Human-Computer Interaction & Emerging Technologies, Vol. 233, 2026, 218–227
https://doi.org/10.54941/ahfe1007521

A Wearable AR and Gesture-Ring System for Enhancing Presentation Performance and Reducing Cognitive Load

Shangyi Wu

China Academy of Art, Hangzhou, 31002, China

ABSTRACT

As remote and hybrid work models become widespread, in-person meetings remain crucial venues for expression and communication in knowledge-based work. Speakers must simultaneously convey content, manage structure, and adjust pacing during presentations. Existing presentation aids, primarily linear screen-based text, often require frequent attention shifts, making it difficult to support on-the-fly adjustments and alleviate nervousness. This paper introduces Ring-AR Presenter, an augmented reality-based wearable presentation support system. Combining AR glasses with a gesture ring, it provides low-interference real-time expression support for presenters. The system displays key information through spatialized teleprompting and structural cues within the presenter's field of view. It also enables silent, discreet content control via the gesture ring, minimizing disruption to the presentation flow. Based on a high-fidelity prototype, a simulated presentation experiment involving 20 participants was conducted. Results indicate positive effects in presentation fluency, pacing control, and subjective tension reduction. Findings suggest that presentation assistance combining AR displays with wearable gesture interaction holds application potential in real-world conference and presentation settings.

Keywords: Gesture-based, Hontrol/public, Speaking, Support/Augmented, Reality, AR/Human–Computer Interaction, Wearable interaction

INTRODUCTION

In in-person presentation scenarios such as classroom demonstrations, academic reports, project defenses, and corporate meetings, speakers must manage multiple tasks within a limited time, including content delivery, structural control, and pacing adjustments. Throughout the presentation, speakers continuously produce verbal output while dynamically refining narrative focus based on time progression and audience feedback. This multitasking nature represents a primary source of cognitive load in in-person presentations.

However, information in such settings is often delivered verbally, non-linearly, and at high density, imposing significant cognitive load on speakers during delivery. Existing presentation aids primarily rely on screen-based teleprompters or linear text, requiring frequent shifts in gaze and attention. This not only disrupts the flow of expression but also diminishes natural interaction with the audience. Moreover, these tools provide limited support

Received February 4, 2026; Revised March 29, 2026; Accepted April 15, 2026; Available online July 20, 2026

for on-the-fly pacing changes or content adjustments and are often insufficient for helping speakers maintain control over their delivery in complex scenarios.

In real presentations, speakers frequently experience nervousness, pacing imbalances, or the need for impromptu narrative restructuring. Traditional presentation support systems, however, largely depend on static content presentation and lack dynamic prompts and feedback mechanisms tailored to the speaking context.

In recent years, advancements in augmented reality (AR) and wearable interaction technologies have opened new possibilities for designing presentation support tools (Radianti et al., 2020; Azuma, 1997). Some studies have integrated AR technology into public speaking and oral communication training. By employing virtual audiences, immersive feedback, or scenario simulations, these systems help users build confidence and alleviate nervousness during presentations (Jim et al., 2019). Such systems primarily focus on training effectiveness during pre-presentation or rehearsal phases, emphasizing improved psychological states and expressive abilities through immersive experiences. However, they offer limited support for dynamic pacing adjustments and real-time feedback during actual presentations.

Other research explores dynamically generating and presenting information based on speech and context during task execution or face-to-face communication (Gao et al., 2024), thereby reducing attention switching and cognitive load. These systems empha size spatialized information presentation and contextual relevance, providing theoretical foundations for real-time cues and structural guidance during speeches. However, such studies predominantly focus on general conversations or collaborative tasks, lacking specialized designs tailored to speech scenarios.

Regarding interaction methods, wearable devices like gesture rings and wristbands can capture fine-grained gestures, enabling silent, discreet, and low-interference operation (Rekimoto, 2001). Related studies indicate that gesture interaction combined with AR displays can enhance interaction fluidity and reduce interference with the primary task. However, such methods in speech assistance scenarios are mostly single-function explorations, lacking integration with spatial cueing and structural guidance, and lacking systematic empirical research.

In summary, existing research on presentation assistance exhibits several limitations: most systems focus solely on presentation training or single-function support, lacking real-time comprehensive assistance during actual presentations; wearable gesture input and spatial cues remain unintegrated; and support for speaker pacing control and emotional regulation remains inadequate. These research gaps provide clear motivation and design direction for the comprehensive presentation assistance system proposed in this paper, based on AR glasses and gesture rings.

To address these challenges, this paper introduces Ring-AR Presenter, an AR-assisted system tailored for live presentations and conference settings. Integrating AR glasses with a wearable gesture ring, the system delivers real-time expressive support through spatialized teleprompting, structural cues, and natural gesture control. Its goal is to reduce cognitive load while enhancing expressive fluency and pacing control.Specifically, this study addresses three research questions:

- **RQ1**: whether spatialized AR cueing and structural guidance can reduce presenters' cognitive load without introducing additional distractions.
- **RQ2**: whether the silent, discreet interaction enabled by wearable gesture rings more effectively supports real-time content structuring compared to traditional screen-based controls.
- **RQ3**: whether an integrated presentation assistance system combining AR visual cues and gesture interaction can alleviate presentation anxiety under high-pressure or tense conditions.

The main contributions of this paper are threefold: (1) Designing an AR spatial teleprompter and guidance interface tailored for live presentation scenarios; (2) Proposing a wearable gesture ring-based interaction method to achieve discreet and highly fluid presentation content control; (3) Evaluating the system through user experiments, validating its effectiveness in enhancing presentation fluency, pacing control, and alleviating nervousness.

SYSTEMS DESIGN

Ring-AR Presenter is designed as a wearable presentation assistance system for offline conferences and live speaking scenarios. It aims to provide speakers with real-time, adjustable, and low-intrusive expressive support without disrupting their presentation flow. The system employs a closed-loop interaction architecture of "perception-generation-presentation-control," comprising a wearable gesture ring, a backend intelligent processing module, and AR glasses working in tandem.

Within this architecture, the gesture ring serves as the primary input device, capturing the speaker's interactive intent during presentations. The ring employs fine-grained gesture recognition to detect user control actions—such as swiping, rotating, or pinching—and transmits these inputs to the backend processing module. Computational tasks and content generation are primarily handled by the backend system on a PC or cloud platform. This includes managing speech content status, scheduling cue information, and generating teleprompter text and structural prompts based on multimodal language models. This approach avoids imposing high computational loads on the wearable device itself.

AR glasses serve as the output terminal, primarily responsible for information presentation and interface display without engaging in complex computations. The system spatially projects generated cue content and structural prompts into the speaker's field of view, enabling access to critical information without looking down or frequently shifting attention. This lightweight display design enhances device battery life and wearing comfort while minimizing disruption to the presentation flow.

During the presentation, the system continuously updates prompts based on content progression and user gesture inputs, creating a real-time closed-loop interaction between visual cues and gesture control. The overall system architecture comprises three core modules: (1) Spatial AR Prompting and Structural Guidance Interface, presenting presentation content and structural information; (2) Gesture-Driven Interaction Module via Wearable Ring, enabling silent, discreet content control; (3) An intelligent scheduling module for content generation and state management, coordinating content output, system responses, and interface updates. These modules operate synergistically to achieve design objectives of reducing cognitive load, enhancing pacing control, and alleviating presentation anxiety.

Spatial AR Prompting and Guidance Interface

Addressing issues inherent in traditional linear teleprompters—such as frequent attention shifts and rigid information presentation during live speeches—this system explores spatial and layered information delivery within an AR environment. It aims to provide speakers with support that better aligns with their natural delivery rhythm. The design prioritizes embedding essential presentation information into the speaker's natural field of view with low visual prominence. This allows speakers to maintain holistic control over content structure, pacing, and delivery status without interrupting their flow or relying on frequent eye shifts.

The system supports not only the live presentation itself but also pre-speech training simulations and post-speech review analysis, forming a comprehensive presentation support workflow centered around "Preparation-Presentation-Review."

The system comprises three core functional modules interconnected through unified interface logic:

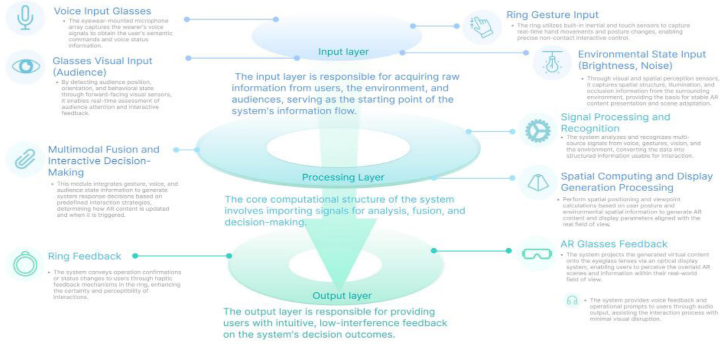

Figure 1: Technical framework.

(1) Presentation Training and Simulation Module

Prior to formal presentations, the system offers reusable training and simulation capabilities. Speakers can enter simulated presentation mode via AR glasses, where the system generates virtual audience distributions and spatial environments based on preset scenarios to replicate diverse presentation scales and formats.

During training, speakers practice with full-sentence-level cueing support. The system records speech rate variations, pause locations, and overall time allocation, providing data for post-performance reviews. The simulation emphasizes "managed pressure," using virtual audiences and stable environmental feedback to gradually acclimate speakers to public speaking states without inducing additional anxiety.

Figure 2: The traning simulation module.

(2) Live Presentation Assistance Module

During actual speeches or meetings, the system provides real-time support to presenters via AR glasses. Prompts appear as complete sentences, serving as structural and expressive references rather than verbatim reading cues. Their display position and prominence are carefully controlled to avoid obstructing vision or disrupting delivery rhythm.

Structural information (e.g., chapter progression, key topic nodes) appears as subtle prompts along the periphery of the field of view, enabling speakers to maintain awareness of overall content flow without frequent script recall or verification.

Additionally, the system introduces a virtual audience mechanism within the AR field of view to alleviate psychological pressure when facing real listeners. This mechanism visually replaces the actual audience, functioning exclusively within the speaker's AR glasses without external display or interference with real-world interactions.Virtual attendees match the spatial distribution and quantity of real attendees, but their appearance and behavior are designed to be neutral, stable, and non-judgmental. This includes minimizing facial expressions, reducing direct eye contact, or using abstracted human silhouettes. This approach reduces the perceived social evaluation pressure on the speaker without altering the actual presentation structure, allowing greater focus on content organization and delivery.

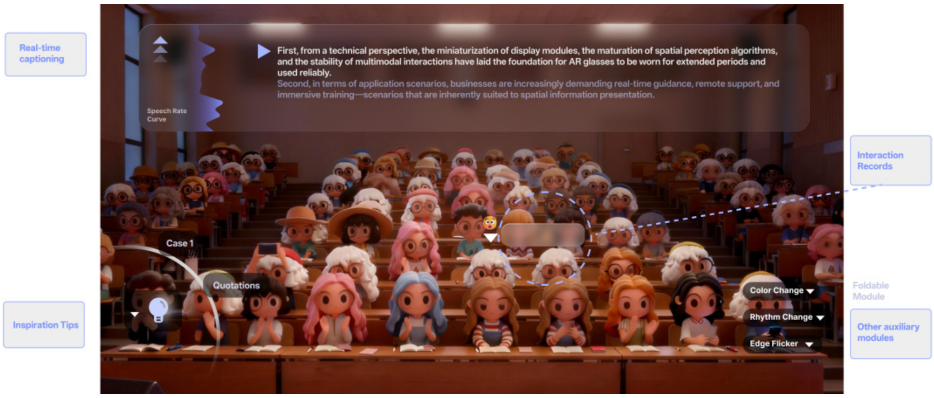

Figure 3: The speech prompter module.

Figure 4: The virtual avatar customization module.

This design does not aim to eliminate the speaker's awareness of the audience's presence. Instead, it reduces the intensity of social evaluation cues to help speakers maintain their pacing and cognitive stability in high-pressure situations, allowing greater focus on content organization and delivery. Pacing and state feedback is conveyed through implicit visual cues, such as subtle positional shifts in prompt elements, transparency adjustments, or progressive fade-ins/fade-outs, reflecting current speaking speed or time allocation. All feedback outputs are presented with low prominence, avoiding explicit warnings or interruptive prompts to preserve the continuity and natural flow of the presentation.

(3) Post-Presentation Review and AI-Assisted Module

Following the presentation, the system enters a review phase. It conducts a structured analysis of the speech based on recorded audio, pacing, and operational data, visually highlighting key points such as pacing fluctuations, time allocation, and content progression status.

Building on this, the AI-assisted speech module compares the original draft with the actual delivery to help speakers identify redundant, disjointed, or unevenly paced sections. It provides actionable, modifiable suggestions for future speech preparation and training.

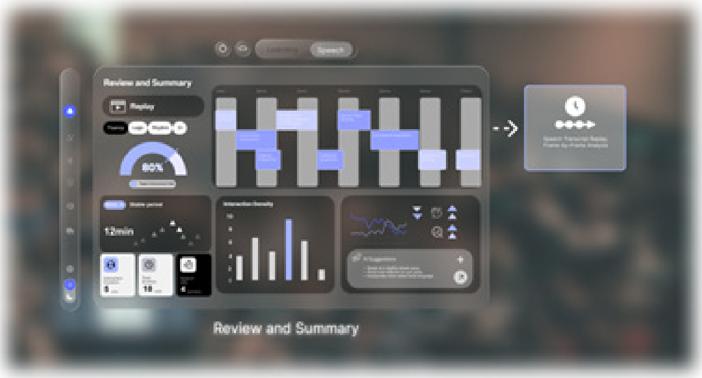

Figure 5: The review and improvement module.

Gesture-driven Interaction via Wearable Ring

In public speaking scenarios, traditional touch or voice control methods can easily draw unwanted attention or interrupt the flow of expression. This module aims to provide a silent, discreet, and highly fluid interaction method, enabling speakers to adjust content structure and prompt states in real time without interrupting their delivery.

The system introduces a wearable gesture ring as the primary input device, allowing speakers to control presentation content in real time without relying on screens or voice commands. The ring is worn on the speaker's thumb and index finger. Interactions are performed by the thumb touching the side of the index finger, pressing or releasing the finger, and natural palm opening/closing motions. These actions originate from everyday hand movements, offering excellent naturalness and controllability, making them suitable for sustained use during presentations.

Figure 6: The gesture interaction module.

Through a set of predefined, easily distinguishable hand gestures, presenters can switch sections, highlight key points, navigate content forward or backward, and display or hide supplementary information. Since interactions primarily occur between fingers or within the palm—requiring no conspicuous limb movements—operations can be performed without disrupting the flow of expression or drawing external attention.

By shifting content control from methods requiring constant screen focus or verbal commands to lightweight, discreet hand gestures, the system reduces the additional demand on visual and verbal resources during presentations. This helps speakers maintain a more stable state of expression, thought, and content management.

EVALUATION

To systematically evaluate the user experience and support effectiveness of the proposed Ring-AR Presenter in live presentations and conference settings, this study employs a mixed-methods approach combining questionnaire assessments and semi-structured interviews. It analyzes speakers' subjective experiences across both quantitative and qualitative dimensions.

This evaluation approach aims to comprehensively examine the system's impact on cognitive load, expressive fluency, pacing control, and emotional experience (Zhang and Yao, 2026), while further understanding users' actual perceptions of system design and interaction methods.

Particpants

This study recruited 20 participants (10 female, 10 male), all possessing prior experience in offline presentations such as classroom demonstrations, academic reports, or project defenses. Participants ranged in age from 18 to 35 years old (M = 24.3, SD = 3.8). All participants had no prior exposure to the system before the experiment to prevent pre-existing knowledge from influencing evaluation outcomes.

Procedure

The experiment employed a simulated presentation task format. Before the experiment began, researchers briefly introduced the system's basic functions and usage methods to participants, followed by a short familiarization period. Subsequently, participants wore AR glasses and a gesture ring to complete a simulated presentation task on a predetermined topic.

During the presentation, participants could freely utilize the system's spatialized teleprompter, structural cues, and gesture control features to navigate content and adjust pacing. After completing the presentation, participants sequentially completed a questionnaire evaluation and participated in a semi-structured interview.

Questionnaire-Based Evaluation

Following the simulated speech task, participants complete a questionnaire to quantify their subjective experience during the presentation. The questionnaire employs a 7-point Likert scale (1 = "Strongly Disagree," 7 = "Strongly Agree") to assess the system's support effectiveness across multiple key dimensions.

The questionnaire primarily covers the following aspects:

(1) Cognitive Load: Assesses psychological stress and attention demands experienced during the presentation, focusing on cognitive burden related to content recall, structural control, and system usage.
(2) Expressive Fluency: Evaluates the coherence and smoothness of the presentation flow, including whether the system reduces pauses, stuttering, or interruptions.
(3) Pacing and Content Control: Measures participants' subjective sense of control over presentation progress, structure, and overall rhythm.
(4) Emotional Experience: Focuses on tension levels and psychological comfort during presentations, analyzing the system's potential role in alleviating presentation anxiety.
(5) System Usability and Acceptance: Evaluates the system's ease of use, learning curve, and willingness to use it in future real-world presentation scenarios.

Additionally, the questionnaire includes open-ended questions to gather participants' subjective feedback on the system's strengths, weaknesses, and potential application scenarios, supplementing the interpretation of quantitative results.

Semi-Structured Interview

Following the questionnaire assessment, semi-structured interviews were conducted with participants to gain deeper insights into their overall experience and subjective perceptions during system use. Interviews centered on user experience, emphasizing the system's support mechanisms during presentations, interaction feedback, and potential improvement directions. Each interview lasted approximately 10–15 minutes and was recorded with participant consent.

Key interview questions included:

(1) Did the system assist in organizing and structuring presentation content?
(2) Did spatial AR teleprompting and structural cues distract attention or disrupt on-site interactions?
(3) Was content control via the gesture ring intuitive? Did it affect presentation movements or pacing?
(4) Did the virtual avatar help alleviate presentation anxiety?
(5) The system's applicability across different presentation scenarios and areas for improvement.

Results

This study analyzed the effectiveness of the proposed system in speech scenarios through a combination of questionnaire surveys and semi-structured interviews.

Questionnaire results indicate that participants' ratings across dimensions—cognitive load, expressive fluency, perceived control over pacing and content, and emotional experience—were predominantly positive. This suggests the system did not significantly increase psychological burden under current experimental conditions and to some extent supported content organization and expressive performance during presentations. Furthermore, most participants gave high ratings for system usability and acceptability, indicating the system holds potential for further application in real-world presentations or training scenarios.

Interview findings further supplemented the questionnaire results. Multiple participants noted that spatialized AR teleprompting and structural cues helped them maintain awareness of the overall structure during their speeches without significantly disrupting attention or on-site interactions. One participant remarked: "These cues remained in my field of vision without demanding my attention, allowing me to know what to say next."

Regarding emotional experience, some participants felt the introduction of virtual avatars somewhat reduced nervousness when facing a real audience, enabling greater focus on content delivery. One participant stated: "With virtual avatars replacing a real audience, I wasn't constantly concerned about others' reactions and felt more confident to keep speaking."

Additionally, participants generally agreed that the combination of the gesture ring and AR glasses provided a low-interference, empowering interaction method (Kim et al., 2019). One participant noted: “The ring and glasses combo made me feel in control of the entire presentation. That sense of control boosted my confidence.”

Synthesizing quantitative and qualitative findings, the proposed system demonstrates overall positive support for presentation experiences under current prototype conditions. It shows particular potential in alleviating nervousness, enhancing control over the presentation process, and maintaining expressive coherence.

CONCLUSION

This paper addresses the cognitive load, pacing control, and nervousness challenges faced by speakers in offline presentations and conferences. We designed and implemented Ring-AR Presenter, a presentation assistance system utilizing AR glasses and a wearable gesture ring. By integrating spatial AR teleprompting and structural cues with silent, discreet gesture interactions, the system provides real-time content references and process control support without disrupting the presentation flow.

Research findings indicate that the system effectively reduces cognitive strain during presentations, helps speakers maintain awareness of content structure and pacing, and alleviates nervousness when facing audiences. By embedding cues within the speaker’s natural field of view and shifting control operations to low-visibility hand gestures, Ring-AR Presenter demonstrates the application potential of wearable AR technology as an auxiliary tool in real-world presentation scenarios. This research provides practical insights for designing interaction methods and optimizing user experience in presentation assistance systems, laying the groundwork for future exploration of smarter, context-aware tools tailored to public speaking scenarios.

REFERENCES

Azuma, R.T. (1997). A survey of augmented reality. Presence: Teleoperators and Virtual Environments, 6(4), 355–385.

Gao, X., et al. (2024). Context-aware augmented reality assistance for real-time face-to-face conversations.

Jim, M.E., et al. (2019). Speak with confidence: Designing an augmented reality training tool for public speaking. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.

Kim, M., Choi, S.H., Park, K.-B., Lee, J.Y. (2019). User interactions for augmented reality smart glasses: A comparative evaluation of visual contexts and interaction gestures. Applied Sciences, 9(15), 3171.

Radianti, J., Majchrzak, T.A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers & Education, 147, 103778.

Rekimoto, J. (2001). GestureWrist and GesturePad: Unobtrusive wearable interaction devices. Proceedings of the Fifth International Symposium on Wearable Computers (ISWC), 21–27.

Zhang, X., Yao, Y. (2026). Can augmented reality technology reduce learners’ cognitive load? A metaanalysis, Smart Learning Environments, 13, 9.