Human-Computer Interaction & Emerging Technologies, Vol. 233, 2026, 188–198
https://doi.org/10.54941/ahfe1007518

Wearable Cue Design: A Comparative Research of Different Modalities of Smartwatches in Microtasks

Xinyu Huang[1], Zhiwei Yao[2], Diyu Zhou[3], Xing Fang[1,4], and Yuxuan Li[1]

[1]Department of Information Design, Wuhan University of Technology, Wuhan, China
[2]School of Design, Jiangnan University, China
[3]The College of Architecture, Art, and Planning, Cornell University, United States
[4]Yangtze River Culture Research Institute, China

ABSTRACT

Despite the widespread adoption of smartwatches, rigorous experimental verification of how their multimodal interaction (visual + haptic) impacts performance and cognitive load in dyadic collaboration remains limited. Adopting a 3 (Modality) × 2 (Task Type) within-subjects design ($N = 14$), we constructed a dual-task framework based on an intragroup dependency mechanism. This framework comprised two individual micro-tasks and one collaborative micro-task, requiring participants to synchronously respond to target stimuli for performance assessment, while secondary Probe Tasks were employed to evaluate cognitive load. Additionally, Performance Decrement Rate (PDR) was used to measure collaborative adaptability. Results indicate that multimodal cues significantly reduce distraction costs by offloading visual demands. While efficiency benefits are context-dependent and moderated by individual baseline capability—specifically benefiting lower-capability users via a compensation effect—multimodal interaction significantly enhances collaborative stability by suppressing performance fluctuations during high-load switching. These findings validate the Cognitive Resource Release Hypothesis and highlight the necessity for adaptive interaction designs in collaborative wearables.

Keywords: Smartwatch, Multimodal interaction, Collaborative micro-task, Cognitive load, Stability

INTRODUCTION

Smartwatches show significant potential in individual monitoring (Årsand et al., 2015; Triantafyllidis et al., 2024), yet their role in multi-user collaboration remains underexplored. Existing research (Olteanu et al., 2025; Wong et al., 2017) relies heavily on qualitative prototypes rather than controlled experiments. We focus on “mobile collaborative micro-tasks”—high-frequency, low-complexity operations repeated within seconds (Bernstein et al., 2011; Luce, 1991; Quinn & Bederson, 2011; Wickens, 2008), which provide an ideal scenario for investigating multimodal cues in collaboration.

Previous research indicates visual cues accelerate reaction times, while haptic or cross-modal cues offer advantages under visual load (Warm et al.,

Received January 31, 2026; Revised April 7, 2026; Accepted April 26, 2026; Available online July 20, 2026

2008; Wickens, 2008). Mixed Reality studies also suggest haptics effectively attract attention (Lin et al., 2022), making them suitable for collaboration. However, experimental verification regarding their impact on collaborative synchrony and efficiency is lacking.

Furthermore, synchrony is decisive in collaboration; subtle desynchronization can amplify efficiency decline (Heath & Luff, 1992; Salas et al., 2008). While multimodal cues theoretically optimize attention allocation, their specific utility in high-frequency collaboration remains unclear. Our objective is to rigorously quantify standard performance indicators to determine if the value of multimodality lies in efficiency gains or other dimensions. This paper systematically compares smartwatch multimodal cues (visual + haptic) versus smartphone cues (visual-only). Integrating metrics such as IES and PDR, we address:

- **RQ1 (Task Efficiency):** How does multimodal efficiency compare to smartphone interactions across individual and collaborative tasks?
- **RQ2 (Distraction Index):** Can multimodal cues reduce distraction by releasing visual cognitive resources?
- **RQ3 (Collaborative Stability):** Do multimodal cues demonstrate stronger stability and adaptability during high-load task switching?

We make three contributions: (1) Validating a dual-task paradigm based on intragroup dependency to simulate collaborative micro-tasks; (2) Demonstrating that multimodal cues serve as effective substitutes for smartphones, enhancing stability by releasing resources; and (3) Verifying the "Compensation Effect" to propose adaptive cueing strategies based on user capability.

RELATED WORK

Micro-Tasks and Collaborative Task Performance

Computer-Supported Cooperative Work (CSCW) emphasizes technology's role in facilitating team communication and synchronization (Bannon & Schmidt, 1989; Grudin, 2002). While tasks are typically classified by complexity and frequency, "mobile collaborative micro-tasks"—characterized by low complexity but high repetition—present unique challenges. High-frequency operations help form rhythmic expectations but may lead to cumulative attentional consumption and vigilance fatigue (Sarter & Woods, 1995; Warm et al., 2008). In collaborative contexts, subtle desynchronization or delays between individuals can amplify into an overall performance decline (Salas et al., 2008). Existing research has enhanced team performance through shared displays or workspace awareness (Greenberg et al., 1996; Heath & Luff, 1992). Wearable devices have also been introduced, such as smartwatch prototypes for aviation alerts (Wong et al., 2017) or factory assembly (Olteanu et al., 2025). However, most studies rely on qualitative prototypes, lacking systematic verification under controlled conditions for high-frequency micro-tasks.

Task Performance and Cognitive Load

Cognitive load describes the extent to which limited cognitive resources are occupied (Sweller, 1988). Classical theory categorizes it into three types: intrinsic (task complexity), extraneous (interface interference), and germane load (learning effort) (Paas & Van Merriënboer, 1994). In high-frequency micro-tasks, while intrinsic load is low, the cumulative effect of triggers is prone to vigilance decrement (Parasuraman, 1986). However, most experiments focus on high-complexity tasks, with insufficient attention paid to the cumulative load in low-complexity collaborative scenarios.

Multimodal Cues and Sensory Allocation. In micro-task environments, operators' visual and manual resources are often occupied, creating "modal ambiguity" that limits visual-only cues (Sarter & Woods, 1995; Wickens, 2008). Recent immersive research reveals a dual effect: while visual overlays improve efficiency, continuous reliance on vision can cause attentional bottlenecks and increase extraneous load (Alessa et al., 2023). Multimodal cues (e.g., visual-haptic) optimize attention allocation by utilizing alternative channels. Individual task studies show haptics accelerate reaction times under high visual load (Lin et al., 2022; Luce, 1991). In collaboration, these cues help operators perceive actions timely, reducing error accumulation. Despite these theoretical advantages, systematic experiments targeting high-frequency collaborative micro-tasks remain limited.

Measurement and Manipulation of Cognitive Load. Cognitive load measurement includes subjective and objective methods. Subjective methods like NASA-TLX capture psychological perception but require retrospective reporting, potentially interrupting high-frequency tasks (Hart & Staveland, 1988). In contrast, objective behavioral metrics such as Reaction Time (RT), Accuracy, and Inverse Efficiency Score (IES) directly reflect resource consumption and are better suited for micro-tasks (Wickens, 2008). Additionally, the secondary task probe paradigm infers real-time attention margins (Kahneman, 1973; Wickens, 2002). Combining task difficulty adjustments (e.g., CRT) with probe insertion allows for objectively quantifying distraction costs in collaborative contexts.

Impact of Individual Differences. Cognitive load effects are not uniform; individual differences in attentional capacity significantly influence performance (Davies et al., 2013; Just & Carpenter, 1992). To control this, experiments typically establish baseline measurements (e.g., SRT/CRT). Since team performance often depends on the weakest link, failing to control these differences can obscure modality effects. Notably, Jeong et al. (2024) highlighted that Cognitive Capacity moderates multisensory efficacy: low-capability individuals often derive greater benefits. This suggests that individual baseline capability must be considered a core variable.

METHODS

This study proposes the following hypotheses:

- **H1 (Task Efficiency):** Smartwatch multisensory notifications significantly improve task efficiency in individual micro-tasks and outperform smartphone interactions in collaborative micro-tasks.

- **H2 (Distraction Index):** Compared to smartphone screen interaction, smartwatch multisensory cues significantly reduce the distraction index for secondary tasks, releasing visual cognitive resources.
- **H3 (Collaborative Stability):** Compared to single visual interaction, smartwatch multimodal cues demonstrate stronger Collaborative Stability, manifested by the dispersion of Performance Decrement Rate (PDR).

Experimental Design and Overview

The experiment employed a 3 (Cue Modality: Multimodal Smartwatch vs. Visual-only Smartwatch vs. Smartphone) × 2 (Task Type: Individual CRT vs. Collaborative C-CRT) within-subjects design (see Figure 1). We adapted a reaction time task combined with a progress-checking task to simulate collaborative micro-task scenarios. Dependent variables included efficiency metrics (RT, Accuracy, IES), attention allocation metrics (Probe RT, Distraction Cost), and collaborative adaptability metrics (PDR).

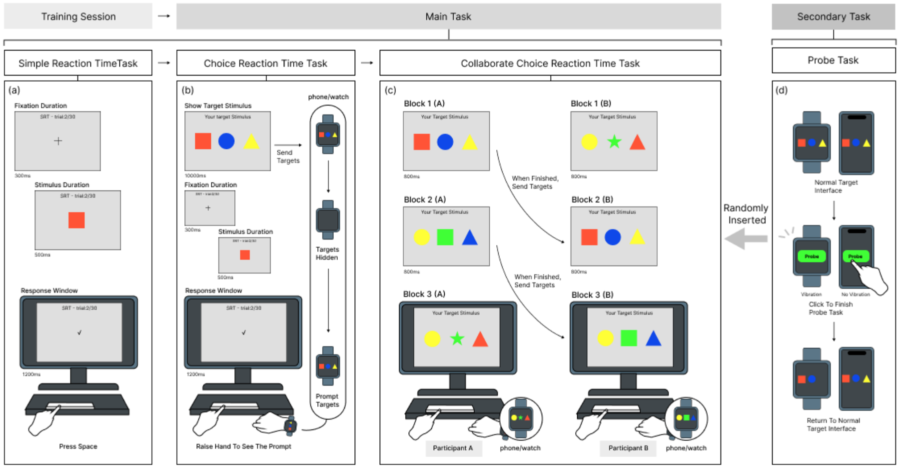

Figure 1: Overview of the experimental procedure and task design.

Primary Task Paradigm and Procedures

The primary tasks consisted of a progressive hierarchy of cognitive load: SRT (Simple Reaction Time, baseline), CRT (Choice Reaction Time, 3 targets with 40% probability), and C-CRT (Collaborative Choice Reaction Time). Notably, C-CRT introduced an inter-personal dependency mechanism where the follower's (Participant B) target stimuli were determined by the leader's (Participant A) output from the previous block, simulating a "lockstep" collaborative constraint. Each task type comprised three blocks, with each block consisting of 30 trials. To minimize fatigue and carryover effects, a 30-second rest was enforced between blocks, and a 60-second rest between different task types. We calculated Inverse Efficiency Score (IES = RT / Accuracy) to provide a balanced efficiency metric that avoids speed-accuracy trade-offs.

Secondary Task and Cognitive Load Measures

To assess real-time attentional resources, a Secondary Probe Task was randomly interspersed (accounting for 12–13% of trials) between primary task trials. Participants were required to quickly confirm these progress checks while maintaining primary task accuracy. We utilized Probe Reaction Time and Distraction Cost (the RT difference between trials with and without probes) to quantify the interference caused by specific cue modalities.

Collaborative Stability Measurement

To verify robustness in complex contexts, we adopted Performance Decrement Rate (PDR) as the core metric for collaborative stability. PDR quantifies the relative cognitive cost of switching from independent (CRT) to collaborative (C-CRT) tasks, defined as $\mathrm{PDR} = (\mathrm{IES}_{\mathrm{Collab}} - \mathrm{IES}_{\mathrm{Indiv}}) / \mathrm{IES}_{\mathrm{Indiv}}$. The magnitude of PDR reflects the switching cost, while its dispersion (Standard Deviation) directly indicates the stability of the interaction system. Lower variance suggests that the modality possesses higher error tolerance across users.

Stimuli and Apparatus

As detailed in Figure 1, temporal parameters were controlled: stimulus duration was 500 ms, Inter-Stimulus Interval (ISI) was randomized between 500 and 1500 ms, and the response window was 1200 ms. Regarding apparatus, the Smartwatch Group received synchronized multimodal cues (visual + 500 ms vibration), while the Smartphone Group received visual-only cues. To facilitate discrimination, the smartwatch Probe task utilized a distinct 1000 ms vibration.

Participants and Analysis

From 16 recruits, 14 valid participants (4 males, 10 females; $M_{\mathrm{age}} = 22.00$, SD = 1.18) were retained after excluding device anomalies. All were right-handed with normal vision. Smartwatch experience was counterbalanced via block randomization. Crucially, Simple Reaction Time (SRT) was collected as a baseline covariate. Data analysis ($\alpha = .05$) retained 94.94% of valid trials after removing outliers (>2 SD). We constructed a Repeated Measures ANCOVA model on IES, setting Cue Modality and Task Type as within-subject factors and SRT Baseline as a continuous covariate to examine the three-way interaction. Normality and sphericity were verified, applying Huynh-Feldt and Bonferroni corrections where necessary.

RESULTS

Speed-Accuracy Trade-Off Check

Before analyzing the Inverse Efficiency Score (IES), we correlated Reaction Time (RT) and Accuracy to rule out potential "speed-accuracy trade-off" effects. Pearson analysis showed no significant negative correlations between RT and Accuracy across modalities in either individual (CRT) or collaborative (C-CRT) tasks ($p > .05$). This indicates participants did not sacrifice accuracy for speed; thus, all subsequent efficiency analyses are based on IES values.

Task Efficiency

To examine cue modality efficacy, we conducted a 3 (Cue Modality) × 2 (Task Type) Repeated Measures ANCOVA on IES, including individual SRT Baseline as a covariate.

Main Effects and Two-way Interactions. The analysis results showed that after controlling for the moderating effect of baseline capability, the main effect of Task Type did not reach a significant level ($F(1, 11) = 0.38$, $p = .551$, $\eta^2_p = .033$). This suggests that simply comparing the overall difficulty of individual versus collaborative tasks is insufficient to explain performance differences without considering specific contexts.

However, a significant interaction between Cue Modality and Task Type was observed ($F(2, 22) = 4.88$, $p = .018$, $\eta^2_p = .307$) (see Figure 2a). This validates the core hypothesis of H1, confirming that multimodal cue efficacy possesses high "context-dependency". Simple effects analysis indicated that in CRT, the Multimodal Smartwatch (C1) showed a trend towards optimal efficiency (M = 704.00, SD = 94.45), superior to both Visual-only Smartwatch (C2, M = 752.43, SD = 86.63) and Smartphone (C3, M = 726.46, SD = 78.43). In C-CRT, the difference between the Multimodal Smartwatch and Smartphone was not statistically significant, demonstrating a "Substitution Effect." This confirms that in high-load, hands-busy collaborative scenarios, the smartwatch serves as an effective alternative to the smartphone.

The Moderating Role of Individual Capability (Three-way Interaction). More critically, the model detected a significant three-way interaction among Modality, Task Type, and SRT Baseline ($F(2, 22) = 4.60$, $p = .021$, $\eta^2_p = .295$). This indicates that multimodal efficacy is significantly moderated by individual capability. Regression trends revealed that for participants with slower baselines (Low Capability), multimodal cues yielded the largest IES improvement in collaborative tasks, demonstrating a clear "Compensation Effect" (see Figure 2b). Conversely, for High Capability participants (extremely fast baselines), vibration cues introduced potential processing burdens, leading to diminished benefits. This underscores that individual differences must be incorporated as a core variable when evaluating wearable interaction efficacy.

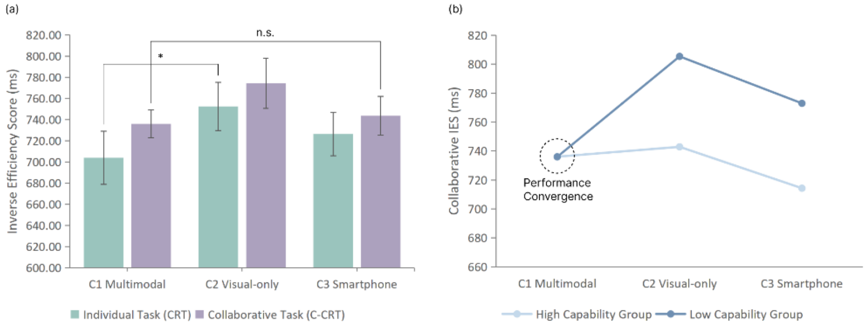

Figure 2: Interaction effects on inverse efficiency score (IES). (a) Two-way interaction between cue modality and task type, showing the general advantage of multimodal cues. (b) Three-way interaction revealing the moderating role of individual capability in collaborative tasks, highlighting the "performance convergence" effect in the multimodal condition. Cognitive load and distraction cost.

Cognitive Load and Distraction Cost

To verify H2 (cognitive resource release), we conducted a One-way Repeated Measures ANOVA on Probe RT. Prior to analysis, one outlier (P004) was excluded due to anticipatory responses, retaining 13 valid samples. Mauchly's test met sphericity assumptions ($p = .556$). Results revealed a significant main effect of cue modality ($F(2, 24) = 6.293$, $p = .006$, $\eta^2_p = .344$) (see Figure 3). The Multimodal Smartwatch (C1) achieved the shortest Probe RT (M = 529.52, SD = 148.93), significantly outperforming both Visual-only Smartwatch (C2, M = 648.14, SD = 152.07, $p = .041$) and Smartphone (C3, M = 679.65, SD = 132.64, $p = .038$). No significant difference was found between C2 and C3 ($p = 1.000$). These findings confirm that haptic cues effectively offloaded visual burden in resource-constrained tasks, enabling responses approximately 120–150ms faster, thereby demonstrating a lower distraction cost.

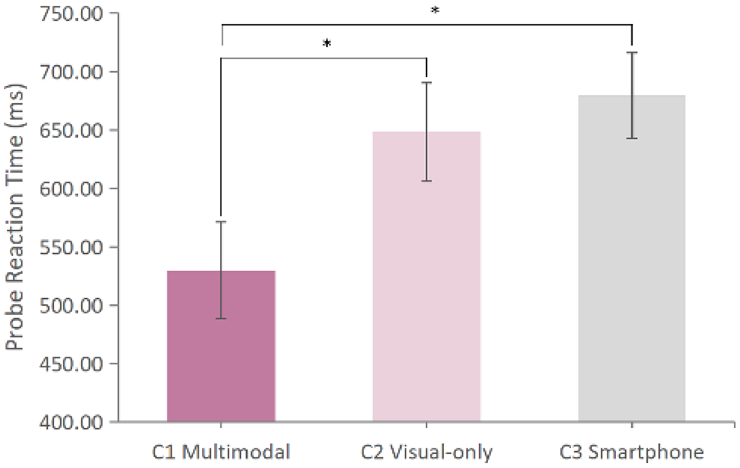

Figure 3: Probe reaction time (Probe RT) across different cue modalities.

Collaborative Stability

Regarding H3, a One-way Repeated Measures ANOVA on mean Performance Decrement Rate (PDR) found no significant differences ($F(2, 26) = .475$, $p = .627$), indicating that the three devices performed comparably in terms of average switching cost.

However, analysis of Stability (PDR dispersion) revealed distinct patterns (see Figure 4). The Visual-only Smartwatch (C2) exhibited the highest fluctuation (SD = 13.93%) suggesting susceptibility to "Performance Collapse" due to substantial individual differences. In contrast, Multimodal cues (C1) reduced this fluctuation to 11.30%, effectively converging performance dispersion. While Smartphones (C3) showed the lowest fluctuation (SD = 9.05%)—likely due to high user proficiency—multimodal interaction demonstrated superior robustness compared to visual-only wearables. This confirms that multimodal cues successfully rectify the instability inherent in single-channel visual interfaces during high-load switching, making them a more reliable collaboration tool.

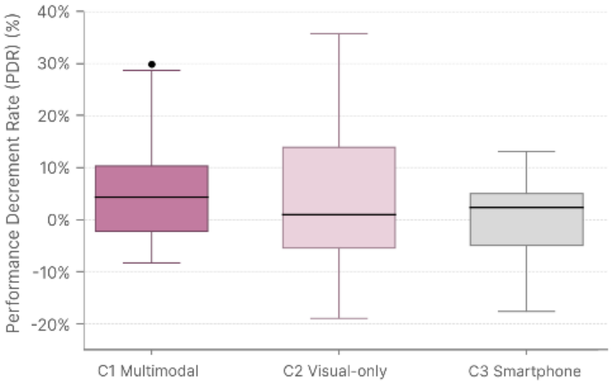

Figure 4: Boxplots of performance decrement rate (PDR) distribution across three modalities.

DISCUSSION

This study investigated the efficacy of smartwatch multimodal cues in mobile collaborative micro-tasks. Through controlled experiments, we assessed task efficiency, cognitive load, and, for the first time, the moderating role of individual baseline capability in a collaborative context. The findings reveal a mechanism more complex than simple efficiency improvement: multimodal efficacy exhibits significant context-dependency (acting as an efficient substitute in collaboration), is critically modulated by individual capability (benefiting low-capability users more), and demonstrates unique advantages in maintaining collaborative stability.

Context-Dependency of Multimodality

Our results indicate that multimodal cues do not yield an absolute speed advantage in all contexts. While haptic feedback accelerated responses in low-load individual tasks, it demonstrated non-inferiority—rather than significant superiority—to smartphones in high-load collaborative tasks. This aligns with Wickens' Multiple Resource Theory. Collaborative micro-tasks require simultaneous visual processing and motor execution. Although haptic cues offload the visual channel, vibration as an exogenous stimulus may introduce slight interference when the motor channel is saturated. Therefore, the primary value of multimodal interaction in collaboration lies in the "Substitution Effect": it enables users to maintain high-level performance comparable to smartphones without relying on continuous visual monitoring. This offers significant ecological value for real-world scenarios where hands are occupied, such as maintenance or logistics, where visual freedom is as critical as speed.

The Moderating Role of Individual Capability

The most striking finding of this study is the interaction between cue modality and individual baseline capability. Our data confirms that the efficacy of multimodal cues is highly dependent on the user's Cognitive Capacity.

For users with slower baseline reactions (Low Capability), multimodal cues brought significant performance improvements. Acting as a mandatory "interruption," haptic signals effectively compensated for their deficits in visual search or attention maintenance, demonstrating a clear "Compensation Effect." In contrast, for users with extremely fast baseline reactions (High Capability), the introduction of vibration did not yield additional benefits and even showed a trend of interference in some trials. This can be explained by the "Ceiling Effect": the visual processing speed of high-capability users approaches physiological limits, making additional sensory channels not only redundant but potentially increasing the burden of cognitive integration. This finding suggests that future wearable collaborative systems should not adopt a "one-size-fits-all" design but should possess the adaptive capability to dynamically adjust cueing strategies based on user capability.

Cognitive Resource Release and Interaction Robustness

The results of the Probe Task strongly support the cognitive resource release hypothesis. Since haptic cues can be perceived without oculomotor orientation, they effectively release the bandwidth of the visual channel, enabling users to respond to secondary stimuli at a lower cognitive cost (shorter reaction time) while maintaining primary task performance.

This release of resources translates into an implicit safety net—Collaborative Stability. The high-performance fluctuation (high standard deviation) in the Visual-only Smartwatch group indicates that relying solely on the small visual interface of the watch easily leads to cognitive overload and Performance Collapse during task switching. The introduction of haptic feedback effectively converged this fluctuation. In collaborative scenarios with extremely low error tolerance, this predictable stability is often of greater practical value than mere average speed.

Limitations and Future Work

As a pilot study, limitations remain. First, while the sample size (N = 14) detected main effects, it restricted detailed demographic analysis. Future research should expand to diverse populations (e.g., factory workers) to verify generalizability. Second, a Novelty Effect was observed; novices showed larger dispersion than skilled users. Longitudinal studies are needed to determine if advantages stabilize over time. Third, the lab setting cannot fully replicate real-world noise. Future work will migrate this paradigm to field environments to verify the anti-interference utility of haptic cues.

REFERENCES

Årsand, E., Muzny, M., Bradway, M., Muzik, J. and Hartvigsen, G. (2015). Performance of the first combined smartwatch and smartphone diabetes diary application study. Journal of Diabetes Science and Technology, 9(3), pp. 556–563.

Bannon, L.J. and Schmidt, K. (1989). CSCW: Four characters in search of a context. Proceedings of the First European Conference on Computer-Supported Cooperative Work (ECSCW '89), pp. 358–372.

Bernstein, M.S., Brandt, J., Miller, R.C. and Karger, D.R. (2011). Crowds in two seconds: Enabling realtime crowd-powered interfaces. Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, pp. 33–42.

Davies, D.R., Matthews, G., Stammers, R.B. and Westerman, S.J. (2013). Human Performance: Cognition, Stress and Individual Differences. Psychology Press.

Greenberg, S., Gutwin, C. and Cockburn, A. (1996). Awareness through fisheye views in relaxed-WYSIWIS groupware. Graphics Interface, 96, pp. 28–38.

Grudin, J. (2002). Computer-supported cooperative work: History and focus. Computer, 27(5), pp. 19–26.

Gutwin, C. and Greenberg, S. (1998). Design for individuals, design for groups: Tradeoffs between power and workspace awareness. Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work, pp. 207–216.

Hart, S.G. and Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in Psychology (Vol. 52, pp. 139–183). North-Holland.

Heath, C. and Luff, P. (1992). Collaboration and control: Crisis management and multimedia technology in London Underground Line Control Rooms. Computer Supported Cooperative Work (CSCW), 1(1), pp. 69–94.

Jeong, S., Kim, J. and Lee, J. (2024). The differential effects of multisensory attentional cues on task performance in VR depending on the level of cognitive load and cognitive capacity. IEEE Transactions on Visualization and Computer Graphics, 30(5), pp. 2703–2712.

Just, M.A. and Carpenter, P.A. (1992). A capacity theory of comprehension: individual differences in working memory. Psychological Review, 99(1), p. 122.

Kahneman, D. (1973). Attention and Effort. Prentice-Hall.

Lin, T.-C., Krishnan, A.U. and Li, Z. (2022). Comparison of Haptic and Augmented Reality Visual Cues for Assisting Tele-manipulation. 2022 International Conference on Robotics and Automation (ICRA), pp. 9309–9316.

Luce, R.D. (1991). Response Times: Their Role in Inferring Elementary Mental Organization. Oxford University Press.

Olteanu, A., Marian, C.G. and Pietraru, R.N. (2025). Smartwatch-Based Monitoring and Alert System for Factory Operators Using Public Cloud Services. Applied Sciences, 15(5), p. 2806.

Paas, F.G. and Van Merriënboer, J.J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. Journal of Educational Psychology, 86(1), p. 122.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In Handbook of Perception and Human Performance (Vol. 2). Wiley.

Quinn, A.J. and Bederson, B.B. (2011). Human computation: A survey and taxonomy of a growing field. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1403–1412.

Salas, E., Cooke, N.J. and Rosen, M.A. (2008). On teams, teamwork, and team performance: Discoveries and developments. Human Factors, 50(3), pp. 540–547.

Sarter, N.B. and Woods, D.D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human Factors, 37(1), pp. 5–19.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), pp. 257–285.

Triantafyllidis, A., Kondylakis, H., Katehakis, D., Kouroubali, A., Alexiadis, A., Segkouli, S., Votis, K. and Tzovaras, D. (2024). Smartwatch interventions in healthcare: A systematic review of the literature. International Journal of Medical Informatics, 190, p. 105560.

Warm, J.S., Parasuraman, R. and Matthews, G. (2008). Vigilance requires hard mental work and is stressful. Human Factors, 50(3), pp. 433–441.

Wickens, C.D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science, 3(2), pp. 159–177.

Wickens, C.D. (2008). Multiple resources and mental workload. Human Factors, 50(3), pp. 449–455.

Wong, S., Singhal, S. and Neustaedter, C. (2017). Smart crew: A smart watch design for collaboration amongst flight attendants. Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 41–44.