Neuroergonomics and Cognitive Engineering, Vol. 167, 2025, 93–104
https://doi.org/10.54941/ahfe1006131

Optimizing Multimodal Alarm Design for Attention Allocation in Discrete Monitoring Tasks

Wenzhe Cun[1], Hao Fan[1,2], Jianjie Chu[1,2], and Dengkai Chen[1,2]

[1]Ningbo Institute of Northwestern Polytechnical University, Ningbo, Zhejiang, 315103, China
[2]Northwestern Polytechnical University, Xi'an, Shanxi, 710072, China

ABSTRACT

This study examined the efficacy of six alarm modalities (visual, auditory, visual-auditory, visual-tactile, auditory-tactile, and visual-auditory-tactile) under four workload conditions in discrete monitoring tasks. A high-simultaneous task scenario was used to assess operators' attention control modes, focusing on shifts between exogenous and endogenous control. Twenty-two participants performed a simulated emergency task, with performance metrics such as hit rates, errors, choice response times, workload, and user experience collected. Results showed that the visual-auditory-tactile alarm provided superior performance, demonstrating its robustness in achieving superior effectiveness across different workload scenarios. Additionally, auditory alarm was more effective than visual alarm in modulating attention control modes. These findings underscore the potential of multimodal alarms to enhance operator performance and demonstrate that auditory alarm effectively modulate ACMs, thereby ensuring system safety in discrete monitoring tasks.

Keywords: Discrete monitoring task, Multimodal alarm, Workload

INTRODUCTION

In modern human-machine systems, operators are increasingly responsible for monitoring multiple processes rather than performing manual tasks (Causse et al., 2025), facing challenges related to performance, workload, and situational awareness (Ferraro & Mouloua, 2021). Discrete monitoring tasks, common in fields like aviation, air traffic control, nuclear power, and healthcare, require operators to monitor multiple independent and intermittent routine operations. These operations are scheduled at specific intervals, and operators must also respond to sudden, unpredictable alarms. In certain cases, routine operations and alarms may occur concurrently, placing considerable demands on attention allocation and cognitive resource management (Clark & Bustamente, 2008).

Alarm systems in discrete monitoring tasks typically rely on visual and auditory alarms; however, these modalities can be overwhelmed in high workload conditions (Dehais et al., 2014; Cun et al., 2024). Multimodal alarms, integrating visual, auditory, and tactile cues, show promise in enhancing operator performance under high-demand conditions (Oskarsson et al., 2012; Lu et al., 2013), but their effectiveness remains debated (Jakus et al., 2015). Studies have suggested that while multimodal alarms can enhance attention and reduce response times, they may also lead to increased processing time (Lu et al., 2012) and pose challenges in integrating semantically inconsistent or asynchronous signals (Wickens et al., 2011). Thus, comparing the effectiveness of uni-, bi-, and tri-modal alarms under varied workload conditions is essential, particularly in discrete monitoring tasks. However, little evidence from controlled experiments under different levels of workload is available to determine the effectiveness of multimodal alarms.

A key challenge in discrete monitoring tasks is to prioritize alarms over routine operations (Zirk et al., 2020), which can be hindered by automatic stimulus-response associations due to ingrained stimulus-response associations of training and long-term operation. Different sensory modalities can modulate attention control modes (ACMs) (Oskarsson et al., 2012), shifting between exogenous (stimulus-driven) and endogenous (task-driven) control. For example, auditory presentations of routine information may dominate attention via auditory preemption (Oskarsson et al., 2012) (exogenous control), overshadowing visual alarms. In contrast, augmenting visual alarms with auditory and tactile channels—both of which have strong attention-capturing characteristics (Krausman et al., 2007)—may induce competition between routine signals and alarm signals, thereby facilitating the shift to endogenous attention and prioritizing critical alarms. Understanding how alarm modalities influence ACMs is crucial for optimizing alarm designs to improve task prioritization. However, despite the potential, the impact of alarm modalities on ACMs in dual-task settings remains unclear.

This study aims to investigate the effectiveness of uni-, bi-, and tri-modal alarms under increasing workload levels in discrete monitoring tasks and examine how alarm modalities influence ACMs in simultaneous task scenarios.

MATERIALS AND METHODS

Participants

Twenty-two participants (11 males, 11 females) aged 22–30 from Northwestern Polytechnical University took part in the study. All had normal vision, color perception, and no hearing or touch impairments. Informed consent was obtained, and the study was approved by the Ethics Committee of the University.

Protocol

Experimental Materials

A discrete monitoring task scenario, based on a flight take-off task, was developed. Participants performed routine operations sequentially as per an operation manual and responded to sudden alarms. The routine operating procedure is outlined in Table 1. During the task, the airspeed display starts at 0 and increases to 150 knots, at which point the altitude begins to rise. Participants follow these changes in airspeed and altitude to complete the routine operations.

Table 1: Routine operations.

Occurrence Time	Task Operations
The 0th s	Push the mouse to 40% of the scale.
The 11th s	Maintenance/installation time to complete
The 14th s	Monitor the airspeed reaching 40 knots, then gently push the mouse forward.
The 61st s	Verify the airspeed reaches 80 knots and announce "eighty check".
The 96th s	When "V1" sounds, verify whether the airspeed reaches 128 knots. If the speed is less than 128 knots, announce "airspeed too slow"; if the speed is 128 knots, announce "normal speed"; if the speed is greater than 128 knots, announce "airspeed too fast".
The 110th s	Monitor the airspeed reaching 150 knots and move the mouse to 60% of the scale.
The 113th s	Announce "gear up" when monitor the height is 15 meters.
The 134th s	Announce "roll mode" when monitor the height is 121 meters.
The 168th s	Press the "P" key on the keyboard when monitor the height is 2000 meters.

When a specific alarm was detected and confirmed by the participants, they were required to complete the corresponding alarm response accurately and quickly (see Table 2). The time window for key-press action was 8s.

Table 2: Four types of alarms and their corresponding operation.

Alarm Types	Alarm-Response Task Operation
ENGINE FAIL	Press the 'S' key (key-press action), then execute decision action, which involves verbally reporting "ENGINE FAIL", current airspeed, altitude, and the decision (either "terminate take-off" if airspeed <128 knots or "continue take-off" if airspeed >128 knots).
WINDSHEAR	Press the 'F' key, then execute decision action, which involves verbally reporting "WINDSHEAR", current airspeed, altitude, and the decision (either "terminate take-off" if airspeed <128 knots or "continue take-off" if airspeed >128 knots).
SPEEDBRAKE FAIL	Press the 'H' key, then execute decision action, which involves verbally reporting " SPEEDBRAKE FAIL", current airspeed, altitude, and the decision (either "terminate take-off" if airspeed <150 knots or "continue take-off" if airspeed >150 knots).
PULL UP	Press the 'K' key, then execute decision action, which involves verbally reporting "PULL UP", current airspeed, altitude, and the decision (either "terminate take-off" if airspeed <150 knots or "continue take-off" if airspeed >150 knots).

Participants were instructed to prioritize alarm responses over routine operations when both occur simultaneously, ensuring accuracy before considering response times.

The display used in this experiment was derived from flight take-off task in aircraft cockpit, in which the information on routine operations and visual alarm are exhibited (see Figure 1). The visual alarm was displayed as a flash pattern (500ms on, 200ms off, repeated for 8 seconds). Auditory alarm was synthesized according to the Boeing 787's cabin alarm language (40–60 dB), with 200 ms intervals between consecutive alarms, lasting for 8 seconds.

A tactile alarm was delivered via a vibration module placed on participant's left wrist (see Figure 1). The vibration frequency was set to 250 Hz (Salzer et al., 2011). Pulse duration was 500 ms with a 200 ms inter-pulse interval (Meng & Spence, 2015; van Erp et al., 2015), with a total duration of 8 seconds.

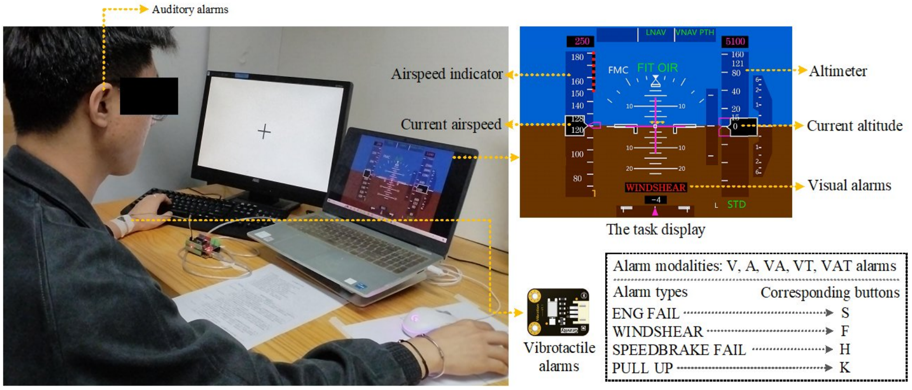

Figure 1: The apparatus and experimental scenario.

Experimental Design

A within-participant design was used, with four workload levels (Baseline, Low, HighSucc, HighSimu) and six alarm modalities (visual (V), auditory (A), visual-auditory (VA), visual-tactile (VT), auditory-tactile (AT), and visual-auditory-tactile (VAT)). A total of 24 trials were conducted and trials were counterbalanced across participants, with alarm types appearing randomly.

Experimental Procedures

The experiment was conducted in a controlled room (23 ± 2 °C, 30–40 dB noise). After an initial briefing, participants underwent a practice phase, ensuring at least 90% correct responses in routine and alarm tasks before proceeding to the main experiment. A total of 24 trials were performed, each lasting 2 minutes 50 seconds. After each trial, participants completed NASA-TLX and UEQ questionnaires and rest for 2-minute. The experiment lasted approximately 3 hours, and participants received a monetary incentive for top performance.

Measures

Independent and Dependant Variables

The independent variables were workload levels and alarm modalities.

1) Four levels of workload when alarm occurred

Alarms occurred at four specific points in the routine operation timeline (see Figure 2), generating four distinct workload conditions, as follows:

a) Baseline: the alarms occurred at the 46th second, with no specific routine operation to be performed.

b) HighSimu (high simultaneous): the alarms appeared simultaneously with the "V1 verifying task", where participants were instructed to prioritize alarm-response tasks before continuing with the routine operation. This meant that participants needed to quickly perceive and process the routine tasks in their minds, then execute the alarm-response task. After completing the alarm-response task, participants verbally reported their airspeed judgment.

c) HighSucc (high successive): The alarms appeared 3 seconds after the "V1 verifying task".

d) Low: The alarms occurred 3 seconds after the altitude reached 121 feet, and participants announced "roll mode" at the 135th second.

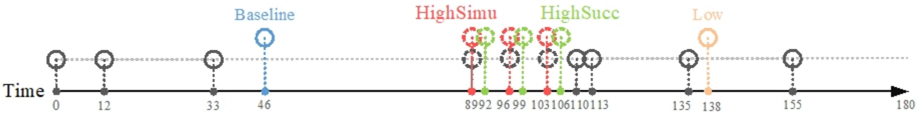

Figure 2: Illustration of four load levels. The solid gray circles represent routine operations, while the solid colored circles indicate alarms. The dashed colored circles represent the three time points when alarm may occur under the HighSimu and HighSucc workload conditions.

The workload values for the alarm-response task were derived from the VACP method (Rusnock & Borghetti, 2018) and four levels of workload—Baseline, HighSimu, HighSucc, and Low—were established. This workload refers to the load generated by the resource consumption of the perception, cognition, and psychomotor from routine operations at the time the alarms occur. These levels were based on verbal anchors and their corresponding values most closely matching the routine operations in VACP Scales (Rusnock & Borghetti, 2018), as seen in Table 3. The time interval of 3 seconds for Low and HighSucc workloads coincided with the completion of routine operations before the alarms.

Table 3: Four levels of workload of alarm-response task.

Alarm occurrence	Perception	Cognition	Psychomotor	Workload Levels
46th s	None (0)	None (0)	None (0)	Baseline
89th/96th/104th s	Interpret sematic content (3.0)	Judgment (4.6)	None (0)	HighSimu
92th/99th/107th s	Visually inspect (3.0)	Judgment (4.6)	Simple speech (2.0)	HighSucc

Continued

Table 3: Continued

Alarm occurrence	Perception	Cognition	Psychomotor	Workload levels
138th s	Interpret sematic content (3.0)	Automatic (1.0)	Simple speech (2.0)	Low

2) Alarm modalities

The alarm modalities examined were: visual (V), auditory (A), visual-auditory (VA), visual-tactile (VT), auditory-tactile (AT), and visual-auditory-tactile (VAT) alarms, which combine varying combinations of visual, auditory, and tactile channels presented simultaneously. Tactile-only alarm was not used due to their limited effectiveness (Ho et al., 2014) in conveying detailed alarm information.

The dependent variables in this study included: 1) performance in the alarm-response task, 2) perceived workload for the six alarm modalities, 3) user experience of the six alarm modalities, and 4) attention control mode (ACM) for the six alarm modalities in the HighSimu condition. The specifics are as follows:

1) Performance in the alarm-response task

Performance was measured by the number of hits, errors, misses, and choice response times (CRTs) of the key-press action. Hits refer to accurate key-press actions, errors to inaccurate actions, and misses to instances where no key-press occurred within the 8-second time window. CRTs refer to the time from alarm onset to the accurate key-press action.

2) Perceived workload of six alarm modalities

Perceived workload was assessed using NASA-TLX (Hart, 2006).

3) User experience of six alarm modalities

User experience was assessed using the User Experience Questionnaire (UEQ) (Schrepp et al., 2014), which included eight items: "obstructive/-supportive," "complicated/easy," "inefficient/efficient," "confusing/clear," "boring/exciting," "not interesting/interesting," "conventional/inventive," and "usual/leading edge." The last three items were excluded due to lower relevance for alarm modality evaluation. The remaining five items were labelled UEQ-1 to UEQ-5.

4) ACMs of six alarm modalities in HighSimu condition

In the HighSimu condition, the "V1" sound coincided with the alarm, and participants was required to prioritize the alarm-response task. In the other workload conditions, the "V1" sound occurred without an accompanying alarm. Upon hearing the "V1" sound, participants' routine operation was to immediately make a judgment and provide a verbal response. This routine operation had developed into an "automatic association" through repeated operation (Verbruggen et al., 2014; Henson et al., 2014). This process was governed by exogenous attention, where the salient appearance of the "V1" sound instantly triggered the verbal response. In the HighSimu condition, however, participants were required to make priority judgments when both the "V1" sound and the alarm occurred simultaneously, necessitating focused cognitive engagement. This decision-making process, in contrast to the automatic association in the other three workload conditions, demanded endogenous control. Participants were unaware in advance whether a given trial would involve the HighSimu condition. As a result, despite being instructed to prioritize the alarm-response task in the HighSimu condition, the strong attentional capture elicited by the "V1" sound may still have led to exogenous control, causing participants to prioritize the routine operation instead.

Considering that different alarm modalities may influence operators' ACMs, the effects of six alarm modalities on operator ACMs were examined. If participants prioritize the alarm-response task, it is considered endogenous control; otherwise, it is regarded as exogenous control.

Data Analysis

Jamovi was used for statistical analysis, with a significance level set at $\alpha = 0.05$. Outliers in CRT data (more than 3 standard deviations) were excluded. A chi-square test analyzed hits, errors, and misses. Linear mixed models were used for CRT, NASA-TLX, and UEQ data, with workload and alarm modalities as fixed effects and subjects as random effects. A generalized mixed model was used to analyze ACMs. Post hoc comparisons were performed using the Bonferroni method.

RESULTS

Performance of the Alarm-Response Task

1) Number of hits, errors, misses

A chi-square test showed no significant differences in hits, errors, or misses among alarm modalities in the Baseline ($p = .167$), HighSucc ($p = .145$), and HighSimu ($p = .689$) conditions. However, significant differences were observed in the Low workload condition (p <.001), where the V alarm had significantly more misses than other modalities. The detailed results are presented in Table 4.

Table 4: Four levels of workload of alarm-response task.

Workload		Alarm Modalities						χ^2	p
		V	A	VA	VT	AT	VAT		
Baseline	Hit	19	22	19	21	22	19	14.130	.167
	Error	1	0	1	0	0	3		
	Miss	2	0	2	1	0	0		
Low	Hit	5	21	22	22	22	22	72.692	<.001
	Error	1	1	0	0	0	0		
	Miss	16	0	0	0	0	0		
HighSucc	Hit	22	18	22	20	22	22	14.667	.145
	Error	0	2	0	1	0	0		
	Miss	0	2	0	1	0	0		
HighSimu	Hit	21	22	22	21	21	22	3.070	.689
	Error	1	0	0	1	1	0		
	Miss	0	0	0	0	0	0		

2) Choice Response Times (CRTs)

A mixed model revealed no significant differences in CRTs for the Baseline (F (5, 105) = 1.79, p = .121) and HighSucc workloads (F (5, 105) = 1.63, p = .159). Significant differences were observed in the Low (F (5, 105) = 7.38, p <.001) and HighSimu workloads (F (5, 105) = 7.38, p <.001).

Post hoc analyses (Bonferroni correction) are presented in Figure 3. The results showed that the CRTs for the V alarm were the longest, with significant differences compared to the A (p <.001), VA (p <.001), AT (p <.001), and VAT alarms (p <.001) in the Low and HighSimu workload. The VAT alarm produced the shortest CRTs in the Low workload.

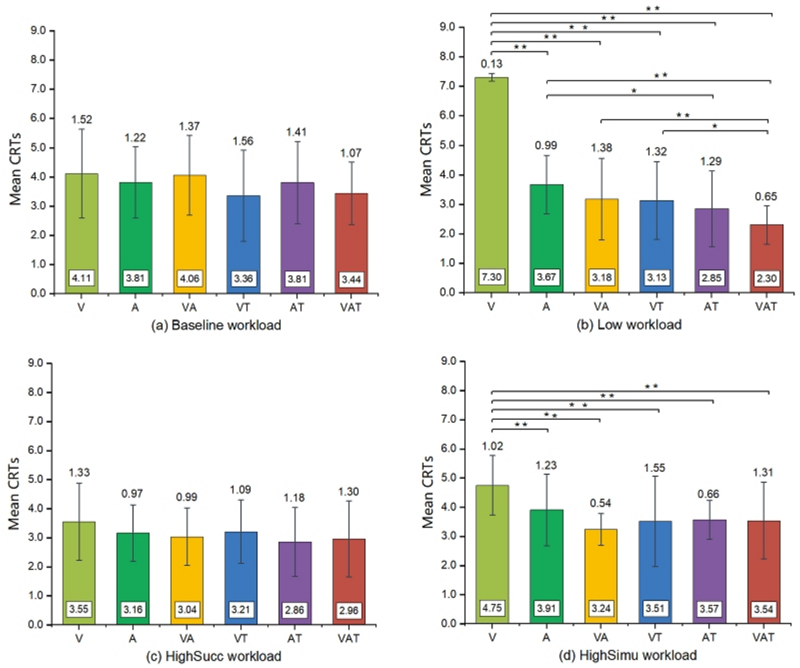

Figure 3: Comparison of CRTs for six alarm modalities across four workload levels.

Perceived Workload

NASA-TLX results showed no significant differences in workload across alarm modalities in Baseline, HighSucc, and HighSimu conditions. Under Low workload, significant differences were found in the "Performance" category (F = 3.56, p = .005), with post hoc analysis indicating that the A and VAT alarms differed (p = .011) as well as the V and VAT alarms (p = .011).

User Experience of Six Alarm Modalities

Overall, the results showed that the V alarm performed the worst, while the VAT alarm received the highest scores across all five items.

Attention Control Modes (ACMs) in HighSimu condition

A generalized mixed model revealed no significant effects of six alarm modality on ACMs (χ^2 (5) = 9.43, p = .093). However, when comparing only the V and A alarms, a significant difference was found (χ^2 (1) = 6.04, p = .014), indicating that A alarm was more effective than V alarm in eliciting the appropriate ACM response. Descriptive data is presented in Table 5.

Table 5: Descriptive data of ACMs of six alarm modalities.

ACMs	V	A	VA	VT	AT	VAT
Endogenous control	14	20	17	18	13	17
Exogenous control	8	2	5	4	9	5

DISCUSSION

Effects of Alarm Modalities Under Varying Levels of Workload

In the Baseline condition, no significant difference in performance was observed across alarm modalities, supporting previous research (Stanton et al., 2013), which suggests that uni-modal alarms are sufficient in less demanding tasks. However, under Low workload, where participants had just completed routine operations, significant differences emerged. The V alarm had the lowest hit rate, while VAT alarm led to the best CRTs. This may be because participants, after completing routine task operations, were in a relaxed state detached from the task context, resulting in lower vigilance (Al-Shargie et al., 2019; Robison et al., 2021). VAT alarm, combining auditory preemption (Ferraro & Mouloua, 2021) and strong attention-grabbing characteristics of tactile stimuli (Krausman et al., 2007), improve response speed compared to other modalities.

Interestingly, the V alarm had a higher miss rate in Low workload than in HighSucc, likely because participants were more engaged in the complex routine tasks of HighSucc, maintaining higher alertness (Pop et al., 2012). This suggests that interventions, such as periodic signals aimed at preventing attention decline during prolonged engagement in low-stimulation tasks and maintaining task engagement (Pop et al., 2012; Mishler & Chen, 2024), must reach a specific threshold to effectively activate operators. These interventions are crucial for re-engaging operators in a human-in-the-loop state (Cimini et al., 2020; Wu et al., 2023), especially in monotonous monitoring tasks.

In the HighSucc condition no significant difference was found between alarm modalities, likely due to participants' high task engagement, which prevented performance degradation from being reflected in the differences between the alarm modalities, despite resource consumption. In the HighSimu condition, except for the V alarm, which had the longest CRTs, the performance of the other five alarm modalities did not differ significantly. This may be due to participants were instructed to prioritize alarm-response tasks, which resulted in performance degradation in routine operations of the dual-task interference (Wickens et al., 2021).

Attention Control Modes and Task Priority

This study found that A alarm was more effective than V alarm in modulating ACMs. This may be because A alarm can effectively "compete" within the same sensory stream as the routine operations. This competition helps shift attention from automatic, exogenous control to endogenous control, enabling operators to prioritize the alarm over routine operations. Despite multimodal alarms of VAT configuration have been shown to enhance perceptual abilities through exogenous attention (Oskarsson et al., 2012), their impact on task prioritization was not significant in this study. This difference may be attributed to the nature of the tasks. While exogenous attention is effective for rapid threat detection, the current study required more complex cognitive processing for task prioritization, which redundant multimodal cues do not fully address.

Limitations

A limitation of this study is the simplified representation of real-world workload conditions in the controlled laboratory setting. In practice, environmental factors such as noise vibrations, or temperature may influence alarm effectiveness. Additionally, as each trial included one alarm occurrence, participants may have anticipated alarms, potentially influencing their responses. Despite this, the benefits of multimodal alarms observed in the study are likely to persist in more realistic environments.

CONCLUSION

This study evaluated the effectiveness of uni-, bi-, and tri-modal alarms across four workload conditions in discrete monitoring tasks. Considering the overall performance across all workload conditions, the VAT alarm demonstrated the best overall effectiveness, showcasing its robustness in achieving superior results across varying workload scenarios. Additionally, auditory alarm was more effective than visual alarm in modulating ACMs, improving task prioritization in simultaneous dual-task scenarios. These findings underscore the potential of multimodal alarms to enhance operator performance and demonstrate that auditory alarm effectively modulate ACMs, thereby ensuring system safety in discrete monitoring tasks.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Mr. Yi Xie from the Commercial Aircraft Corporation of China for his valuable insights on alarm design and task settings.

REFERENCES

Al-Shargie, F., Tariq, U., Mir, H., Alawar, H., Babiloni, F., & Al-Nashash, H. (2019). Vigilance decrement and enhancement techniques: A review. Brain sciences, 9(8), 178.

Avril, E., Cegarra, J., Wioland, L., & Navarro, J. (2021). Automation Type and Reliability Impact on Visual Automation Monitoring and Human Performance. *International Journal of Human-Computer Interaction*, 38(1), 64–77.

Causse, M., Mercier, M., Lefrançois, O., & Matton, N. (2025). Impact of automation level on airline pilots' flying performance and visual scanning strategies: A full flight simulator study. *Applied Ergonomics*, 125, 104456.

Cimini, C., Pirola, F., Pinto, R., & Cavalieri, S. (2020). A human-in-the-loop manufacturing control architecture for the next generation of production systems. *Journal of manufacturing systems*, 54, 258–271.

Clark, R. M., & Bustamente, E. A. (2008). Enhancing decision making by implementing likelihood alarm technology in integrated displays. *Modern Psychological Studies*, 14(1), 5.

Cun, W., Yu, S., Chu, J., Chen, Y., Sun, J.,... Fan, H. (2024). Exploration of multimodal alarms for civil aircraft flying task: A laboratory study. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 34(4), 279–291.

Dehais, F., Causse, M., Vachon, Regis,... Tremblay. (2014). Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness. *Human Factors*, 4(56), 631–644.

Ferraro, J. C., & Mouloua, M. (2021). Effects of automation reliability on error detection and attention to auditory stimuli in a multi-tasking environment. *Applied Ergonomics*, 91, 103303.

Fisk, A. D., & Lloyd, S. J. (1988). The role of stimulus-to-rule consistency in learning rapid application of spatial rules. *Human Factors*, 30(1), 35–49.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 9, pp. 904–908). Sage CA: Los Angeles, CA: Sage publications.

Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus–response bindings in priming. *Trends in cognitive sciences*, 18(7), 376–384.

Ho, C., Gray, R., & Spence, C. (2014). Reorienting driver attention with dynamic tactile cues. *IEEE Transactions on Haptics*, 7(1), 86–94.

Krausman, A. S., Pettitt, R. A., & Elliott, L. R. (2007). Effects of redundant alerts on platoon leader performance and decision making. Army Research Laboratory.

Lu, S. A., Wickens, C. D., Sarter, N. B., Thomas, L. C., Nikolic, M. I., & Sebok, A. (2012). Redundancy gains in communication tasks: A comparison of auditory, visual, and redundant auditory-visual information presentation on NextGen flight decks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 1476–1480). Sage CA: Los Angeles, CA: SAGE Publications.

Lu, S. A., Wickens, C. D., Prinet, J. C., Hutchins, S. D., Sarter, N.,... Sebok, A. (2013). Supporting interruption management and multimodal interface design: Three meta-analyses of task performance as a function of interrupting task modality. *Human Factors*, 55(4), 697–724.

Meng, F., Ho, C., Gray, R., & Spence, C. (2015). Dynamic vibrotactile warning signals for frontal collision avoidance: Towards the torso versus towards the head. *Ergonomics*, 58(3), 411–425.

Meng, F., & Spence, C. (2015). Tactile warning signals for in-vehicle systems. *Accident analysis & prevention*, 75, 333–346.

Mishler, S., & Chen, J. (2024). Boring but demanding: Using secondary tasks to counter the driver vigilance decrement for partially automated driving. *Human factors*, 66(6), 1798–1811.

Oskarsson, P. A. E. L., & And Carlander, O. (2012). Enhanced perception and performance by multimodal threat cueing in simulated combat vehicle. *Human Factors*, 54, 122–137.

Robison, M. K., Unsworth, N., & Brewer, G. A. (2021). Examining the effects of goal-setting, feedback, and incentives on sustained attention. *Journal of Experimental Psychology: Human Perception and Performance*, 47(6), 869.

Park, N. W., Lombardi, S., Gold, D. A., Tarita-Nistor, L., Gravely, M., Roy, E. A.,... Black, S. E. (2012). Effects of familiarity and cognitive function on naturalistic action performance. *Neuropsychology*, 26(2), 224.

Pop, V. L., Stearman, E. J., Kazi, S., & Durso, F. T. (2012). Using engagement to negate vigilance decrements in the NextGen environment. *International Journal of Human-Computer Interaction*, 28(2), 99–106.

Rusnock, C. F., & Borghetti, B. J. (2018). Workload profiles: A continuous measure of mental workload. *International Journal of Industrial Ergonomics*, 63, 49–64.

Salzer, Y., Oron-Gilad, T., Ronen, A., & Parmet, Y. (2011). Vibrotactile "on-thigh" alerting system in the cockpit. *Human Factors*, 53(2), 118–131.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). Applying the user experience questionnaire (UEQ) in different evaluation scenarios. In *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience: Third International Conference*, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I 3 (pp. 383–392). Springer International Publishing.

Scott, J. J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Human factors*, 50(2), 264–275.

Stanton, N. A., Harvey, C., Plant, K. L., & Bolton, L. (2013). To twist, roll, stroke or poke? A study of input devices for menu navigation in the cockpit. *Ergonomics*, 56(4), 590–611.

van Erp, J. B., Toet, A., & Janssen, J. B. (2015). Uni-, bi-and tri-modal warning signals: Effects of temporal parameters and sensory modality on perceived urgency. *Safety science*, 72, 1–8.

Verbruggen, F., McLaren, I. P., & Chambers, C. D. (2014). Banishing the control homunculi in studies of action control and behavior change. *Perspectives on Psychological Science*, 9(5), 497–524.

Wickens, C. D., Helton, W. S., Hollands, J. G., & Banbury, S. (2021). *Engineering psychology and human performance* (5th Editioned.). New York: Routledge.

Wickens, C., Prinet, J., Hutchins, S., Sarter, N., & Sebok, A. (2011). Auditory-visual redundancy in vehicle control interruptions: Two meta-analyses. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 1155–1159). Sage CA: Los Angeles, CA: SAGE Publications.

Wu, J., Huang, Z., Hu, Z., & Lv, C. (2023). Toward human-in-the-loop AI: Enhancing deep reinforcement learning via real-time human guidance for autonomous driving. *Engineering*, 21, 75–91.

Yang, S., & Ferris, T. K. (2019). Supporting multitracking performance with novel visual, auditory, and tactile displays. *IEEE Transactions on Human-Machine Systems*, 50(1), 79–88.

Zirk, A., Wiczorek, R., & Manzey, D. (2020). Do we really need more stages? Comparing the effects of likelihood alarm systems and binary alarm systems. *Human factors*, 62(4), 540–552.