, it would serve to inform the operator that a change had taken place and that attention should perhaps be directed towards the visual display. In line with Wickens' (2008) multiple resource theory of attention, such an auditory alert would be preferable to a visual notification system due to the already high processing load in the visual domain. Accordingly, performance should be facilitated if information to be processed is not constrained to a single sensory modality but is rather distributed across modalities and/or codes.

Moreover, the use of an auditory alert may extend beyond a simple warning effect, with studies from experimental psychology illustrating the perceptual benefits of visual-auditory integration for visual search. The pip and pop effect (van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008) for example, demonstrates how a spatially non-specific auditory signal can facilitate the identification of a target visual item within a cluttered and changing environment. When the auditory and visual signals are in temporal synchrony, they are integrated perceptually, generating a salient emergent feature that captures attention. That is, an auditory 'pip' helps the target visual item 'pop' out from amongst competing single-modality distractors, resulting in faster identification. This effect has further been extended to moving stimuli, with participants better able to detect moving dots that abruptly change orthogonal direction when accompanied by a tone (Staufenbiel, van der Lubbe, & Talsma, 2011). These findings invite the possibility that a spatially uninformative tone may facilitate visual change detection within a dynamic and multitasking task environment.

While an auditory alert certainly has the potential to enhance visual change detection, if operators fail to consciously detect the tone itself then cross-modal benefits will be unlikely. The processing of elements in an auditory scene is obligatory and not affected by explicit head/eye movements (Edworthy, Loxley, & Dennis, 1991), but still this auditory information may not reach conscious awareness. Although less well-researched than its visual counterpart, inattentional deafness can be an issue with operators failing to detect seemingly obvious features of an auditory stream (Vitevitch, 2003). This phenomenon is influenced by the allocation of attention (Eramudugolla, Irvine, McAnally, Martin & Mattingley, 2005), and within a complex, dynamic and multitasking environment there is no guarantee that the alert will be successfully perceived. Moreover, it has been shown that presenting auditory alarms among other irrelevant sounds to promote change detection may tend to bias threat evaluation judgment towards increased perceived hostility (Vachon, Tremblay, Nicholls, & Jones, 2011), which can be disastrous in environments related to security.

In the current study, we explored the impact of using visual-auditory cues to promote change detection in a realistic command and control environment. We used a microworld in which participants have to monitor a radar screen representing the airspace around the ship, evaluate the threat level of every aircraft moving in the vicinity of the ship based on a visual list of parameters, and take appropriate defensive measures against hostile aircraft. Critical changes consist of an aircraft passing unexpectedly from a non-threatening to a hostile status. Participants must detect such changes and perform the appropriate self-defense action. In past studies using this microworld, each critical change was accompanied by a change on the radar screen (i.e., a change in the direction and/or the speed of the aircraft) that makes it visually noticeable (e.g., Vachon, Vallières, Jones, & Tremblay, 2012; Vallières et al., 2012). Using unimodal cues to promote change detection, the authors reported that a significant number of critical changes remained undetected by participants. In the present study, we examined whether presenting an auditory alarm in temporal synchrony with visual cues could enhance visual change detection. In line with past studies on the pip and pop effect, auditory alarms always indicated a critical change and no irrelevant sound was presented throughout the simulation. The effectiveness of the visual-auditory cues was evaluated according to a holistic approach, i.e. by evaluating the impact of auditory alarms not only on detection performance, but also on the capacity to support the monitoring of a complex dynamic situation. Therefore, the efficiency of auditory alarms was determined by the percentage of detected changes (as compared to that found in the condition where no auditory alarm was presented) as well as by its impact—positive or negative—on threat evaluation performance.

METHOD

Participants

Thirty-two students from Université Laval (18 females, 14 males, *M* = 23.84 years old, *SD* = 4.44) took part in the experiment in exchange of a small honorarium. All reported normal or corrected-to-normal vision and audition.

Microworld

The experiment used the Simulated Combat Control System (S-CCS) microworld (see Hodgetts, Vachon, & Tremblay, 2014; Vachon et al., 2012) run on a PC. This microworld provides a functional simulation of threat evaluation and combat power management processes (i.e., response planning, execution, and monitoring) that can also be generalized to other command and control situations. The visual interface includes three parts; a black radar screen, a list of parameters relating to the aircraft selected, and a set of action buttons (Figure 1). At the center of the screen is the ownship with multiple aircraft moving in the vicinity in real time. An aircraft is represented by a white dot surrounded by a green square with a line attached; this line indicates the direction of the aircraft, and its length is proportional to the aircraft speed. Sixteen 4-min scenarios were created for the experiment. Each scenario involved from 24 to 28 aircraft in total (not presented together at the same time on the radar) and was different from the others in terms of parameter values and aircraft trajectories.

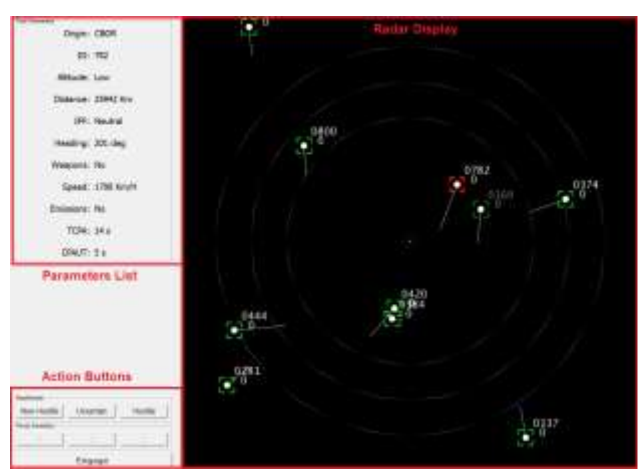

Figure 1. Screenshot of the Simulated Combat Control System (S-CCS) microworld visual interface

Task

Participants' assignment was threefold: (1) to determine the threat level (non-hostile, uncertain, hostile) of all the aircraft on the radar screen; (2) determine the threat immediacy of hostile aircraft (i.e., how long until they hit the ship); and (3) engage a missile to neutralize a hostile aircraft. Clicking with the mouse on an aircraft icon would turn the surrounding square red and display five parameters relating to that aircraft in the parameters list: (a) country of origin (ADRK, WEIV, CBOR; ADRK = threatening), (b) altitude (low, high; low = threatening), (c) identification friend or foe (IFF) (friend, neutral, foe; foe = threatening), (d) weapons detected (yes, no; yes = threatening), and (e) military emissions (yes, no; yes = threatening). Other parameters were also displayed that were not part of the threat assessment task (e.g., heading, distance, and speed). Participants were asked to classify each aircraft as either non-hostile (0 or 1 threatening parameters), uncertain (2 or 3 threatening parameters), or hostile (4 or 5 threatening parameters), and click on the corresponding action button. For aircraft classified as hostile, further actions were required because they were programed to hit the ship. Participants were asked to classify the threat immediacy of those hostile aircraft (on a scale of 1 to 3; < 15 s, 15-30 s, or > 30 s, respectively). Participants should then choose to launch a missile in defense, taking into account the probability of hitting and destroying the hostile aircraft (the radar screen was divided into hit-accuracy zones: 0%, 25%, 50%, and 100% according to the distance from the ship). Clicking on the 'engage' button launched a missile with a 2-s delay, and only one could be airborne at any one time.

Change detection

When an aircraft appeared on the radar (either at the beginning or during the ongoing scenario), it was either non-hostile or uncertain. However, aircraft parameters could change over time, so it was necessary to check back at the parameters of classified aircraft on a regular basis in order to reassess threat level. Aircraft status could turn from non-hostile to uncertain, from non-hostile to hostile, of from uncertain to hostile. When an aircraft status changed to hostile, it was considered a critical change because hostile aircraft were programmed to hit the ship. All other changes regarding aircraft parameters were considered non-critical. Each 4-min scenario included a total of 8 unexpected critical changes and 25 non-critical changes. Critical changes were separated by a minimum of 15 s and were accompanied by a change in aircraft speed (increase) and/or in aircraft direction (heading in the direction of the ship) visible on the radar. Participants were required to detect these critical changes in order to further investigate and ultimately protect the ship. A critical change was considered detected only if the aircraft was classified as hostile within the 10 s following this change.

In half of the scenarios, critical changes were also accompanied by an auditory alarm in order to promote change detection. Those alarms were designed following Patterson's (1982) recommendations for optimal detection of high priority warnings. They consisted of five 1000-Hz tones (44.1 kHz sample rate, 16 bit, mono) of 100-ms duration (including a 20-ms fade-out to avoid clicks), each separated by a 100-ms interval. Auditory alarms always signal a critical change (no false alarms) and were presented via headphones at ~75 dB.

Procedure and Design

Participants read though a tutorial describing the context of the simulation and the tasks to execute (including detecting critical changes). They were told that critical changes were always accompanied by a change in aircraft speed and/or direction visible on the radar, and by auditory alarms in some scenarios. Participants were asked first to complete two training sessions, each including four 3-min scenarios. They then completed four experimental sessions, each including four 4-min scenarios, for a total of 16 scenarios in the whole experiment. After each scenario, participants were presented with the mental demand scale of the NASA-TLX subjective workload questionnaire (Hart & Staveland, 1988). They had to indicate how much mental and perceptual activity was required in the simulation on a scale from 1 (low) to 10 (high).

Auditory alarms were presented in two of the four experimental sessions (selected according to a Latin square design). Those two sessions were blocked and presented in a counterbalanced fashion across participants. The experimental session lasted about two hours.

RESULTS

We first examined how auditory alarms were efficient in promoting critical status change detection and then looked at the influence of these alerts on the level of mental demand perceived by participants during the simulation and the performance at the threat evaluation task.

Change detection performance

The percentage of detected changes was averaged across scenarios included in each experimental condition (no alarm, alarm). As shown in Figure 2A, the percentage of detected changes was greater for scenarios with auditory alarms (88%) than without (81%). That result was confirmed by a paired-sample t-test, $t(31) = -4.64$, $p < .001$, indicating that auditory alarms were successful in promoting change detection. Subsequent analyses indicated that 92% of undetected critical aircraft were not classified by participants (a classification omission) while only 8% were misclassified, $\chi^2(1, N = 622) = 431.39$, $p < .001$.

We were also interested in examining whether participants were faster at detecting critical changes when both auditory and visual cues were presented rather than visual cues only. To this end, we computed the mean detection time (in ms) according to whether auditory alarms were presented or not during critical changes (see Figure 2B). A paired-sample t-test revealed that participants were faster at detecting critical changes in scenarios with alarms (3,940 ms) than without alarms (4,504 ms), $t(31) = 5.68$, $p < .001$.

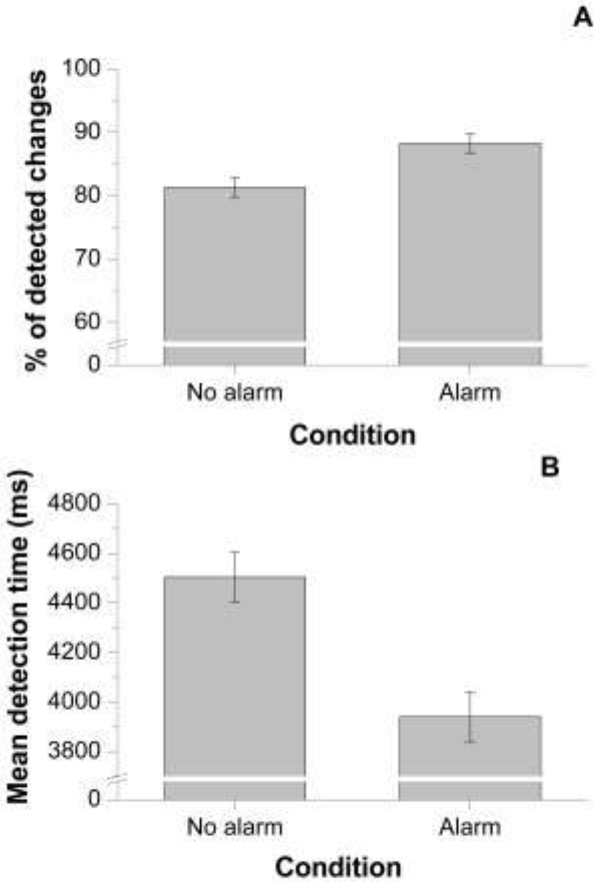

Figure 2. Percentage of detected changes (panel A) and mean detection time in ms (panel B) as a function of whether auditory alarms were presented or not during critical changes. Error bars represent the standard error of the mean.

Mental demand

The score at the mental demand scale was averaged across scenarios included in each experimental condition (no alarm, alarm). As shown in Figure 3, the score on the mental demand scale was lower for scenarios with alarms (5.1) than without (5.7). This difference was found to be significant, $t(31) = 2.93$, $p = .006$, suggesting that participants felt less mentally loaded when critical changes were cued by auditory alarms.

Threat evaluation

In order to examine the influence of auditory alarms on threat evaluation, we compared the number of classification omissions made in scenarios with and without alarms. Classification omission was computed each time participants failed to classify or reclassify any aircraft (critical or not) during the scenario. The analysis of classification omissions was preferred over classification accuracy since previous analyses showed that omissions were highly more frequent than misclassification. Results showed that 47% of classification omissions observed in the experiment were made during scenarios with alarms compared to 53% during scenarios without alarms (see Figure 4). An adjustment chi-square test confirmed that classification omissions were less frequent in scenarios with auditory alarms than without, $\chi^2(1, N = 1,410) = 6.27$, $p = .01$.

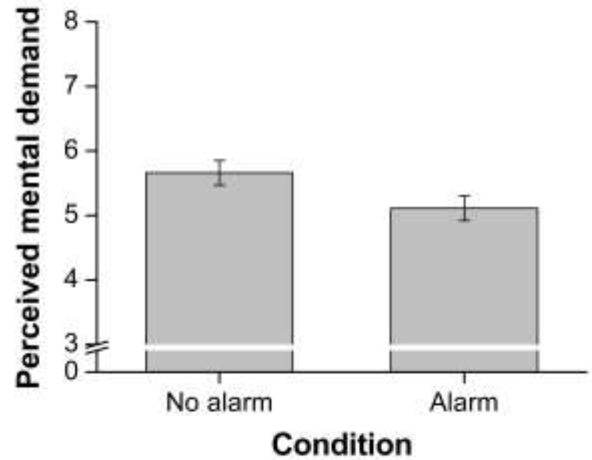

Figure 3. Mean score at the mental demand scale as a function of whether auditory alarms were presented or not during critical changes.

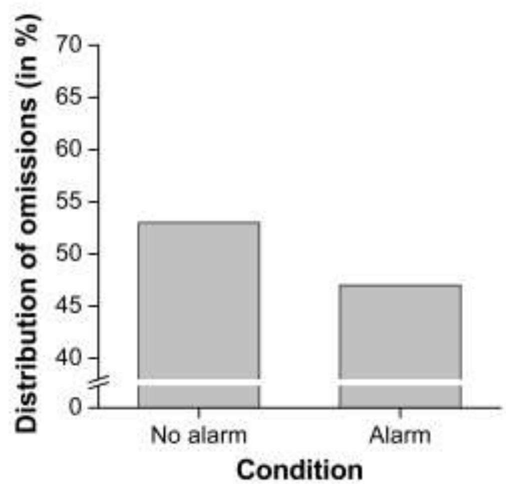

Figure 4. Distribution of classification omissions in percentage as a function of whether auditory alarms were presented or not during critical changes.

DISCUSSION

Accompanying auditory alarms were able to increase the speed and the success with which visual changes were detected on the radar screen. As well as improved detection performance, the auditory alert was also subjectively perceived as presenting more optimal task conditions with participants reporting significantly lower mental load with the alarm than without. This reduced mental demand associated with change detection also had a positive effect on the aircraft classification task, with fewer omissions made in the alarm condition.

Previous research has shown that although on-screen support tools such as CHEX can be useful for identifying changes in a change-detection-only task, the increased demand in the visual modality can overload the operator and actually hinder overall performance when change detection is just one of a number of subtasks that must be performed (Vallières et al., 2012). Our finding that an auditory alarm promotes visual change detection – as well as benefiting performance on the primary task – suggests that utilizing the auditory modality may represent a more optimal distribution of workload. This is in line with the multiple resource theory (Wickens, 2008), which purports a benefit to performance when limited modality-specific resources share the processing demands of the task.

The current results are also consistent with findings from experimental psychology which demonstrate how a spatially uninformative but temporally congruent tone can improve the detection of change within a visual scene (the so-called pip and pop effect; van der Burg et al., 2008). When an auditory stimulus (or vibrotactile stimulus; Ngo & Spence, 2010) temporally coincides with a visual change (e.g., color, motion), a multisensory integration process is thought to occur pre-attentively to increase the salience of the visual target. Evidence from neuroscience studies find that multisensory integration is associated with enhanced neural firing (Stein, Jiang, & Stanford, 2005) resulting in lower detection thresholds than for unisensory stimuli.

One suggestion is that a deviant tone has a subjective ‘freezing’ effect on the visual configuration for a short period, increasing the perceived duration of visual targets (Vroomen & de Gelder, 2000), and thus temporally extending information sampling such that changes are more readily detected. This is supported by eye movement studies that reveal how sounds have a freezing effect on scanning behavior; fixation duration increases with the occurrence of a tone, and the mean number of saccades decreases (Zou, Müller, & Shi, 2012). An auditory tone has also been associated with a more efficient search, with participants more likely to scan away from already searched areas than repeat already-covered ground (Zou et al., 2012).

The fact that clear benefits were observed in the alarm condition suggest that participants were able to perceive the warning tone easily, and indicate that inattentional deafness was not an issue in the current multitasking setting. This is perhaps because the tone was the only auditory stimulus played during the experiment, and so having established a basic effect of the tone, further study would need to determine whether the same benefits could be observed within a more complex auditory environment. Many command and control settings are characterized by background sounds and conversation between other personnel that might be critical to the state of the mission, but that can nonetheless be disruptive to other features of the task (Hodgetts et al., 2014). Nonetheless, our study showed that multisensory integration is a promising avenue to support change detection and decision making in complex and dynamic environments.

REFERENCES

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors*, *33*(2), 205-231. doi:10.1177/001872089103300206

Eramudugolla, R., Irvine, D. R., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates ‘change deafness’ in complex auditory scenes. *Current Biology*, *15*(12), 1108-1113. doi:10.1016/j.cub.2005.05.051

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.) *Human mental workload*. Amsterdam: North Holland Press.

Hodgetts, H. M., Vachon, F., & Tremblay, S. (2014). Background sound impairs interruption recovery in dynamic task situations: Procedural conflict? *Applied Cognitive Psychology*, *28*(1), 10–21. doi:10.1002/acp.2952

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. In R. D. Wright (Ed.), *Visual attention* (Vol. 8, pp. 55-76). New York, NY: Oxford University Press.

Ngo, M. K., & Spence, C. (2010). Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, Perception, & Psychophysics*, *72*(6), 1654–1665. doi:10.3758/APP.72.6.1654

Patterson, R. D. (1982). *Guidelines for auditory warning systems on civil aircraft (Paper No. 82017)*. London: Civil Aviation Authority.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373. doi:10.1111/j.1467-9280.1997.tb00427.x

Smallman, H. S., & St. John, M. (2003). CHEX (Change History EXplicit): New HCI concepts for change awareness. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (Vol. 47, pp. 528-532). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/154193120304700358

Staufenbiel, S. M., van der Lubbe, R. H. J., & Talsma, D. (2011). Spatially uninformative sounds increase sensitivity for visual motion change. *Experimental Brain Research*, *213*(4), 457-464. doi:10.1007/s00221-011-2797-6

Stein, B. E., Jiang, W., & Stanford, T. R. (2005). Multisensory integration is single neurons of the midbrain. In G. Calvert, C. Spence, & B. E. Stein (Eds). *The handbook of multisensory processes.* MIT press, Cambridge.

Vachon, F., Tremblay, S., Nicholls, A. P., & Jones, D. M. (2011). Exploiting the auditory modality in decision support: Beneficial "warning" effects and unavoidable costs. In *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (Vol. 55, pp. 1402–1406). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1071181311551292

Vachon, F., Vallières, B. R., Jones, D. M., & Tremblay, S. (2012). Nonexplicit change detection in complex dynamic settings: What eye movements reveal. *Human Factors*, *54*(6), 996–1007. doi:10.1177/0018720812443066

Vallieres, B. R., Hodgetts, H. M., Vachon, F., & Tremblay, S. (2012). Supporting change detection in complex dynamic situations: Does the CHEX serve its purpose? In *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (Vol. 56, pp. 1708–1712). Santa Monica, CA: Human Factors and Ergonomics Society. doi:10.1177/1071181312561342

van der Burg, E., Olivers, C. N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053-1065. doi:10.1037/0096-1523.34.5.1053

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*(4), 389-422. doi:10.1207/s15327051hci1904_9

Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 333-342. doi:10.1037/0096-1523.29.2.333

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human perception and performance*, *26*(5), 1583-1590. doi:10.1037/0096-1523.26.5.1583

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449-455. doi:10.1518/001872008X288394

Zou, H., Müller, H. J., & Shi, Z. (2012). Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision*, *12*(5):2, 1-18. doi:10.1167/12.5.2