Design Science

Designing adaptive feedback systems for managing cognitive load in augmented reality

Jiacheng Sun and Ting Liao

Department of Systems Engineering, Stevens Institute of Technology, Hoboken, NJ, USA

Abstract

Managing cognitive load is central to designing interactive systems, particularly within augmented reality (AR) environments that impose complex and immersive demands. This study investigates two complementary approaches in parts to managing cognitive load in AR: refining interaction modalities and integrating adaptive physiological feedback. In Part 1, eye-tracking and hand-based modalities are evaluated across tasks of varying difficulty, using skin conductance responses (SCRs) as a proxy for cognitive load. Results show that while hand gestures improved task performance in simple tasks, cognitive load levels are comparable across modalities. In Part 2, an adaptive feedback system based on a signal-derived metric, cumulative SCR (CSCR), is developed to trigger short rest interventions during sustained cognitive load. Statistical analyses illustrate that rest interventions significantly reduced cumulative cognitive load, though their effect on task performance was inconclusive. These findings emphasize the trade-offs between cognitive relief and performance continuity and highlight the potential of physiologically adaptive systems in supporting cognitive-aware interaction design.

Keywords: Cognitive Load, Human Computer Interaction, Skin Conductance Response, Adaptive Feedback System, Physiological Feedback

Received 12 January 2025
Revised 29 October 2025
Accepted 30 October 2025

Corresponding author
Ting Liao
tliao@stevens.edu

Des. Sci., vol. 11, e49
journals.cambridge.org/dsj
DOI: 10.1017/dsj.2025.10040

1. Introduction

Cognitive load refers to the mental effort required to process information (Sweller 2011). Managing cognitive load is one of the fundamental considerations for designing interactive systems (Kosch *et al.* 2023; Gkintoni *et al.* 2025), particularly in immersive environments like augmented reality (AR), where the system complexity and information interference can increase users' cognitive load and consequently undermine task performance, as observed in higher dwell time, increased gaze switching, and poorer task recall compared to traditional 2D interfaces (Alessa *et al.* 2023; Suzuki, Wild & Scanlon 2024). AR has been widely adopted across domains, such as education, health care, and industry, to deliver context-aware information by integrating virtual and real environments. In the context of engineering design, AR offers unique potential by enabling users to visualize, manipulate, and evaluate digital prototypes in physical spaces, thereby enhancing spatial reasoning, iterative exploration, and collaborative decision-making. For instance, in industrial contexts, AR has been used to guide construction workers through safety procedures and to visualize underground geological structures during infrastructure planning (Garzón 2021; Purwinarko, Hardyanto & Adhi 2021). Understanding cognitive load in AR interaction design is thus essential for improving user experience and performance.

A wide range of strategies has been proposed to manage users' cognitive load and thereby enhance task performance and overall experience, particularly in AR environments (Hou *et al.* 2025). For instance, animated AR guidance has been applied in industrial maintenance tasks to help workers repair complex mechanical systems, significantly reducing task ambiguity and completion time (Alessa *et al.* 2023). In educational AR scenarios, multimodal interfaces combining voice commands and hand gestures have enabled users to manipulate virtual molecules or annotate diagrams more naturally, thereby improving engagement and learning efficiency (Chen *et al.* 2024). Smart assistants using gaze and speech have also been used to answer user queries or highlight relevant objects in the environment, lowering search effort and attentional load (Wang *et al.* 2025). Building on these approaches, multimodal interaction, such as eye-gazing-based and hand-based controls, provides users with flexible and intuitive ways to interact within augmented environments. This multimodal approach is particularly valuable for cognitively demanding tasks, as it reduces cognitive load by allowing users to split effort across channels or choose the most intuitive input method (Lystbæk *et al.* 2022). To evaluate the effectiveness of these management strategies, it is essential to measure cognitive load accurately and reliably. Traditionally, cognitive load assessment has relied on subjective measurement tools, such as the National Aeronautics and Space Administration (NASA) Task Load Index (NASA-TLX), which captures users' perceived psychological, physical, and temporal demands experienced during a task (Hart & Staveland 1988). While widely used, subjective measurements have limitations, particularly in lengthy tasks, where post–task ratings are subject to bias and may fail to reflect dynamic changes in cognitive load (Hart 2006).

To overcome these limitations, recent studies have increasingly used physiological signals as objective, real-time indicators of cognitive load. Among these, skin conductance response (SCR) has emerged as a promising method (Soshi *et al.* 2021). Derived from galvanic skin response (GSR) data, SCR captures changes in skin's electrical conductance in response to stimuli and has been widely used to assess arousal, attention, and cognitive load in paradigms involving visual search, attention cueing, and reaction to auditory stimuli (Benedek & Kaernbach 2010; Yoshida *et al.* 2014). Previous researchers have experimentally demonstrated that SCR varies with the degree of attention and used SCR as a basis for assessing the degree of attention (Yoshida *et al.* 2014). Assessing the activity of a stimulus or intervention-related event by measuring the SCR is a common process in empirical research (Benedek & Kaernbach 2010).

Despite progress in cognitive load assessment, most studies have focused on single interaction methods or domain-specific task scenarios, leaving a gap in understanding how different interaction modalities influence cognitive load. Furthermore, while physiological feedback systems have shown promise in areas like health monitoring (Rodrigues, Postolache & Cercas 2020), their application in adaptive cognitive load management remains relatively limited.

To address these gaps, this study integrates multimodal interaction with a physiologically adaptive feedback system for managing cognitive load in AR. This study consists of two parts, making two key contributions. First, it investigates how

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

different interaction modalities affect cognitive load across tasks of varying complexity. Second, it explores the potential of the adaptive system for managing cognitive load. Together, these contributions clarify the relationship between interaction methods and mental demands, informing the design of interactive systems that use real-time feedback to enhance user experience.

2. Background

2.1. Development of interaction methods in AR

As AR continues to support increasingly complex tasks, such as conceptual development and system integration, the design of intuitive and cognitively compatible interaction methods has become a critical concern. Over the past two decades, researchers have explored various interaction methods to improve user performance and experience in AR environments. Early developments focused on hand-based interaction, allowing users to manipulate virtual objects through natural gestures. These techniques were found to be intuitive and effective in enhancing task engagement and spatial understanding (Kumar, Paepcke & Winograd 2007). As sensing technologies advanced, eye-tracking emerged as a supplementary input modality, enabling users to select objects or trigger actions using gaze. Studies have shown that gaze-based input can accelerate object selection and reduce physical fatigue (Sibert & Jacob 2000; Duchowski 2017). The introduction of AR devices, such as Microsoft HoloLens, in 2015 further improved the tracking accuracy and responsiveness of hand and eye inputs, making them more accessible and robust for practical use. For example, Nguyen, Gouin-Vallerand & Amiri (2023) demonstrated that modern hand-gesture systems significantly improve user immersion and sense of control in complex AR tasks. These advancements have catalyzed growing interest in combining multiple interaction modalities to offer greater flexibility and task performance. Despite promising results in existing applications (Kolla & Plapper 2023; Zhang *et al.* 2023; Rasch *et al.* 2025), the impact of interaction modalities on aspects of user experience across areas such as cognitive load, attention distribution, and task strategy remains underexplored.

2.2. Cognitive load theory and assessment methods

According to cognitive load theory, cognitive load refers to the mental effort required to process information (Kalyuga 2011; Sweller 2011; Minkley, Xu & Krell 2021). Elevated cognitive load has been associated with higher error rates, longer task completion time, and impaired learning outcomes (Hepsomali *et al.* 2019). Traditionally, cognitive load has been measured using subjective assessment tools, such as the NASA-TLX, which are limited by self-report bias and lack of real-time resolution (Hart 2006). Recent advancements in biosensing technologies have enabled more objective approaches to measure cognitive load. One widely adopted method uses GSR, also known as electrodermal activity (EDA), which measures electrical conductance changes on the skin, reflecting variations in sweat gland activity associated with arousal or cognitive processing (Boucsein 2012). These changes reflect autonomic arousal, which has been closely associated with cognitive effort during task performance (Shi *et al.* 2007; Boucsein 2012). A growing

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

body of research supports the use of GSR for inferring cognitive load. For example, Ekin *et al.* (2025) demonstrated that GSR, combined with heart rate, heart rate variability (HRV), and skin temperature, can successfully distinguish between intrinsic and extraneous cognitive loads. Jukiewicz and Marcinkowska validated the effectiveness of EDA features in differentiating task demands (Jukiewicz & Marcinkowska 2025), while Cai and Demmans Epp applied EDA signals to predict learner workload during educational tasks (Cai & Demmans Epp 2024). Buchner *et al.* reviewed the increasing application of physiological indicators, including EDA, in evaluating cognitive load within AR environments (Buchner *et al.* 2023). These studies collectively support the integration of GSR-based feedback into interactive systems as a real-time, nonintrusive proxy for monitoring cognitive demand. Unlike post–task assessments, the GSR-based approach enables adaptive systems that can dynamically respond to users' cognitive states, offering new opportunities to enhance user support in immersive or cognitively demanding scenarios.

2.3. Adaptive systems based on real-time physiological feedback

As sensing technologies continue to evolve, adaptive systems leveraging real-time physiological feedback have been adopted in various domains. These systems are particularly effective in monitoring users' physical conditions and emotional states (Sun *et al.* 2023; Li & Liao 2025). Many studies utilize physiological signals such as heart rate, skin temperature, or electroencephalography (EEG) to detect the physical health status or mood fluctuations of users (Shu *et al.* 2018). For example, HRV is commonly used to assess emotional changes and stress levels (Shaffer & Ginsberg 2017), while EEG demonstrates significant advantages in detecting users' concentration and emotional stability (Zhu *et al.* 2024; García-Hernández *et al.* 2023). These systems are prevalent in telemedicine, sports health management, and emotion detection, providing real-time feedback to help users adjust their physical or emotional states (Liu, Sourina & Nguyen 2010; Nandi *et al.* 2021). However, their application in cognitive load management remains relatively limited. The challenge lies in managing dynamic cognitive load in complex task environments, where users must process layered information and respond to multiple interaction inputs in real time (Alessa *et al.* 2023). To address this gap, this study introduces an adaptive feedback system that utilizes real-time physiological input to trigger brief rest-based interventions, allowing users to recover from elevated cognitive load without interrupting the overall task flow. The system includes a task-specific signal interpretation method designed to approximate elevated cognitive load states in real time, supporting more responsive task adjustment.

2.4. Research objectives and hypothesis

This study addresses the following research challenges: (1) Although eye- and hand-based interactions have been examined, their effects on cognitive load in complex AR tasks are not fully understood (Moncur, Galvez Trigo & Mortara 2023). (2) Existing studies emphasize task performance, leaving a limited understanding of how interaction methods affect cognitive load management (Chen, Paas & Sweller 2023). (3) Although cognitive load theory and subjective measures are widely used, they may fail to reflect real-time cognitive states. Objective

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

assessment methods show promise for cognitive load management but require further validation in AR contexts (Alessa *et al.* 2023). (4) Adaptive systems using physiological feedback show promise in emotion and health monitoring, but their use for real-time cognitive load regulation in AR remains limited (Moncur, Galvez Trigo & Mortara 2023).

To address these gaps, the study comprises two parts:

Part 1 examines challenges (1)–(3) by analyzing the effects of interaction methods and task difficulty on cognitive load and task performance. Eye-tracking and hand-based interactions are across tasks of varying difficulty, simulating different levels of cognitive demands. Tasks are categorized into two difficulty levels: easy and hard. Easy tasks require straightforward interactions with minimal steps and decision-making points. Hard tasks require more steps and decision-making points, increasing interaction complexity and requiring greater focus to complete. Task performance is evaluated as task completion time, with faster performance indicating higher fluency and lower interaction complexity. Based on this setup, the following hypotheses are proposed and evaluated:

1. Easy Tasks
   - 1a. The task completion time for eye-tracking and hand-based operations is similar.
   - 1b. The cognitive load for eye-tracking and hand-based operations is similar.
2. Hard Tasks
   - 2a. The task completion time for eye-tracking operations is lower than that for hand-based operations.
   - 2b. The cognitive load for eye-tracking operations is lower than that for hand-based operations.
3. A longer task completion time leads to a greater SCR occurrence, reflecting greater cognitive load.

Part 2 addresses challenges (3) and (4) by implementing an adaptive feedback system driven by real-time physiological data. To support this function, a new metric, cumulative SCR (CSCR), is introduced to track sustained increases in skin conductance and approximate periods of elevated cognitive load. The system dynamically provides rest interventions as structured opportunities for recovery. These rest interventions are hypothesized to support sustained attention, reduce task completion time, and maintain cognitive stability.

The following hypotheses are proposed and evaluated:

4. Tasks with the adaptive rest module will show a reduction in CSCR per minute after rest events compared to before.
5. Tasks with rest interventions have shorter completion times compared to tasks without rest interventions.

The primary contribution of this study is the introduction of an adaptive feedback system that operates across different interaction modalities and utilizes real-time physiological monitoring to support cognitive load management in AR environments.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

3. Methods

To test the proposed hypotheses, a human-subject study was conducted in two parts, using an AR-based maze application as the experimental platform. The maze task was selected for its intuitive goal-directed nature and flexibility in operational complexity. The study utilized a 2×2 design to examine two interaction methods (gaze-based and hand-based) and two levels of task difficulty (easy and hard). The design led to four experimental conditions, which simulate various scenarios of AR interaction. Part 1 examined the effects of these conditions on performance and cognitive load. In Part 2, the effectiveness of the adaptive feedback system was evaluated in the same conditions.

Participants' task performance and cognitive load were the key dependent variables. Task performance was evaluated as task completion time, while cognitive load was assessed using both subjective rating and objective physiological measures. Subjective rating was measured using NASA-TLX, averaged across six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. For the physiological measure, in Part 1, cognitive load was assessed using SCR-based metrics. In Part 2, a real-time adaptive feedback system was introduced to manage cognitive load, supported by CSCR-based metrics. The system triggered rest interventions during tasks based on participants' cognitive load.

3.1. Experiment setup

The experiment was conducted in a closed laboratory environment, and the lab layout is shown in Figure 1. The setup included a computer, a Mindfield® eSense GSR sensor (shown in Figure 2), an Android device, and a HoloLens 2 headset. All devices were connected via an isolated local area network (LAN) to ensure stable and secure data transmission. To ensure procedural consistency, all instructions were shown on the computer or the headset, and the research assistants only responded to specific participant inquiries. The framework of the system is shown in Figure 3.

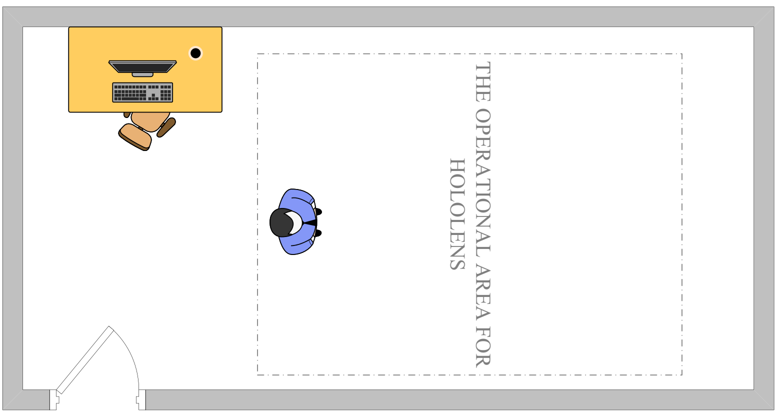


Figure 1. The lab layout. The layout shows the setup used in the experiment, with designated areas for the HoloLens and participants.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

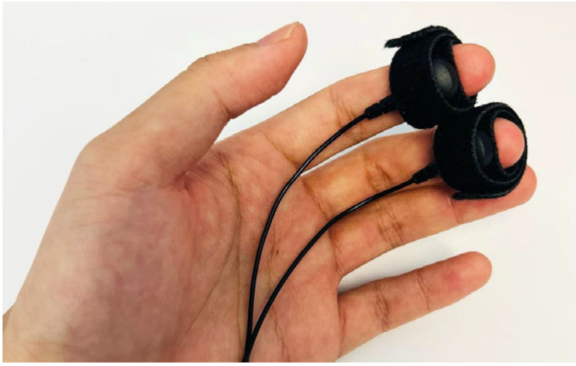


Figure 2. GSR wired electrodes. The sensor consists of two wired electrodes attached to the fingers and collects data at a sampling rate of 10 Hz.

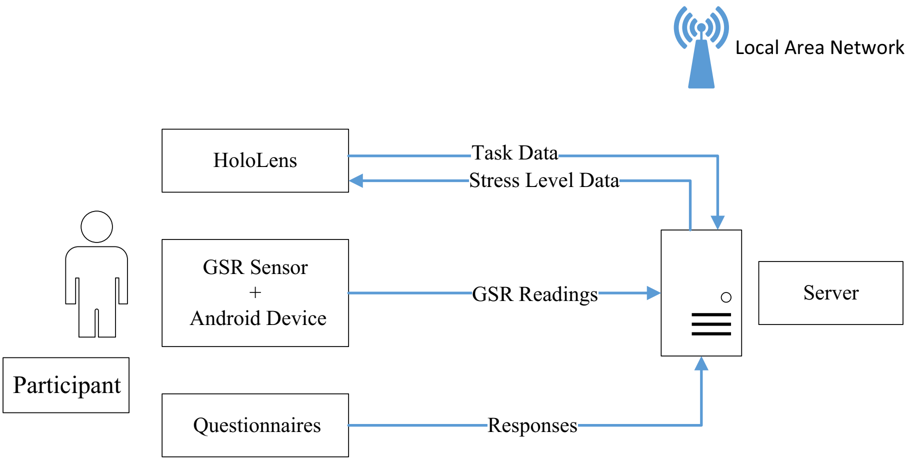


Figure 3. The system overview. This figure illustrates the system architecture, including the HoloLens, GSR sensor, Android device, and local area network.

The HoloLens 2 hosted a maze application designed for the experiment. GSR data were recorded using the eSense sensor and transmitted to a remote server via an Android device, where it was processed to indicate participants' real-time cognitive load. This cognitive load level was used as the trigger for the rest of the interventions. The computer presented the study instructions, questionnaires, and the relaxation video during the task.

3.2. Procedure

Participants were invited to the lab and provided informed consent, completed a demographic survey, and reported their familiarity with AR devices and their current emotional state. Participants were equipped with the GSR sensor, familiarized themselves with the headset operations, and completed eye-tracking calibration.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

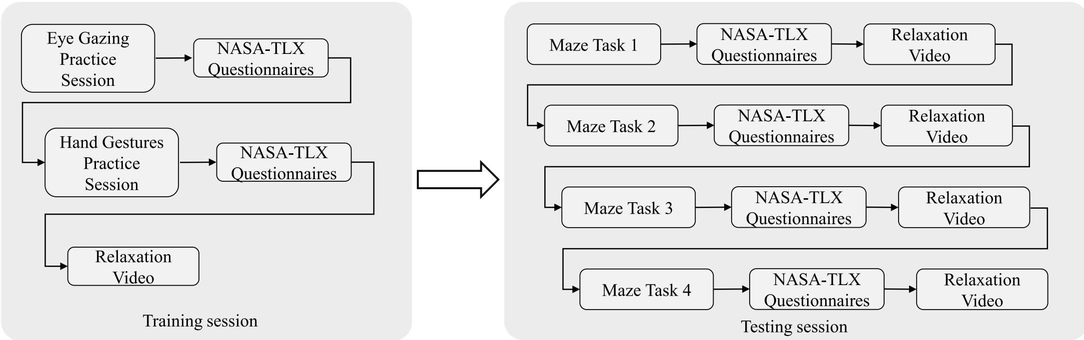


Figure 4. The experimental process, which includes a training session for participants to practice both eye-gazing and hand-gesture interactions and a testing session of four tasks to evaluate task performance and cognitive load. Each task was followed by a NASA-TLX questionnaire and a relaxation video.

The maze application guided participants through the training session and the testing session, as shown in Figure 4. In the training session, participants experienced two interaction methods: (a) eye-tracking control and (b) hand-based control. Participants practiced selecting a virtual cube using a pinch gesture and gaze control in the headset, as shown in Figure 5. Participants completed two standard NASA-TLX questionnaires and a two-minute relaxation video on the computer. The questionnaires and the video here were for familiarization purposes. These data were not used in the final experimental analysis. After the video, participants returned to the headset to start the testing session.

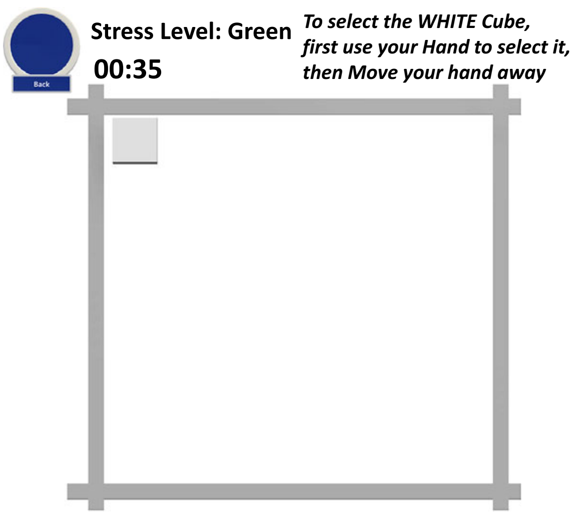
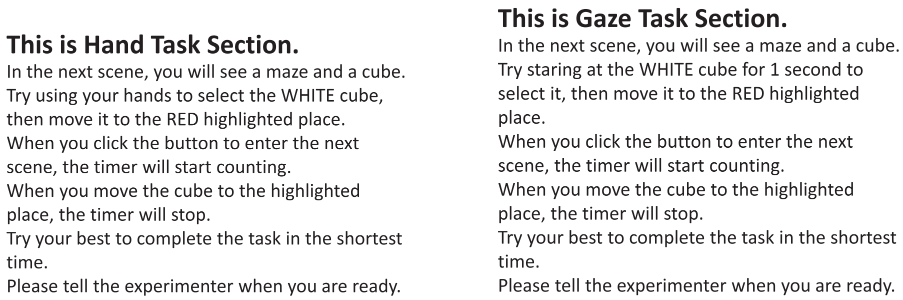


Figure 5. An illustration representing the user's view during the hand-gesture practice scenario. The user interface displays the cognitive load indicator and timer on the left, with a text prompt providing instructions on the top right. In this scenario, participants practiced selecting the virtual cube using a pinch gesture. *Note:* In this figure only, the white cube is shown in gray for visibility against the white background.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press



(a) Hand Task Section (b) Gaze Task Section

Figure 6. Interactive methods and task prompts.

In the testing session, participants experienced four maze tasks. After each task, participants completed the NASA-TLX questionnaire and the relaxation video. The maze order was randomized, and physiological and task performance data were collected continuously throughout the session. Upon completing all tasks, participants ranked the difficulty of the four tasks based on their subjective impressions.

3.3. Maze design

The maze task was developed using Unity3D and deployed on HoloLens 2. Participants navigated through the maze by controlling a virtual cube to reach target positions. Maze navigation provides an intuitive, goal-oriented task and allows flexibility and interaction complexity by manipulating maze paths.

Before each task, the application informed the participant of the designated interaction method and the task's completion criteria. The interaction method was displayed at the top of the screen, while the completion criteria were communicated through both on-screen text and voice instructions (as shown in Figures 6a,b). Once participants initiated the task, the timer began. When the cube reached the target, the timer stopped, the maze task closed automatically, and the application directed participants to complete the questionnaire. The instructions are shown in Figure 7.

Rather than increasing path complexity, the design emphasized the operational difficulty as the primary trigger for cognitive load. For the hard maze condition, dual path options were introduced to elicit greater active cognitive engagement (as shown in Figure 8).

3.4. Adaptive system design and data collection

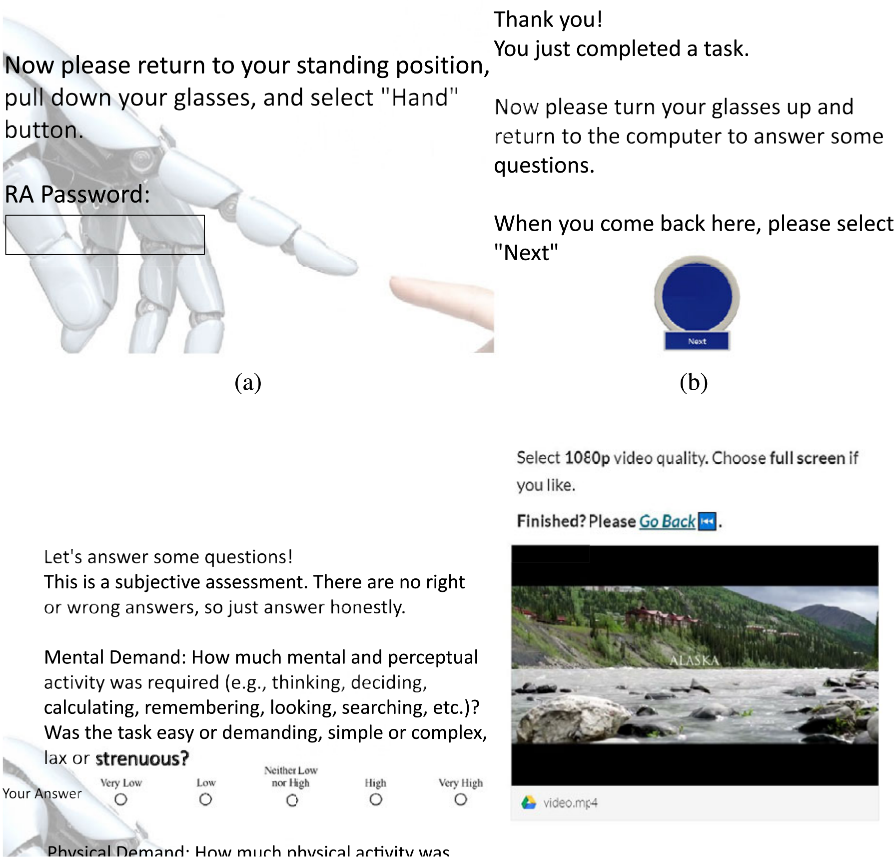
3.4.1. Skin conductance response (SCR)

Based on GSR, SCR specifically captures discrete phasic responses to short-term stimuli and is commonly used to monitor momentary fluctuations in cognitive load.

The detection and computation of SCR follow a four-step process, as shown in Figure 9.

1. Listening state: The default state of the application, where it continuously monitors incoming signals.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press



(a) (b)

(c) (d)

Figure 7. The workflow for switching between the AR environment via a HoloLens headset and the activities on the computer. (a) An instruction on the computer screen prompts the participant to return to the AR task area and resume the experiment. (b) An in-headset prompt instructs the participant to return to the computer after completing a task. (c) A partial view of the NASA-TLX questionnaire presented on the computer. (d) The relaxation video shown on the computer between tasks.

2. Detecting SCR raise: If the signal rises consistently for at least two seconds, or the difference between the current signal and the estimated base value exceeds 0.5 μS (Mindfield Biosystems Ltd, n.d.; accessed October 16, 2024), a potential SCR event is detected.
3. Gathering and calculating SCR during the fluctuation phase: Upon detecting an SCR rise, the application enters the fluctuation phase. The first signal in this phase is marked as the base value. Signal values are tracked during this phase.
4. Initiate recovery phase and end of fluctuation: The fluctuation phase ends when the signal drops by more than 50% from its peak amplitude. At this point, the application returns to the listening state.

*Notably, if the signal drops during the rising phase in step 2, the system aborts the detection and reverts to the listening state.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

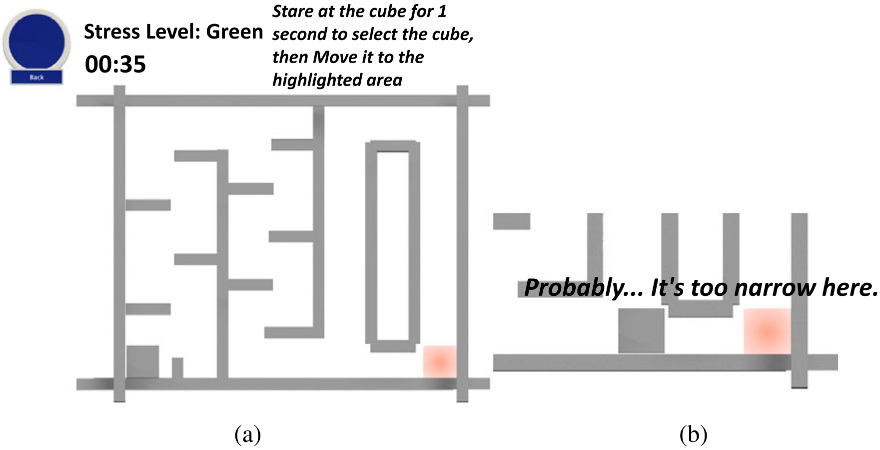
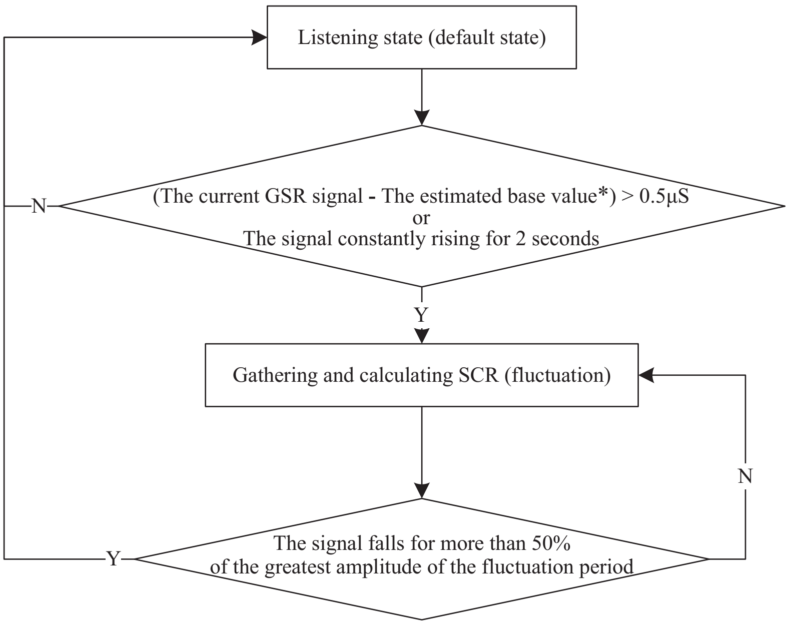


Figure 8. When the user tries to go through the lower path, a prompt will pop up to tell the user to try another path (the upper path) due to the width.



The estimated base value* : The value that is considered to be the first value in the rising signal. But at this stage, it is not confirmed yet as a base value

Figure 9. The judgment criteria for SCR activities.

Momentary cognitive load can be approximated by counting the number of SCR occurrences within a given time window (Ahmadi, Ozgur & Kiziltan 2024). This real-time event detection enables tracking physiological fluctuations during task execution. While SCR is effective in capturing phasic fluctuations in cognitive load, it exhibits certain limitations. Traditional SCR computation relies on identifying prominent signal peaks, which primarily reflect short-term responses to

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

specific stimuli. In scenarios involving sustained cognitive effort, SCR may not adequately capture gradual signal increases, leading to potential underrepresentation of cumulative cognitive load states.

3.4.2. Cumulative skin conductance response (CSCR)

To support the adaptive feedback system for scenarios with sustained cognitive effort, a new indicator termed CSCR was proposed and developed in this study. A CSCR event is identified using the same four-step process and criteria as a standard SCR event, as detailed in the preceding section (Figure 9). The novelty of the CSCR metric lies not in the detection of the event window but in the quantification method used during an active event. As illustrated in Figure 10, while a standard SCR is registered as a single count after the event concludes, CSCR provides a more granular, real-time measure of the response's duration. During an active event window, the CSCR count increments once per second for as long as the signal is in a continuous rising phase. This allows CSCR to capture and quantify smaller signal fluctuations and sustained periods of increase within a larger event, which would be missed by the single-count nature of a standard SCR.

Although tonic components were not formally extracted, CSCR approximates tonic-like signal patterns, which have been associated with cognitive load in extended tasks (Setz *et al.* 2010; Shi *et al.* 2007). CSCR was particularly suited for real-time system feedback, where peak-based or area-based measures may be inefficient to compute. By incorporating CSCR into the adaptive feedback system, it became possible to track users' cognitive load in real time and trigger rest-related interventions in real time, when sustained overload was detected.

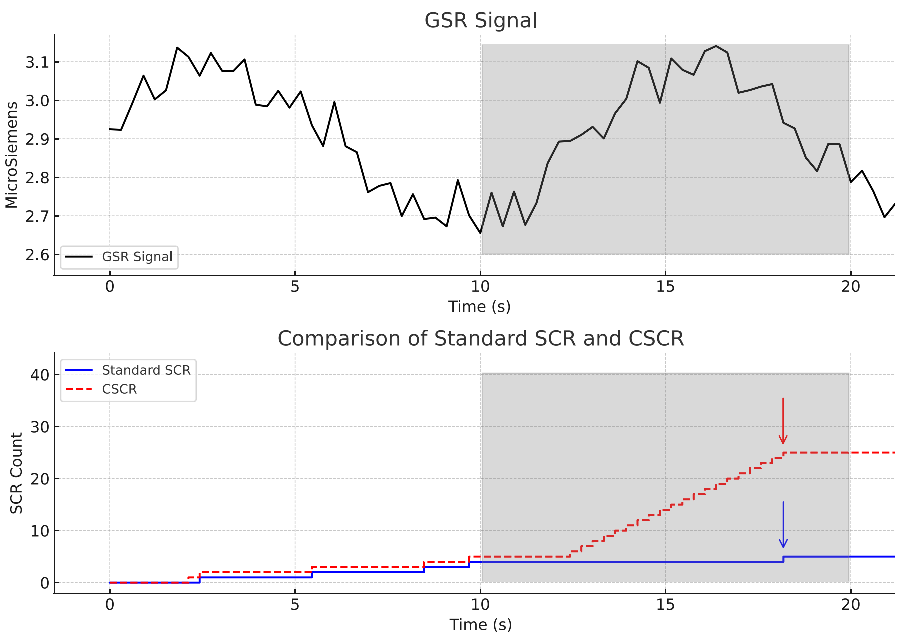


Figure 10. Comparison of CSCR and SCR based on one participant's recorded data.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

3.4.3. Design of real-time feedback

The real-time detection utilized CSCR/min as the core metric. This metric is well-suited because it allows continuous monitoring of sustained conductance increases without relying on discrete peak detection. To enable intuitive feedback, CSCR/min values were categorized into four discrete levels based on adapted thresholds from the SCR/min guidelines in the eSense Skin Response manual (Mindfield Biosystems Ltd, n.d.; accessed October 16, 2024): 0–5 as "green" (low cognitive load), 6–9 as "yellow" (moderate), 10–15 as "orange" (high), and 16 or above as "red" (very high).

CSCR/min was continuously calculated on the server throughout each task. Specifically, this metric was computed using a 60-second sliding window. At any given moment, it represented the cumulative sum of all CSCR counts that occurred in the preceding 60 seconds. The current cognitive load level was transmitted to the headset and displayed in real time as a color-coded indicator. When the indicator reached the "red" level, the system automatically initiated a 15-second rest intervention, as shown in Figure 11. To avoid interference at task onset, the rest module was disabled during the first 10 seconds of each task. The timer paused during the 3-second prompt (Figure 11a) and resumed during the rest intervention (Figure 11b), which was included in the total task time. The phrase "Smile, please:)" was included as a gentle, positive suggestion to help users relax during the break. Participants were allowed to end the rest early, and repeated interventions were permitted if the threshold was exceeded again. Notably, the timer continued running during the rest intervention to preserve the integrity of performance measurements.

3.4.4. Metrics

A set of metrics was used to evaluate participants' task performance and cognitive load during the experiment. These indicators are summarized in Table 1.

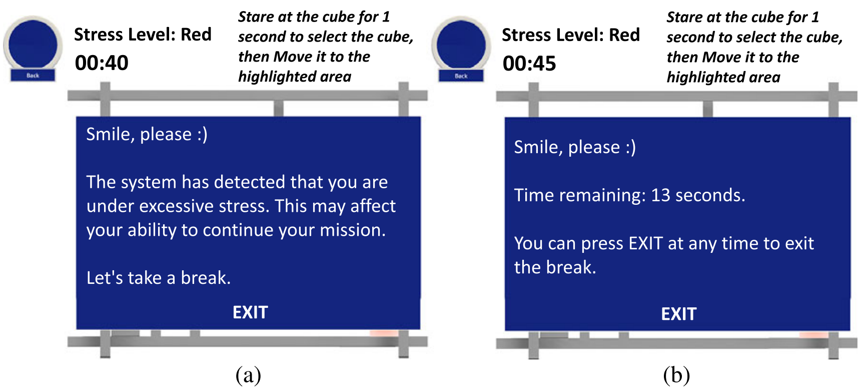


Figure 11. The two-stage process of an adaptive rest intervention triggered when the cognitive load level reaches "red." (a) The system first displays an initial prompt, notifying the user that a period of high cognitive load has been detected. (b) The system then transitions to a 15-second rest period with a countdown timer.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

Table 1. Overview of metrics used to evaluate task performance and cognitive load

Metric	Description	Purpose
Task completion time	The time required to complete each task	To assess the task performance
SCR occurrence	The number of SCR events identified from the GSR signal	To capture instances of fluctuations in cognitive load
SCR/min	The number of SCR events per minute, calculated by the server	To normalize SCR data for comparison across tasks
CSCR occurrence	The number of CSCR events identified from the GSR signal	To capture periods of sustained arousal in cognitive load
CSCR/min	The number of CSCR events per minute, calculated by the server	To normalize CSCR data for comparison across tasks
NASA-TLX scores	The level of subjective effort, measured using the NASA-TLX questionnaire	To measure subjective perceptions of cognitive workload
Rest time	The total rest duration during each task, recorded by the HoloLens system	To quantify intervention duration for evaluating the impact of the adaptive feedback system

3.5. Participants

A total of 40 participants were recruited in two separate groups for the two parts of the study. Four participants were excluded due to incomplete data from equipment malfunction or voluntary withdrawal. This resulted in a final sample of 36 participants, with 17 in Part 1 and 19 in Part 2, whose data were used for the final analysis. All 36 participants reported normal emotional states before the experiment to reduce potential confounds in physiological signal interpretation. The final sample included 22 males and 14 females, with an average age of 24.9 years. Participants' academic levels included 5 undergraduates and 31 graduate students, primarily from STEM disciplines such as computer science, systems engineering, and mechanical engineering. A majority of participants (66.7%) reported limited familiarity with AR devices (categorized as "Not at all" or "Somewhat familiar").

4. Results and analysis

4.1. Part 1. Impact of interaction methods

Part 1 analyzed the impact of the interaction method under two levels of task difficulty, focusing on task completion time and SCR occurrences. To provide a rigorous and transparent analysis, a multimodel strategy was adopted.

A two-way repeated-measures analysis of variance (ANOVA) was applied as baseline analysis. The result confirmed a significant main effect of task difficulty on both task completion time ($F(1,16) = 90.56$, $p < 0.01$, partial $\eta_p^2 = 0.85$) and SCR occurrences ($F(1,16) = 34.29$, $p < 0.01$, partial $\eta_p^2 = 0.68$), but no significant main effect was observed for interaction method nor for the interaction between the two factors. While this test suggests that the effect of the interaction method is

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

consistent across difficulty levels, we would like to further explore the localized effects, which the ANOVA analysis might obscure, particularly under different conditions.

To analyze the data from a population-average perspective, a generalized estimating equation (GEE) model was applied. GEE is particularly effective for repeated-measures data in which the correlation structure among within-subject observations must be considered (Zeger & Liang 1986; Ballinger 2004). In contrast to ANOVA, which estimates subject-specific effects, GEE models population-level average effects and provides robustness against misspecification of the within-subject correlation structure. These features make GEE well-suited for experimental designs with modest sample sizes and an emphasis on identifying general performance trends across participants. The main effect term for the interaction method in GEE directly evaluates the performance difference in the baseline (easy task) condition. This analysis showed a statistically significant advantage for hand gestures in the easy task condition ($\beta = -8.25, p < 0.01$).

As a cross-validation of this finding, a permutation test and a bootstrap analysis were conducted on the easy task data. The permutation test confirmed a highly significant performance advantage for hand gestures in the easy condition ($p < 0.01$). The bootstrap 95% confidence interval for the mean difference (hand–gaze) was $[-13.29, -3.43]$, further reinforcing the practical significance of this effect.

Taken together, while the ANOVA did not detect a significant interaction, the GEE model and direct tests support our conclusion that hand gestures were significantly faster than eye gazing in easy tasks, allowing us to reject Hypothesis 1a.

Regarding Hypothesis 1b, the same models were applied. Both the ANOVA and GEE models consistently indicated no significant difference between the two interaction methods for easy tasks. This result supports Hypothesis 1b.

For Hypothesis 2, results from both the ANOVA and GEE models showed no significant differences between eye gazing and hand gestures for hard tasks, in terms of either task completion time or SCR occurrence. Therefore, Hypotheses 2a and 2b are not supported.

A correlational analysis between the SCR occurrence and task completion time across all tasks delivered a Pearson correlation coefficient of 0.70 ($p < 0.01$), indicating a large effect size. This suggests that participants experiencing greater physiological arousal, as reflected by more SCR occurrences, tended to take longer to complete tasks. The correlation test supports Hypothesis 3, indicating that higher cognitive load is associated with longer task completion time.

4.1.1. Summary of findings for Part 1

The results for Part 1 provide the following insights:

- Hypothesis 1 was partially supported: While hand gestures resulted in faster task completion than eye gazing for easy tasks, cognitive load was similar for both methods.
- Hypothesis 2 was not supported: Task completion time and cognitive load were similar for both interaction methods under high-task difficulty, suggesting that increased complexity may reduce the performance differences between interaction modalities.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

- Hypothesis 3 was supported: A strong positive correlation was observed between the number of SCR occurrences and task completion time across all tasks.

These findings suggest that task difficulty moderates the influence of interaction methods on cognitive load and task performance. Based on these insights, Part 2 shifts the focus to the role of an adaptive rest module, examining how real-time physiological feedback can support cognitive load management and enhance task performance.

4.2. Part 2. Impact of the adaptive rest module

In Part 2, an adaptive rest module was introduced to examine its effects on cognitive load and task performance.

To validate the CSCR occurrence as an indicator of cognitive load, a pairwise comparison has been conducted regarding NASA-TLX scores, as the ground truth of cognitive load. A correlation analysis between SCR and NASA-TLX scores showed a small-to-moderate positive correlation ($r = 0.28$, $p = 0.03$, indicating a small-to-moderate effect size), indicating that SCR may reflect cognitive responses. In comparison, CSCR occurrence demonstrated a stronger correlation with NASA-TLX scores ($r = 0.41$, $p < 0.01$, indicating a moderate effect size), suggesting that CSCR occurrence may reflect cumulative aspects of participants' cognitive experience. These results support the use of CSCR occurrence as a task-sensitive indicator for reflecting cognitive load trends in dynamic or extended task scenarios.

To assess Hypothesis 4, the impact of the adaptive rest module on cognitive load was examined using CSCR/min. A total of 47 rest events were recorded among 18 participants, with an average duration of 11.45 seconds and a median of 15 seconds. Notably, 61.7% of these breaks reached the maximum allowable duration, suggesting that participants often accepted and used the full rest intervention when prompted. Each task triggered an average of 1.34 rest events, with some tasks prompting up to 4, confirming that the system was responsive to real-time physiological changes and actively engaged during task execution.

Specifically, a t-test was used to compare the CSCR/min values within identical 60-second sliding windows recorded immediately before and after the rest intervention. This comparison showed a statistically significant reduction in CSCR/min ($t = -2.55$, $p = 0.01$, Cohen's $d = 0.60$), supporting that the adaptive rest module leads to a reduction in cumulative cognitive load.

To further examine the influence of rest intervention while considering the interaction methods and task difficulty holistically, the ordinary least squares (OLS) regression model was performed. The change in CSCR occurrence before and after rest was used as the dependent variable (ΔCSCR Occurrence), and rest time (in seconds), interaction method (gaze versus hand), and task difficulty (easy versus hard) were included as predictors. While CSCR/min captures the real-time cognitive load state surrounding each rest event, the regression model examined the change in total CSCR occurrences to represent the cumulative reduction associated with different rest durations and task conditions. The model is specified as

$$\begin{aligned}\Delta \text{CSCR Occurrence} = {} & \beta_1 \cdot \text{RestTime} + \beta_2 \cdot \text{TaskType} \\ & + \beta_3 \cdot \text{DifficultyLevel} + \epsilon\end{aligned} \tag{1}$$

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

Regression results showed that rest time had a significant negative effect on CSCR occurrence difference ($\beta_1 = -0.43$, $p < 0.01$, $R^2 = 0.66$), indicating that longer rest durations were associated with greater reductions in cumulative cognitive load. The other two predictors were not statistically significant, suggesting that rest time was the primary explanatory factor. This result provides additional evidence that the adaptive rest module contributed to lowering cognitive load, further supporting Hypothesis 4.

Although the experiment used a 2×2 design, the adaptive rest module provided feedback based solely on real-time physiological state and did not differentiate between interaction method and task difficulty. To ensure this did not bias the results, both factors were included as dummy-coded predictors in the regression model. Neither was significant, supporting the decision to analyze the overall effect without stratifying by task condition.

To evaluate Hypothesis 5, which hypothesized that rest interventions would improve task performance, task completion time was compared between Part 1 and Part 2.

For our primary analysis, we focused on the adjusted completion time for Part 2 ($M = 59.51s, SD = 32.75$), which excludes rest durations, to better isolate the effect of workflow interruptions on performance. A t-test on these adjusted times showed a statistically significant difference ($t = 2.34, p = 0.02$, Cohen's $d = 0.55$), indicating that tasks in Part 1 ($M = 47.39s, SD = 26.00$) were completed more quickly. For completeness, we also analyzed the total completion time including rest breaks ($M = 68.13s, SD = 42.27$ for Part 2), which yielded the same significant conclusion ($t = 3.34, p < 0.01$, Cohen's $d = 0.79$). This discrepancy may be attributed to interruptions in task continuity caused by the inserted rest interventions, which could have affected participants' concentration and workflow.

While the adaptive rest module effectively reduced cognitive load, its impact on task performance was inconclusive. Hypothesis 5 was not supported. These findings suggest that although physiological feedback can inform rest timing, optimizing the frequency and duration of rest interventions may be essential for improving both cognitive state and task performance.

4.2.1. Summary of findings for Part 2

The results for Part 2 can be summarized as follows:

- Hypothesis 4 was supported: CSCR/min values were significantly reduced after rest interventions compared to before, confirming its potential in moderating cognitive load during task execution.
- Hypothesis 5 was not supported: Although the adaptive system reduced cognitive load, the inclusion of rest interventions led to longer task completion time, suggesting a trade-off between cognitive recovery and task performance.

To validate the task difficulty manipulation, participants were asked to rank the four task types from most to least difficult (1 = most difficult, 4 = easiest) at the end of all tasks. Significant negative correlations were found between these rankings and NASA-TLX scores ($r = -0.34$, $p < 0.01$), task completion time ($r = -0.32$, $p = 0.01$), and SCR ($r = -0.29$, $p = 0.02$), indicating that more difficult tasks were consistently associated with higher subjective and physiological load. These results support the effectiveness of the task differentiation and confirm alignment between task design and participant perception.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

According to the previous analysis, the stronger correlation between the CSCR and subjective workload (NASA-TLX) than SCR suggests its potential utility in future cognitive monitoring systems. These findings highlight the potential of adaptive physiological feedback systems for cognitive load management. Further work is needed to refine rest timing and duration to better balance load reduction with performance efficiency.

5. Limitations

Several limitations of this study should be acknowledged. First, the sample size was relatively small. Most participants were STEM students, which may introduce biases related to greater familiarity with technical interfaces and multitasking. This characteristic might limit the generalizability of the findings to a broader population. Future studies could consider increasing the sample size to improve statistical power and population diversity.

Second, this study did not include a direct comparison between CSCR and other established cumulative electrodermal indicators, such as peak amplitude sum or area under the curve. Future work will benefit from benchmarking CSCR against these techniques to further validate its responsiveness and applicability. Its broader applicability could be assessed by comparing it with behavioral or performance-based metrics across varied contexts.

Third, the tasks in the experiment were relatively brief, with average completion times ranging from approximately 20–70 seconds. This short duration may limit the ability to observe the cumulative effects of adaptive rest interventions on cognitive performance over longer periods of sustained engagement. Additionally, time-resolved analyses were not performed due to the high variability and reduced stability of physiological signals within short time windows. As a result, the correlation analysis with post–task NASA-TLX scores focused on overall load estimation rather than real-time fluctuation. Future studies involving longer tasks could enable time-series analysis and capture dynamic cognitive load trajectories. Longitudinal studies could further examine how sustained task engagement and recurring rest interventions influence cognitive load and task performance over extended timeframes.

Finally, while the adaptive feedback system reduced cognitive load, it was associated with longer task completion time. This suggests a trade-off between cognitive relief and performance efficiency. Moreover, although CSCR occurrence correlated with task time, this relationship may reflect individual strategies rather than direct task demand. Some participants may have slowed their actions to manage internal states, making CSCR more indicative of regulation strategies than task duration. Future work could design adaptive systems that adjust not only for physiological thresholds but also for diverse self-regulation patterns, using dynamic or time-normalized measures.

6. Discussion

The findings from Part 1 suggest that the performance advantages of hand gestures over eye gazing are primarily observed under low-task difficulty conditions. This pattern may be attributed to participants' prior experience with gesture-based interfaces, which facilitates more efficient execution in cognitively undemanding

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

contexts. However, under high-task difficulty, the performance gap between the two interaction modalities narrowed considerably, implying that increased cognitive demands may attenuate the benefits associated with specific modalities. In this study, task difficulty was operationalized based on the number of decision points and required path selections, reflecting the concept of element interactivity (Chen, Paas & Sweller 2023). This conditional effect of interaction modality aligns with the framework of cognitive load theory. The observed convergence in performance under high-task difficulty thus likely reflects the cognitive saturation imposed by complex task structures, which can override modality-specific advantages. These results offer implications for the design of AR systems. When cognitive resources are heavily taxed, no single interaction modality can guarantee superior efficiency across users. Instead, allowing users to select the modality that best fits their familiarity and comfort may help reduce subjective effort. Therefore, in tasks with high element interactivity, providing users with flexibility in choosing interaction modalities may help accommodate individual preferences and reduce cognitive load. Conversely, for streamlined tasks characterized by low cognitive complexity, gesture-based interactions may be more effective in enhancing operational efficiency. Adapting input modalities based on task complexity may enable more cognitively sustainable and user-centered AR interface designs. Such findings are particularly relevant in applied AR settings, such as industrial inspection, assembly guidance, or educational simulations, where users often alternate between manual and visual operations. Enabling flexible modality choice in these contexts may help reduce cognitive strain and enhance overall usability.

The successful implementation of the dynamic feedback system in Part 2 demonstrates the potential of biosensor-driven adaptive systems for managing cognitive load in real time based on users' internal physiological states. The analysis showed that the rest of the interventions triggered by elevated cognitive load resulted in statistically significant reductions in subsequent load levels. These findings underscore the efficacy of closed-loop adaptive strategies in dynamically mitigating cognitive demands during task execution. Although CSCR requires further validation, its observed responsiveness in detecting cognitive load and initiating timely interventions highlights its promise for real-time cognitive load monitoring. In practical terms, such adaptive feedback can be beneficial in AR-assisted learning, remote operations, and safety-critical monitoring, where sustained attention is essential and cognitive overload can directly impact decision accuracy or user well-being. This initial evidence supports using physiological markers like CSCR in designing neuroadaptive interfaces, especially in areas where maintaining performance under different cognitive loads is essential.

Our findings resonate with a subtle but important pattern observed in prior AR research: Interventions that reduce cognitive load may sometimes come at the cost of task efficiency. For example, Ghasemi *et al.* (2021) compared head-locked versus world-locked AR modes in a data-entry task. They found that the head-locked mode reduced task time but increased perceived workload, while the world-locked mode, though slower, felt less mentally taxing. More directly, a neurophysiological study using AR-based maintenance instructions revealed that while AR reduced overall task time, it also increased mental workload – as measured by EEG and NASA-TLX – particularly for high-demand tasks (Alessa *et al.* 2023).

These prior observations and our findings highlight a similar performance-cognition trade-off in AR systems – although the adaptive system effectively

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

reduced cognitive load, it marginally increases task completion time. This may be due to disruptions in task continuity introduced by rest interventions, which could interrupt users' concentration. This outcome underscores a central design tension in AR environments: balancing cognitive recovery and task efficiency. In time-sensitive scenarios such as emergency response or surgical procedures, minimizing delay is paramount, and continuous task flow may take precedence. In contrast, contexts like training, prolonged monitoring, or knowledge-intensive tasks may benefit more from cognitive relief, even at the cost of extended duration. Therefore, designers should adjust rest strategies based on the task objectives. Customizing rest thresholds or implementing task-aware intervention rules may serve as valuable mechanisms.

Unlike earlier studies that focus on interface design, our work introduces a physiologically aware rest mechanism. The use of CSCR allows rest to be inserted in response to internal state changes, offering a novel means of supporting cognitive recovery without fundamentally altering task content or interface layout. This approach provides a generalizable framework for future AR and virtual reality (VR) applications that aim to balance user workload dynamically, enabling adaptive pacing and rest scheduling in both professional and educational settings. Future designs may examine whether well-timed, context-aware interventions can achieve cognitive load reduction without negatively impacting task efficiency.

7. Conclusion and future work

This study explored how different interaction modalities and a real-time adaptive rest system affect cognitive load and task performance in AR environments. The results showed three major findings: (1) Gesture-based interactions resulted in faster task completion than gaze-based interactions in simple tasks, while performance converged under higher task difficulty; (2) cognitive load, as reflected by SCR occurrences, positively correlated with task duration; and (3) the proposed adaptive feedback system, guided by real-time CSCR measures, effectively reduced cumulative cognitive load but did not improve task performance due to interruptions in task continuity.

While the adaptive system showed promise, it requires further validation across diverse tasks and populations. The participant sample was relatively small and primarily composed of science, technology, engineering, and mathematics (STEM) students, and the task durations were relatively short. These factors may have constrained the generalizability and temporal scope of the findings.

Future studies will focus on refining the adaptive feedback system, exploring optimal rest strategies under varying task demands. Expanding the participant base, incorporating additional physiological indicators, and applying the model to complex real-world environments will help improve both the accuracy of CSCR and the practicality of adaptive rest-based systems. These efforts aim to enhance the intelligence of adaptive human–computer interfaces and inform future AR and cognitive-aware interaction design.

References

Ahmadi, N. K., **Ozgur, S. F.** & **Kiziltan, E.** 2024 Evaluating the effects of different cognitive tasks on autonomic nervous system responses: implementation of a high-precision, low-cost complementary method. *Brain and Behavior* **14** (10), e70089; doi:10.1002/brb3.70089.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

Alessa, F. M., **Alhaag, M. H.**, **Al-harkan, I. M.**, **Ramadan, M. Z.** & **Alqahtani, F. M.** 2023 A neurophysiological evaluation of cognitive load during augmented reality interactions in various industrial maintenance and assembly tasks. *Sensors* **23** (18), 7698; doi:10.3390/s23187698.

Ballinger, G. A. 2004 Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods* **7** (2), 127–150; doi:10.1177/1094428104263672.

Benedek, M. & **Kaernbach, C.** 2010 A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods* **190** (1), 80–91; doi:10.1016/j.jneumeth.2010.04.028.

Boucsein, W. 2012 *Electrodermal Activity*. Springer Science & Business Media.

Buchner, J., & **Kerres, M.** 2023 Media comparison studies dominate comparative research on augmented reality in education. *Computers & Education*, **195**, 104711; doi:10.1016/j.compedu.2022.104711.

Cai, Y. & **Demmans Epp, C.** 2024 *Predicting cognitive load in language learning with physiological signals. Preprint, arXiv:2405.05543*. https://arxiv.org/abs/2405.05543.

Chen, L., **Zhao, H.**, **Shi, C.**, **Wu, Y.**, **Yu, X.**, **Ren, W.**, **Zhang, Z.** & **Shi, X.** 2024 Enhancing multi-modal perception and interaction: an augmented reality visualization system for complex decision making. *Systems* **12** (1), 7; doi:10.3390/systems12010007.

Chen, O., **Paas, F.** & **Sweller, J.** 2023 A cognitive load theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review* **35**, 63; doi:10.1007/s10648-023-09782-w.

Duchowski, A. T. 2017 *Eye Tracking: Methodology, Theory and Practice*. Springer; doi:10.1007/978-3-319-57883-5.

Ekin, M., **Krejtz, K.**, **Duarte, C.**, **Duchowski, A. T.**, & **Krejtz, I.** 2025 Prediction of intrinsic and extraneous cognitive load with oculometric and biometric indicators. *Scientific Reports*, **15**(1), 5213; doi:10.1038/s41598-025-89336-y.

Lystbæk, M. N., **Rosenberg, P.**, **Pfeuffer, K.**, **Grønbæk, J. E.**, & **Gellersen, H.** 2022 Gaze-hand alignment: Combining eye gaze and mid-air pointing for interacting with menus in augmented reality. *Proceedings of the ACM on Human-Computer Interaction*, **6** (ETRA), Article 145; doi:10.1145/3530886.

García-Hernández, R. A., **Celaya-Padilla, J. M.**, **Luna-García, H.**, *et al.* 2023 Emotional state detection using electroencephalogram signals: a genetic algorithm approach. *Applied Sciences* **13**, 6394; doi:10.3390/app13116394.

Garzón, J. 2021 An overview of twenty-five years of augmented reality in education. *Multimodal Technologies and Interaction* **5** (7), 37; doi:10.3390/mti5070037.

Ghasemi, Y., **Singh, A.**, **Kim, M.**, **Johnson, A.** & **Jeong, H.** 2021 Effects of head-locked augmented reality on user's performance and perceived workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **65** (1), 1094–1098; doi:10.1177/1071181321651169.

Gkintoni, E., **Antonopoulou, H.**, **Sortwell, A.** & **Halkiopoulos, C.** 2025 Challenging cognitive load theory: the role of educational neuroscience and artificial intelligence in redefining learning efficacy. *Brain Sciences* **15**(2), 203; doi:10.3390/brainsci15020203.

Wang, Z., **Rao, M.**, **Ye, S.**, **Song, W.**, & **Lu, F.** 2025 Towards spatial computing: Recent advances in multimodal natural interaction for XR headsets. *Preprint*, arXiv:2502.07598. https://arxiv.org/abs/2502.07598.

Hart, S. G. 2006 NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, pp. 904–908. Human Factors & Ergonomics Society; doi:10.1177/154193120605000909

Hart, S. G. & **Staveland, L. E.** 1988 Development of NASA-TLX (task load index): results of empirical and theoretical research. *Human Mental Workload* **52**, 139–183; doi:10.1016/S0166-4115(08)62386-9.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

Hepsomali, P., **Hadwin, J. A.**, **Liversedge, S. P.** & **Keane, G.** 2019 The impact of cognitive load on processing efficiency and performance effectiveness in anxiety: evidence from event-related potentials and pupillary responses. *Experimental Brain Research* **237**, 897–909; doi:10.1007/s00221-018-05466-y.

Jukiewicz M., **Marcinkowska J.** 2025 Analysis of Electrodermal Signal Features as Indicators of Cognitive and Emotional Reactions—Comparison of the Effectiveness of Selected Statistical Measures. *Sensors*. **25**(11):3300; doi:10.3390/s25113300

Kalyuga, S. 2011 Informing: a cognitive load perspective. *Informing Science* **14**; doi:10.28945/1349.

Kolla, S. S. V. K. & **Plapper, P.** 2023 Interaction modalities for augmented reality applications in manufacturing. In *Advances in Transdisciplinary Engineering*, Vol. **35**, pp. 379–390. IOS Press BV; doi:10.3233/ATDE230063

Kosch, T., **Karolus, J.**, **Zagermann, J.**, **Reiterer, H.**, **Schmidt, A.** & **Woźniak, P. W.** 2023 A survey on measuring cognitive workload in human-computer interaction. *ACM Computing Surveys*, **55**(13s), Article 283, 1–39; doi:10.1145/3582272.

Kumar, M., **Paepcke, A.** & **Winograd, T.** 2007 EyePoint: practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2007*, pp. 421–430. Association for Computing Machinery; doi:10.1145/1240624.1240692

Li, C. & **Liao, T.** 2025 Can I catch up later? Design of personalized intervention for online learning using eye-tracking-based video reconstruction and replay. *Proceedings of the Design Society* **5**, 851–860; doi:10.1017/pds.2025.10099.

Liu, Y., **Sourina, O.** & **Nguyen, M. K.** 2010 Real-time EEG-based human emotion recognition and visualization. In *Proceedings of the 2010 International Conference on Cyberworlds*, pp. 262–269. IEEE; doi:10.1109/CW.2010.37.

Mindfield Biosystems Ltd. 2024 *eSense Skin Response Manual.* https://help.mindfield.de/en/skin-response-manual.

Minkley, N., **Xu, K. M.** & **Krell, M.** 2021 Analyzing relationships between causal and assessment factors of cognitive load: associations between objective and subjective measures of cognitive load, stress, interest, and self-concept. *Frontiers in Education* **6**; doi:10.3389/feduc.2021.632907.

Moncur, B., **Galvez Trigo, M. J.** & **Mortara, L.** 2023 Augmented reality to reduce cognitive load in operational decision-making. In *Augmented Cognition, HCII 2023. Lecture Notes in Computer Science* (ed. D. D. Schmorrow & C. M. Fidopiastis), Vol. **14019**. Springer; doi:10.1007/978-3-031-35017-7_21.

Nandi, A., **Xhafa, F.**, **Subirats, L.** & **Fort, S.** 2021 Real-time emotion classification using EEG data stream in E-learning contexts. *Sensors* **21**, 1589; doi:10.3390/s21051589.

Nguyen, R., **Gouin-Vallerand, C.** & **Amiri, M.** 2023 Hand interaction designs in mixed and augmented reality head-mounted display: a scoping review and classification. *Frontiers in Virtual Reality*; doi:10.3389/frvir.2023.1171230.

Purwinarko, A., **Hardyanto, W.** & **Adhi, M. A.** 2021 Development of learning media for earth physics based on augmented reality as an interactive learning media. *Journal of Physics*; doi:10.1088/1742-6596/1918/4/042131.

Rasch, J., **Wilhalm, M.**, **Müller, F.**, & **Chiossi, F.** 2025 AR you on track? Investigating effects of augmented reality anchoring on dual-task performance while walking. *In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Article 1217, 21 pp., Association for Computing Machinery; doi:10.1145/3706598.3714258.

Rodrigues, M. J., **Postolache, O.** & **Cercas, F.** 2020 Physiological and behavior monitoring Systems for Smart Healthcare Environments: a review. *Sensors* **20** (8), 2186; doi:10.3390/s20082186.

https://doi.org/10.1017/dsj.2025.10040 Published online by Cambridge University Press

Setz, C., **Arnrich, B.**, **Schumm, J.**, **La Marca, R.**, **Tröster, G.** & **Ehlert, U.** 2010 Discriminating stress from cognitive load using a wearable EDA device. *IEEE Transactions on Information Technology in Biomedicine* **14** (2), 410–417; doi:10.1109/TITB.2009.2036164.

Shaffer, F. & **Ginsberg, J. P.** 2017 An overview of heart rate variability metrics and norms. *Frontiers in Public Health* **5**, 258; doi:10.3389/fpubh.2017.00258.

Shi, Y., **Ruiz, N.**, **Taib, R.**, **Choi, E.** & **Chen, F.** 2007 Galvanic Skin Response (GSR) as an index of cognitive load. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery; doi:10.1145/1240866.1241057

Shu, L., **Xie, J.**, **Yang, M.**, **Li, Z.**, **Li, Z.**, **Liao, D.**, **Xu, X.** & **Yang, X.** 2018 A review of emotion recognition using physiological signals. *Sensors* **18** (7), 2074; doi:10.3390/s18072074.

Sibert, L. E. & **Jacob, R. J. K.** 2000 Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems, CHI 2000*, pp. 281–288. Association for Computing Machinery; doi:10.1145/332040.332445

Soshi, T., **Nagamine, M.**, **Fukuda, E.** & **Takeuchi, A.** 2021 Modeling skin conductance response time series during consecutive rapid decision-making under concurrent temporal pressure and information ambiguity. *Brain Sciences* **11**, 1122; doi:10.3390/brainsci11091122.

Sun, Y., **Lu, T.**, **Wang, X.**, **Chen, W.**, **Chen, S.**, **Chen, H.**, & **Zheng, J.** 2023 Physiological feedback technology for real-time emotion regulation: a systematic review. *Frontiers in Psychology*; doi:10.3389/fpsyg.2023.1182667.

Hou, Y., **Xie, Q.**, **Zhang, N.**, & **Lv, J.** 2025 Cognitive load classification of mixed reality human–computer interaction tasks based on multimodal sensor signals. *Scientific Reports* **15**(1), 13732; doi:10.1038/s41598-025-98891-3.

Suzuki, Y., **Wild, F.** & **Scanlon, E.** 2024 Measuring cognitive load in augmented reality with physiological methods: a systematic review. *Journal of Computer Assisted Learning* **40** (2), 375–393; doi:10.1111/jcal.12882.

Sweller, J. 2011 Cognitive load theory. *Psychology of Learning and Motivation* **55**, 37–76; doi:10.1016/B978-0-12-387691-1.00002-8

Yoshida, R., **Nakayama, T.**, **Ogitsu, T.**, **Takemura, H.**, **Mizoguchi, H.**, **Yamaguchi, E.**, **Inagaki, S.**, **Takeda, Y.**, **Namatame, M.**, **Sugimoto, M.** & **Kusunoki, F.** 2014 Feasibility study on estimating visual attention using electrodermal activity. *International Journal on Smart Sensing and Intelligent Systems* **7** (5), 1–4; doi:10.21307/ijssis-2019-050.

Zeger, S. L. & **Liang, K.-Y.** 1986 Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42** (1), 121–130; doi:10.2307/2531248.

Zhang, Y., **Nowak, A.**, **Xuan, Y.**, **Romanowski, A.** & **Fjeld, M.** 2023 See or hear? Exploring the effect of visual/audio hints and gaze-assisted instant post-task feedback for visual search tasks in AR. In *Proceedings of the 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1113–1122. IEEE; doi:10.1109/ISMAR59233.2023.00128

Zhu, X., **Liu, C.**, **Zhao, L.** & **Wang, S.** 2024 EEG emotion recognition network based on attention and spatiotemporal convolution. *Sensors* **24** (11), 3464; doi:10.3390/s24113464.