Int. J. Human–Computer Studies 198 (2025) 103467

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

AF-Mix: A gaze-aware learning system with attention feedback in mixed reality

Shi Liu *, Peyman Toreini, Alexander Maedche

Human-Centered Systems Lab, Karlsruhe Institute of Technology, Kaiserstrasse 89-93, Karlsruhe, 76133, Germany

ARTICLE INFO

Keywords:
Mixed reality
Eye tracking
Attention feedback
Learning

ABSTRACT

Mixed Reality (MR) has demonstrated its potential in various learning contexts. MR-based learning environments empower users to actively explore learning content visualized in multiple formats, such as 3D models, videos, and images. Nonetheless, the sophisticated visualizations in MR learning environments may result in potential visual overload, posing a challenge for users in efficiently allocating their attention. In this paper, we present AF-Mix, a learning support system that leverages eye tracking sensors in Microsoft HoloLens 2 to offer attention feedback for learners. Aiming to design AF-Mix, we conducted a participatory design study and integrated the attention feedback into our system, following users' needs and suggestions. Furthermore, we evaluated AF-Mix in an evaluation study (n = 22) following a quantitative analysis of users' visual behavior, as well as a qualitative analysis of interview transcripts. Our findings show that providing feedback to support the learning process can be achieved effectively with eye tracking. In specific, attention feedback assists learners in retrieving previously missed information and encourages learners to reallocate their attention in the review process. Moreover, providing personalized feedback based on previous attention allocation is more effective in supporting users than a self-review approach without gaze-aware assistance in MR. Such feedback facilitates users in managing their limited attentional resources better and supports the reflection of their learning journey more effectively.

1. Introduction

Mixed Reality (MR) for learning has been intensively explored in recent years (Arici et al., 2019; Maas and Hughes, 2020). For example, it has been applied in medical education to help students better understand anatomy structure (Huang et al., 2018), in math education to create a playful learning scenario (Khan et al., 2018), or in simulation learning to present immersive environments (Ibáñez et al., 2016). In general, MR learning systems come with several advantages, including their potential to visually present and engage learners with abstract concepts to promote a more comprehensive understanding (Khan et al., 2018; Radu and Schneider, 2019). Consequently, MR learning environments can better motivate learners and improve their overall learning performance (Khan et al., 2019).

Nonetheless, MR learning environments also come with significant limitations, including challenges for learners in attention management and overcoming cognitive overload (Vovk et al., 2018; Knierim et al., 2018). Meanwhile, most MR learning environments follow a constructivist approach, requiring active learning (Hanid et al., 2020), and such an approach also requires sufficient guidance and assistance (Mayer, 2004). Therefore, a need for more effective support for learners in MR environments has been recognized (Thoravi Kumaravel et al., 2019). Besides, users sometimes encounter visual overload and are challenged to manage their attention during the learning process appropriately (Knierim et al., 2018). In physical classrooms, this is typically the task of the teacher to structure the learning materials and provide feedback, which is essential if learners are expected to establish a comprehensive understanding of the learning topic (Quintana et al., 2004).

While some existing MR learning systems are designed for tutoring or collaborative learning with multiple users (Huang et al., 2021; Radu et al., 2021), other systems expect users to learn independently without help from fellow learners or teachers in the MR environment (r and r, 2017; Mohammadhossein et al., 2022). Successful learning outcomes in such scenarios require effective self-regulated learning skills. In the framework of self-regulated learning, designing an effective learning system requires consideration and support of various cognitive and metacognitive activities, such as goal-setting, planning, monitoring, etc. (Panadero, 2017). Among these activities, providing feedback to students is crucial for cognitive support during the learning process (Butler and Winne, 1995). While feedback can be achieved in

* Corresponding author.
E-mail address: shi.liu@kit.edu (S. Liu).

https://doi.org/10.1016/j.ijhcs.2025.103467
Received 2 August 2023; Received in revised form 29 November 2024; Accepted 11 February 2025
Available online 20 February 2025

various ways, we focus specifically on providing attention feedback for MR learning systems in our study. This is due to the reported challenges of attention management in the immersive learning environment (Han et al., 2022; Grogorick and Magnor, 2020). Embedding attention feedback to support learning in MR environments to the best of our knowledge has been not investigated so far.

We provide attention feedback in MR learning systems to support users in managing their attention allocation and promote a more efficient review process of their learning outcomes. Therefore, we name our system *AF-Mix*. Our human-centered design process focused on collecting user needs and potential design solutions in participatory workshops. After implementing the system in Unity and deploying it on HoloLens 2, we evaluated the impact of attention feedback on the review process with a user evaluation, analyzing eye tracking and interview data collected from 22 participants in a between-subject experiment. The data analysis, using a Bayesian t-test framework, suggests that *AF-Mix* enhanced awareness by helping learners retrieve previously missed information while having no significant impact on learners' processing time during the review. Furthermore, participants perceived *AF-Mix* as supportive and effective for learning in MR. Participants also reported that such a system can benefit learners of different learning styles and encourage learners to optimize their review strategy.

Our main contribution in this paper is a new approach for presenting attention feedback using eye tracking in MR learning environments. The results of our user evaluation suggest that *AF-Mix* supported participants in obtaining an overview of their learning progress and successfully helped them to better allocate their attention in the review process. By conducting user evaluation and synthesizing user feedback, we contribute to the HCI community with design recommendations and guidelines for gaze-aware learning support in MR systems, as well as insights on supporting self-monitoring and self-reflection in MR learning environments.

2. Related work

The term MR was originally coined to indicate a spectrum of virtual environments, including Augmented Reality (AR) (Milgram and Fumio, 1994). However, researchers have interpreted the term differently over the past decade (Speicher et al., 2019). One attempt is to define MR as "strong AR", as reported by Speicher et al. (2019), and it has been adopted in works of Yue et al. (2017) and Maas and Hughes (2020). According to this definition, MR environments should create an enhanced user experience compared to their mobile-AR counterparts, constructing a blended and interactive environment of physical surroundings and virtual elements, not just static virtual overlays on top of the physical environment. In this paper, MR specifically refers to using head-mounted displays (HMD) to construct immersive AR environments with extended features such as eye tracking.

Using MR for education and learning has been explored in many fields, such as medical education (Moro et al., 2021; Bianchi et al., 2020), science education (Arici et al., 2019; Radu and Schneider, 2019), language learning (Shao et al., 2020), K-12 education (Maas and Hughes, 2020), collaborative programming (Radu et al., 2021), etc. Research has shown that using MR for educational purposes has the potential to improve students' motivation and learning outcomes by making abstract ideas and concepts more tangible (Khan et al., 2019). With MR, students can interact with the virtual learning material naturally, as they would do when learning with physical material. For example, Radu and Schneider (2019) implemented a collaborative learning environment with HoloLens 1, where abstract concepts in physics are visualized in MR (e.g., magnetic fields). Students can easily explore the "dynamic nature of relationships between important variables" in such learning environments (Radu and Schneider, 2019). The importance of giving and receiving feedback in such MR environments is also addressed in their work, specifically the feedback given by other collaborators during the learning process. As another example, Vazquez et al. (2017) proposed language learning by using MR. Compared to traditional learning formats such as phrasebooks, serendipitous language learning is achieved in MR with extensive contexts. For instance, when users discover a physical object in the environment, the MR system not only displays the corresponding word for the object but also augments the physical object with additional contextual information by "embedding 3D models" or a movie clip to the object (Vazquez et al., 2017). Similar to the previous example, Hennerley et al. (2017) implemented an MR application that supports the understanding of abstract concepts in programming with 3D visualizations. Therefore, one major advantage of learning in MR environments is the following: MR systems support students in understanding intangible concepts by embedding 3D visualizations and various formats of learning material in the learning environment. Besides, MR devices motivate kinesthetic learning, a learning strategy that by definition requires physical activities to actively engage in the learning task (Iqbal et al., 2019), whereas other learning formats such as mobile-AR often hinder it (Radu and Schneider, 2019; Oh et al., 2018).

Even though MR systems have advantages over other learning formats in some contexts, researchers have pointed out the limitations of MR learning systems as well. Optimizing cognitive workload caused by information overload, addressing usability issues, and reducing motion sickness are challenges for designing MR learning environments (Vovk et al., 2018; Vergel et al., 2020). In response, some research has explored providing learning support in MR systems. For instance, Giraudeau et al. (2019) claim that intuitive interactions in MR can reduce cognitive overload during the learning process. Ibáñez et al. (2016) adopt scaffolding in AR simulation learning systems to guide users in exploratory learning. However, little research has been done to provide feedback that increases users' self-awareness and supports them in optimizing their review strategies. We believe that this research gap prevents learners from fully utilizing the advantages of MR learning environments.

Compared to traditional learning environments, MR systems often present learners with a self-directed and constructivist scenario, including ones developed by Radu and Schneider (2019) and Khan et al. (2018). Constructivism focuses on motivating learners to actively engage in learning activities and construct their understanding of the knowledge proactively (Schunk, 2012). Examples of constructivist learning strategies include inquiry-based learning, experiential learning, kinesthetic learning, etc. These learning strategies have been commonly adopted for designing MR or AR learning environments (Huang et al., 2016; Radu and Schneider, 2019; Radu et al., 2021). Meanwhile, education research suggests the necessity of providing feedback in constructivist learning environments to promote better self-regulation (Kirschner et al., 2006). Such feedback can increase users' self-awareness and support in improving their learning performance. Receiving feedback and adjusting learning strategies are also seen as an integral part of self-regulated learning (Butler and Winne, 1995).

Nonetheless, most MR learning systems have so far focused on creating innovative interaction and experience during the learning journey, while integrating effective self-awareness feedback for learning in MR has been overlooked. This research gap has been recognized by Thoravi Kumaravel et al. (2019), as they criticize the lack of feedback in most MR learning environments. As a response, they integrate learning feedback in their system by implementing "bidirectional mixed-reality telepresence" tutoring (Thoravi Kumaravel et al., 2019), allowing tutors to give asynchronous feedback on learners' performance. Furthermore, another study conducted by Shao et al. (2020) focuses on American sign language education, which also emphasizes the importance of providing feedback during the learning process, especially in the context of motion learning. The feedback discussed in these studies focuses on addressing whether a user action is correctly performed or not, instead of examining users' attention distribution and promoting self-awareness in the learning process. Moreover, these systems are typically designed for multi-user learning environments. We argue that there is a research gap in providing feedback on users' attention during the learning process in self-directed and single-user MR learning environments.

Attention plays an important role in learning, and managing attention in constructivist learning environments is crucial for learning success (Schunk, 2012). Existing studies have leveraged eye tracking to provide attention feedback in non-MR learning contexts (D'Mello et al., 2012; Sharma et al., 2016). In a review of studies that used eye tracking for learning support, Lai et al. (2013) suggest that eye tracking is a "promising channel for educational researchers to connect learning outcomes to cognitive processes". To the best of our knowledge, providing attention feedback has not been well-studied in MR settings. Therefore, we aim to fill this research gap by designing a gaze-aware system that presents attention feedback to users in MR learning environments.

3. Method

Our study aimed to introduce attention feedback into MR learning systems. For this goal, we followed an iterative research process. Overall, we organized our research process in three phases. Each phase contributes to the development and evaluation of the attention feedback system, denoted as *AF-Mix*. The first phase focused on unveiling user needs and establishing design goals. We conducted a participatory design study and organized three design workshops. The goal of the workshops was to reveal user challenges in using MR learning systems and elicit potential design ideas for *AF-Mix*.

The second phase revolved around the practical implementation of *AF-Mix* based on design solutions derived from the preceding phase. This resulted in the extension of an existing MR learning system with gaze-aware learning support, i.e. attention feedback. The feedback included an analytical overview of learners' attention distribution (Fig. 6-1) and color-coded flashcards (Fig. 6-3) reflecting learners' prior attention allocation.

The final phase comprised a comprehensive evaluation of *AF-Mix* through a dedicated user study. The primary objective was to investigate the impact of attention feedback on learning strategy and outcomes, particularly its influence on the review process during the learning experience. Employing mixed methods, we analyzed eye tracking data and interview transcripts. For the evaluation, we recruited 22 participants and divided them into two groups: Participants in the treatment group received attention feedback after exploring the learning material, while participants in the control group did not.

For the data analysis, we adopt the Bayesian t-test framework for the quantitative eye tracking data (Rouder et al., 2009). Specifically, we aim to determine whether the attention feedback enhances learners' awareness of content previously overlooked during the learning process. Additionally, we explore whether the introduction of attention feedback influences the processing time during the subsequent review phase. Complementing our quantitative analysis, we also conducted interviews to gain insights into the user experience. The interviews had been transcribed and coded by two independent coders in an opening coding process before we conducted a thematic analysis based on the guidelines proposed by Braun and Clarke (2006).

4. Participatory design

Aiming to understand user needs and define our design goals, we conducted a participatory design study. For the participatory design, we organized three design workshops and recruited 15 students (6 female, 9 male) to generate ideas and design solutions. Students were recruited voluntarily from a Human-Computer Interaction (HCI) lecture at the local university. Participation in the workshops did not influence the outcomes or grades for the lecture. The recruitment process received approval through an Institutional Review Board (IRB) review process. The inclusion and exclusion criteria for the workshops include the following (See Table 1):

We used the 6-3-5 brainwriting method to collect ideas and sketches from participants. After the workshops, we summarized user needs and design goals.

Based on the previous work of Liu et al. (2022), we used their MR learning system for HCI education as the foundation of our participatory design. In their MR learning system, users can explore the established HCI concept of Model Human Processor (MHP) proposed by Card (1981), using multimodal interaction such as gesture and eye-gaze-based interaction. In the end, the system will present users with a review option, which allows users to rediscover previously missed learning content. Nonetheless, the system has not been systematically evaluated and the review option requires a more effective design. We believe the learning content is well-suited for the participants, as the concept is directly relevant to the HCI lecture they were attending and had not yet been covered in their lecture at the time of the design workshops. Apart from this system, we also presented various other MR learning systems to the participants, including HoloPatient[1], HoloAnatomy (Wish-Baratz et al., 2019). The goal is to reduce the priming effect of one particular MR learning system on participants during the design workshops.

The goals of our participatory design are the following: First, we aim to reveal user challenges in learning with existing MR systems and validate the potential attention management difficulty indicated in the related work. Second, we would like to collect ideas and design solutions produced by the participants to effectively integrate attention feedback in an MR learning system. For the participatory design study, 15 participants (6 females) were recruited voluntarily and divided evenly into three sessions. All participants were undergraduate students taking an HCI lecture and had a basic understanding of HCI-related learning content and MR technology. Each session lasted around one hour and was organized in a meeting room at the local university. Before the workshops, participants were first informed about the study procedures and gave their consent.

The design workshops unfolded with the following procedure: First, participants were given a brief introduction to MR technology and the HoloLens 2 device. They were invited to use the device hands-on to familiarize themselves with gesture- and eye-gaze-based interaction techniques afterwards. Next, participants explored a selection of existing MR learning applications (HoloPatient, HoloAnatomy, the MHP learning application) at their own pace, allowing them to experience the current design and use cases of MR in education. Following this, we introduced the concept of leveraging eye-tracking technology to support attention management during learning. To minimize potential biases, instead of showcasing existing gaze-aware learning support systems, we only provide basic information to participants: we first demonstrated the technical capabilities of the HoloLens 2 eye-tracking sensors. Afterwards, we provided a concise explanation of attention management principles in learning, informed by current research (Roda and Nabeth, 2007).

Lastly, participants generated and exchanged ideas on how to design attention feedback using the 6-3-5 brainwriting method. The 6-3-5 brainwriting method is a classic ideation technique originally proposed for product innovation and later adopted in interdisciplinary research (VanGundy, 1984; Heslin, 2009). In HCI research, it has been used as an effective participatory method (Muller and Kuhn, 1993; Boy, 1997), and researchers have revealed its advantages for group elicitation and preventing the ideation process from being dominated by verbally dominant participants (Wilson, 2013). For our design workshop, we followed the guidelines proposed in prior research on participatory design (Muller and Kuhn, 1993; Hutt et al., 2021). In our

[1] HoloPatient is an MR learning system developed by GIGXR: https://www.gigxr.com/holopatient/.

Table 1
Inclusion and exclusion criteria for workshop participants.

Inclusion criteria	Exclusion criteria
Willing to test and provide feedback in group discussions	Severe motion sickness, epilepsy, or other medical conditions
Enrolled as a full-time university student	Limited English proficiency to understand learning material
Proficient in English (CEFR B2)	Unable to participate due to constraints, e.g. COVID-19

Fig. 1. One session of the participatory design workshops.

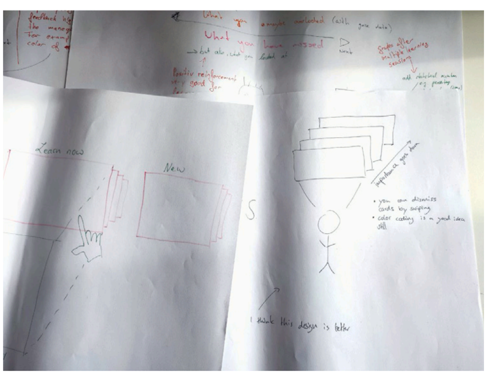

Fig. 2. Brainwriting worksheets produced by participants.

workshop, a brainwriting session typically lasted around 30 min. In this session, participants formulated their ideas by making sketches or writing them down on a worksheet. The session followed a repetitive process consisting of five rounds. In the first round, each participant contributed three ideas on his or her worksheet within five minutes. After the first round, the worksheets were passed around to the participants so that all participants could review others' ideas and engage in this additive process. The session finished when every participant had reviewed the worksheets of all other participants. After the brainwriting session, participants discussed the final results and suggested potential design solutions. Fig. 1 shows participants exchanging design ideas using the brainwriting technique. After organizing three participatory design workshop sessions, the results were collected and synthesized from the worksheets.

4.1. Findings

We coded and clustered all ideas generated by the participants into three themes, representing the main user needs:

1. Attention feedback should be presented in a simple and straightforward way in the review process, without causing additional cognitive overload.
2. The review process should provide a different learning approach from the previous explorative learning process.
3. The review process including attention feedback should help learners to effectively process previously missed information.

Many ideas attempt to reduce users' cognitive overload. The first two themes emphasize this user need. Participants commonly recognized a difficulty users may face in MR environments: cognitive overload. After reviewing the learning system developed by Liu et al. (2022), participants assumed that the initial exploratory learning process requires a high mental workload. Therefore, participants suggested that the review process and attention feedback should feature a simple design, which does not require much effort. Besides, design ideas contributed by participants highlight that a clear distinction between the exploration and the review process should be clarified as affordance in the system design, assisting learners to adjust their learning strategies accordingly. Based on participants' insights, we concluded one potential design solution: the initial learning process may allow users to explore learning material in an arbitrary sequence, making the learning process more autonomous and flexible. Thereafter, the review process can present a predefined learning path that guides users in covering content they might have missed previously.

Overall, the collected ideas suggest a unique design to present attention feedback in a simple format. As one participant wrote in her worksheet, attention feedback should "reduce users' mental workload with a simple interface" instead of further overwhelming users. Additionally, the third theme suggests that participants wanted more than just receiving information on their attention. Particularly, participants wished for additional support for self-reflection in the review process after receiving the attention feedback. One participant explained this during the discussion:

> In my opinion, the goal of getting feedback is to know more about your weakness in the process. If the system is already capable of telling me what I missed, is it possible to help me review and memorize what I missed? (...) Fixing your weakness is more important than finding it out. (P4)

Sketches (see Fig. 2) produced by the participants also corroborate the clustered themes. In the design sketches produced by the participants, one idea attempts to address the user needs in all clustered themes. The idea is to present attention feedback and review options as a deck of virtual flashcards. The MR learning system can potentially display one flashcard at a time. Each flashcard should be color-coded based on users' previous attention. Users can go through the flashcards one by one to review their learning progress and the learning material at their own pace. Fig. 2 shows this idea from the worksheets. This design was favored in the group discussion during the participatory design sessions. Participants from the workshop sessions proposed several advantages of this design. First, this design reminds users of familiar experiences in using a deck of physical flashcards, thus potentially reducing their cognitive workload. Moreover, participants suggested that the new design has the potential to help users reflect on their learning results and even assist them in fostering knowledge. In conclusion, the design of attention feedback for our system was finalized based on the results of the participatory design study: The feedback scene of our system is designed as a set of color-coded flashcards. Details of the design and implementation of the attention feedback are discussed in the next section.

Fig. 3. Low-fidelity prototype of the feedback scene.

5. Design and development of *AF-mix*

Our system *AF-Mix* integrates **A**ttention **F**eedback in **Mix**ed Reality. *AF-Mix* consists of three scenes: a) introduction; b) exploration; c) feedback. The introduction scene assists users in onboarding and presents essential information to help them understand the overall concept to be learned. The exploration scene allows users to explore the learning material with eye-gaze-based interaction and learn the concept at their own pace, The learning material is presented as content items, including 3D models, images, texts, and videos. Lastly, attention feedback is provided in the feedback scene based on users' previous gaze attention.

We selected the established HCI concept MHP as the learning content, based on the previous study of Liu et al. (2022). The concept is well-established and was initially proposed by Card (1981) to integrate Human Information Processing (HIP) knowledge in HCI education and research. The user requirements suggest an engaging and active learning experience. Users should be able to freely explore the learning content and activate content items based on their attention. Therefore, we arranged the content with multimedia material, including 3D models, videos, images, and texts as content items. Users can trigger the content items and explore the information with their gaze, and *AF-Mix* presents users with attention feedback based on the collected eye tracking data. Section 5.2 describes the user interaction in more detail.

Before implementing *AF-Mix*, we created low-fidelity prototypes and made design decisions. Fig. 3 shows the feedback scene with the design of presenting attention feedback as color-coded flashcards, inspired by the idea that emerged from the participatory design study (See Fig. 2). The final design of attention feedback in our system is the following: After exploring the content in the exploration scene, users would receive a summary of how many content items they have explored or missed. Next, users have the opportunity to review all items displayed as a deck of flashcards in a randomized order and go through them one after another. Each flashcard is also color-coded to indicate the level of attention from the previous scene.

The following sections describe the system architecture and each scene with more details, including how users explore the learning content and how attention feedback is presented in *AF-Mix*.

5.1. System architecture

For implementing *AF-Mix* as a HoloLens 2 application, we used Unity as our engine and imported the foundation package of Mixed Reality Toolkit (MRTK) 2.7.0 in Unity. In Unity, our workflow consists of two parts: a) creating 3D models, content items, and the user interface; b) compiling C# scripts to connect the frontend and backend. The first part was achieved with the open-source Computer-Aided Design (CAD) software Blender[2]. In Blender, five 3D models (human body, eye, ear, neuron cells, and brain) were created to represent the main components of the MHP: a human user, visual perception, auditory perception, motor system, cognition, and memory.

After importing the 3D models in Unity, the frontend and backend of *AF-Mix* were implemented. Prefabs from the MRTK library were used for the frontend interface, including prefabs of information panels that hold images, texts, and videos as content items as well as UI buttons. The backend collects eye tracking data when users interact with the items. Based on the collected data, the system categorizes the content items according to users' gaze duration and presents attention feedback accordingly. Additionally, the collected eye tracking data were also used for quantitative analysis in the evaluation study.

5.2. Introduction and exploration scene

Attempting to avoid usability issues, we include a short introductory scene in the final system. The introduction scene provides interactive examples of how to use eye-gaze-based interaction to explore elements in MR.

In the introduction scene, *AF-Mix* introduces eye-gaze-based interaction to users (See Fig. 4-2, Fig. 4-3). At this stage, users are instructed to toggle a button with their gaze. If users successfully toggle the button by gazing at the button for more than two seconds (default setting for button prefabs in MRTK for HoloLens 2), they can proceed to the next page. On the next page, a 3D model of a cheeseburger will appear on top of the information panel. Here, the introduction guides users to look at the model and activate a tooltip before they proceed to the last page (See Fig. 4-3). This interaction technique is used throughout the exploration scene. The last page of the introduction scene presents basic information on the MHP concept. It is designed to give users a general impression of what they will learn later and reduce their cognitive workload when they enter the exploration scene.

The exploration scene contains learning material users can explore, including 3D models and content items. Upon entering the scene, users encounter the main 3D model of a human user interacting with a laptop. Besides, a notification message is fixed above the 3D model, reminding users to employ the previously introduced eye-gaze-based interaction to explore the content (Fig. 5-1). The self-directed learning process follows a drill-down approach, where users explore content arranged in three layers. Fig. 5 illustrates the process of exploring the three layers of content. The main 3D model is the first layer of content: It is static and displayed throughout the scene. When users gaze at particular parts of the main model, including eyes, ears, the brain, and hands, the second layer will be activated. Each second layer comprises another 3D model (secondary model) and a short text explaining one of the following components: visual perception, auditory perception, motor system, cognition, and memory. When exploring the second layer, users can gaze at the secondary models and activate the third layer. The third layer contains content items that elaborate on sub-parts of the MHP concept in detail. For example, the auditory perceptual processor (ears) is part of the second layer, consisting of three sub-parts: inner, middle, and outer ear. Each sub-part includes two content items: one information panel with texts and another with either images or videos.

Throughout the exploration scene, a total of 20 content items are presented. As detailed in the system architecture, each item is associated with a script that logs whether users have fixated on the item and records the gaze duration for each item. This functionality is crucial for providing attention feedback in the subsequent scene. In the exploration scene, users can find a 'Finish Learning' button located in front of the main hologram, enabling them to proceed to the feedback scene when they decide to conclude their content exploration.

[2] https://www.blender.org/

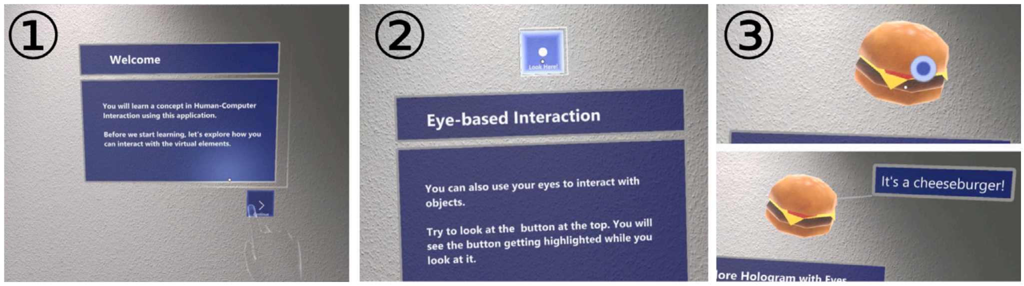

Fig. 4. Guidance on using gesture (1) and eye-gaze-based interaction (2,3) in the introduction scene.

Fig. 5. First, users use eye-gaze-based interaction to explore the main 3D model (1) and activate the second layer of content (2). Next, users gaze at the secondary models (2) to activate and explore the content items (3).

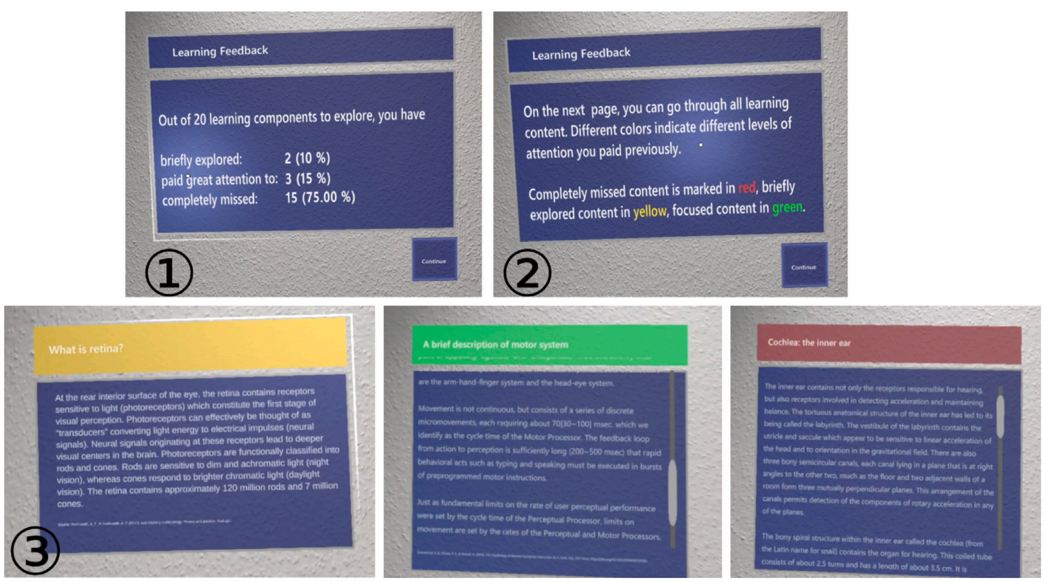

Fig. 6. Attention feedback in our system includes basic learning analytics on users' attention (1) and color-coded flashcards for reviewing the content (2,3).

5.3. Feedback scene

Attention feedback is presented as a deck of color-coded flashcards in the last scene in *AF-Mix*. After finishing the explorative learning in the previous scene, users are presented with a brief analytics of their learning outcome with three numbers (See Fig. 6-1): the number of completely missed items, the number of briefly explored items, and the number of items that have been examined with great attention. The criteria for such three categories are defined based on users' gaze duration. If users' gaze duration on a content item exceeds a predefined threshold, based on reading rate research (Brysbaert, 2019), the component is labeled as "explored with great attention" (color-coded in green). The threshold is derived as follows: On average, each content item contains 85.05 words (SD = 40.19). It is assumed that learners need to read at least the first three lines of text on an item, averaging 27.6 words (SD = 2.1), to comprehend its content. Consequently, the benchmark threshold for a content item to be color-coded green ("explored with great attention") is a gaze duration exceeding 5 s, calculated using a fast reading rate of 337 words per minute for non-fiction English text (Rayner et al., 2010; Brysbaert, 2019). Meanwhile, if the gaze duration falls below this threshold but is greater than zero, the item is labeled as "briefly explored" (color-coded yellow). Items with a gaze duration of zero seconds are marked as "completely missed" (color-coded red). This categorization organizes content items in the feedback scene into three attention levels: high (green), brief (yellow), and none (red).

The analytics is presented in an information panel consisting of the aforementioned three numbers. After reviewing the analytics, the

Fig. 7. A user testing the system.

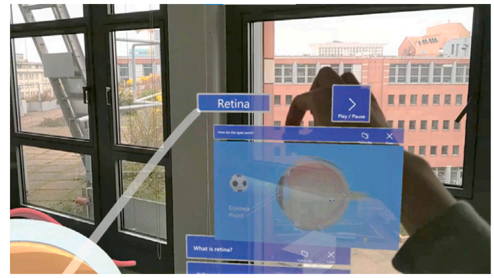

Fig. 8. Screenshot during evaluation.

system presents all content items as a deck of flashcards to users (See Fig. 6-3). The order of the cards is randomized, and the system displays one card at a time. Here, each card is labeled in one of three colors: red, green, and yellow, indicating the three levels of attention. Users can navigate through the cards by selecting either the "back" or "forward" button, arranged on two separate sides of the card. Through presenting the flashcards, *AF-Mix* provides users with individual attention feedback along with a review process. During the review process, users are allowed to check their previous attention to each content item based on the color coding at their own pace. At any time during the process, they can finish the review process by selecting a "Finish Review" button at the bottom of the card deck.

Our design follows the basic idea collected in the participatory design workshop, as most participants were in favor of using flashcards to review the learning content in MR. Aiming to alleviate potential cognitive overload as suggested by the participants, all 3D models from the exploration scene are hidden in the feedback scene. Consequently, only the content items remain, serving as flashcards for attention feedback in the review process.

6. Evaluation study

6.1. Data collection

The evaluation of *AF-Mix* was conducted as a between-subject laboratory experiment with students from the local university in a co-learning space. The learning space had been previously designed for immersive learning experiences and self-study at the university. The evaluation received approval from the university IRB. In the evaluation study, 22 participants were recruited. They were evenly divided into a control group and a treatment group, each comprising 11 participants. Participants who completed the study were incentivized with a monetary remuneration of 12 euros.

The entire evaluation consists of 22 sessions, including 11 sessions for the treatment group and 11 for the control group. Each session was designed to last approximately 60 min. In each session, one of the 22 participants was randomly selected and invited to immerse themselves in the MR learning environment with HoloLens 2. The integration of the eye tracking sensor on HoloLens 2 facilitated the collection of eye tracking data to track users' attention allocation with the *AF-Mix* system. This data collection process involved tracking participants' gaze focus on each content item, determining whether an item was examined in both the exploration and feedback scenes, and capturing the gaze duration on each reviewed item in both scenes (see Figs. 7 and 8).

Additionally, qualitative data were collected through semi-structured interviews conducted afterwards. The goal of the interview is to collect feedback as qualitative data for a thematic analysis. In the coding process for thematic analysis, two coders independently reviewed and coded the interview transcripts to improve inter-coder reliability and validity. An achieved Cohen's kappa of 0.82 suggests a good agreement level through the coding process.

6.2. Task and procedure

The primary objective of this study is to investigate the efficacy of providing attention feedback for MR learning environments. Thus, we introduced two separate review options within the MR learning app: one integrated with attention feedback, as described in Section 5 (assigned to the treatment group), and another without such feedback (assigned to the control group). The introduction and exploration scenes for the two options are identical, only the feedback scene has been modified for the control group. Participants in the control group received flashcards without color coding or analytics, unlike the treatment group, who received attention feedback along with insights into their learning outcomes, as depicted in Fig. 6. This deliberate design meant that participants in the control group were left uninformed about which specific items they may have missed during the initial exploration. Consequently, they had to rely on their memory, while this information was presented explicitly in the feedback scene to the treatment group.

With these design considerations, we organized 22 sessions. Each session began with an onboarding process, presenting participants with study background information, explaining the experimental procedure, and clarifying their rights. After giving informed consent, participants were instructed to use the HoloLens 2 and calibrate the device for eye tracking. This concludes the onboarding process.

After the onboarding process, each participant in both groups was given the same task: They should explore the learning content freely without time constraints in the exploration scene and review their learning progress by going through the flashcards in the feedback scene with *AF-Mix*. During the process, eye tracking data were collected as described in the previous Section 6.1. Upon completing the learning task, each participant was invited to a semi-structured interview. The interviews were designed to take around 10-15 min to complete, addressing the following topics: a) overall impression; b) potential usability challenges; c) opinions on the attention feedback; and d) suggestions for further improvement. The interviews were recorded with the consent of the participants. After the interviews, participants completed the evaluation study and claimed the remuneration.

6.3. Participants

We recruited 22 participants (8 females). The age of the participants ranged from 19–28 years old (M = 24.09, SD = 2.69). The recruitment process was done through the online recruitment platform of the Karlsruhe Design and Decision Lab (KD2Lab) at Karlsruhe Institute of Technology[3]. The KD2Lab has a standard subject pool comprising university students from diverse educational backgrounds and with different academic degrees. The inclusion and exclusion criteria remain

[3] More information on the recruitment platform can be found on the website of the KD2Lab: https://www.kd2lab.kit.edu/

the same as for the design workshops. Through this process, we managed to recruit participants that represent the target user demographic for *AF-Mix*.

During the onboarding process, we inquired about participants' previous experience with immersive technologies and their familiarity with the presented topic. None of the participants had prior experience with HoloLens 2, and none were familiar with the MHP concept presented in *AF-Mix*.

7. Data analysis and results

7.1. Eye tracking data analysis

We collected eye tracking data to assess gaze duration on each content item throughout the evaluation. This allowed for the measurement of the variables NewFeedbackItems (NFI) and TimeNewFeedbackItems (TNFI). NFI represents the total number of content items that were not initially viewed in the exploration scene but were later **discovered and examined** in the feedback scene. TNFI indicates the total gaze duration, measured in seconds, spent examining the newly **discovered and examined** content items in the feedback scene. Note that, to be classified as 'discovered and examined,' a content item must accumulate a gaze duration exceeding 5 s in the feedback scene. We set this minimum threshold based on the assumption that participants, with a fast reading rate of 337 words per minute, would likely need to read at least three lines of text presented on content items (with an average length of 27.6 words) to grasp an overview of the content. We use prior research from Rayner et al. (2010) as a reference for the reading speed, while the average text length is calculated from the presented content items in our design. This approach aims to provide a thoughtful consideration of participants' interactions with the presented content.

Considering our between-subject design and data collection method, we opted for the Bayesian t-test framework to analyze the gathered data, as described by Rouder et al. (2009). Our research goal is to explore the impact of introduced attention feedback in the MR learning environment, and as such, we formulated relevant hypotheses. Given that the attention feedback offers learners a comprehensive overview of their prior learning outcomes, providing a clear indication of what they missed in the exploration phase through color coding, we outline the rival hypotheses regarding the NFI variable as follows: The null hypothesis ($H_0\_NFI : \delta = 0$) assumes that attention feedback does not affect the number of newly discovered and examined items in the feedback scene. In contrast, the one-sided alternative hypothesis ($H_+\_NFI : \delta = 0$) suggests that providing attention feedback results in a higher number of such items for participants in the treatment group. The δ value represents the standardized effect size.

For the TNFI variable, we propose the following rival hypotheses: The null hypothesis ($H_0\_TNFI : \delta = 0$) proposes that attention feedback does not affect the total gaze duration on newly discovered and examined items in the feedback scene. In contrast, the one-sided alternative hypothesis ($H_+\_TNFI : \delta > 0$) suggests that attention feedback leads to a longer gaze duration due to the increased amount of information (i.e., the color coding on each flashcard). Due to limited prior knowledge of the research topic, a Cauchy prior distribution ($r = 1/\sqrt{2}$, truncated at zero) has been assigned for both hypotheses. This choice ensures a conservative and unbiased approach to the Bayesian analysis.

Notably, before delving into the examination of the two variables and associated hypotheses mentioned earlier, we investigated the number of items examined during the exploration scene in a pre-analysis. The results from a t-test ($BF_{10} = 0.481$, $BF_{01}z = 2.079$, $\delta = -0.228$, 95% CI: [−0.987, 0.457]) reveal no significant difference between the two groups. This supplementary pre-analysis is crucial as it helps verify that any potential disparities observed in the feedback scene (e.g., differences concerning NFI and TNFI) are likely attributed to the design intervention rather than variations in participants' learning behavior between the two groups.

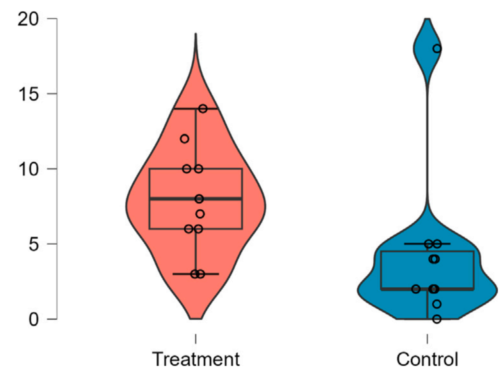

Fig. 9. Number of newly discovered and examined content items in review phase (NFI) for the treatment group (M = 7.909, SD = 3.448) and the control group (M = 4.091, SD = 4.888).

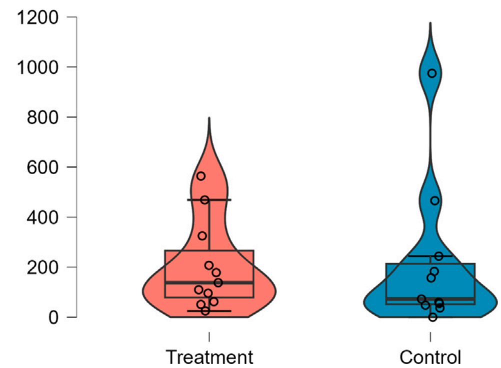

Fig. 10. Total gaze duration (in seconds) on newly discovered and examined items in the review phase (TNFI) for the treatment group (M = 202.273, SD = 177.648) and the control group (M = 208.909, SD = 286.533).

Regarding the NFI variable, the mean item count was 7.909 for the treatment group and 4.091 for the control group, with standard deviations of 3.448 and 4.888 respectively. Therefore, the higher value of NFI in the treatment suggests the efficacy of attention feedback in helping learners retrieve previously missed information. Regarding the TNFI variable, the mean gaze duration was 202.273 s for the treatment group and 208.909 s for the control group, accompanied by standard deviations of 177.648 and 286.533, respectively. Upon plotting the data (See Figs. 9 and 10), we observed non-normality in both variables, accompanied by the presence of outliers. These observations may be attributed to the relatively low sample size and considerable individual differences in learning strategies (Schunk, 2012). To address model misspecification arising from non-normality and outliers, we implemented a log transformation for both variables. Additionally, a Mann–Whitney U test was conducted alongside the two-sample t-test, as it remains unaffected by the log transformation. This dual-method approach has been applied in existing studies to ensure the robustness of a Bayesian t-test analysis (van Doorn et al., 2020). For a comprehensive overview of the collected data, descriptive analyses, and the Bayesian analysis results, please refer to the accompanying .jasp file, accessible at https://osf.io/5t7vf/ (see Figs. 11 and 12).

The findings reveal strong evidence supporting H_+_NFI as the Bayes factor (BF_{+0}) is 11.757, indicative of a great likelihood for the alternative hypothesis. Consistently, the Mann–Whitney U analysis yields a BF_{+0} of 5.171, providing moderate evidence in favor of the alternative hypothesis. This finding underscores the effectiveness of attention feedback in helping learners retrieve previously missed information. On the contrary, for the gaze duration on newly examined items in the feedback scene, no decisive evidence is shown for H_+_TNFI, with $BF_{+0} = 0.754$ and $BF_{0+} = 1.343$. The error percentage in both

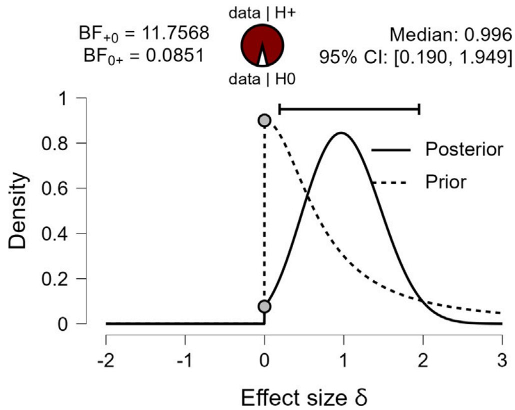

Fig. 11. Hypothesis testing with one-sided Bayesian independent samples t-test for $H_{+}\_NFI : \delta > 0$.

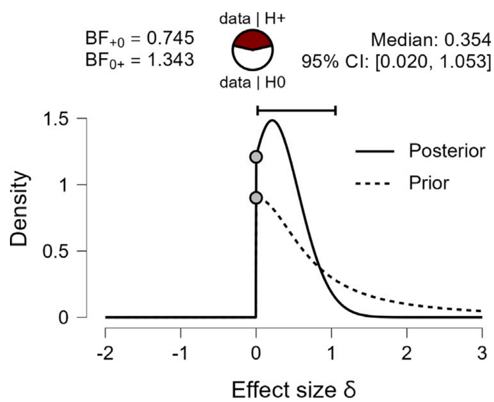

Fig. 12. Hypothesis testing with one-sided Bayesian independent samples t-test for $H_{+}\_TNFI : \delta > 0$.

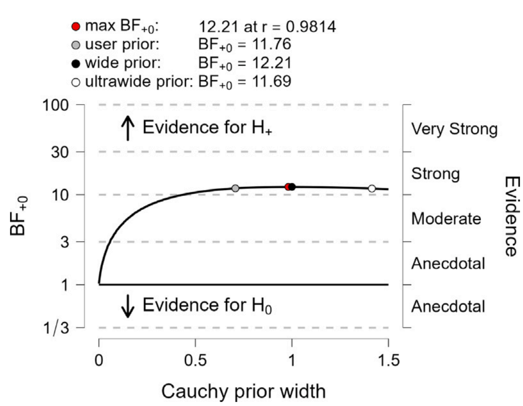

Fig. 13. The Bayes factor robustness plot for BF_{+0} concerning NFI. The maximum BF_{+0} = 12.21 was reached at a prior width r of 0.98. The evidence for the alternative hypothesis is stable throughout different prior distributions.

cases is < 0.001%. Further exploration and analysis may be needed to better understand the impact of attention feedback on the time spent examining newly discovered items.

To ensure the robustness of our findings, we conducted additional analyses. Figs. 13 and 14 depict the Bayes factors concerning the prior width (r). For BF_{+0} regarding variable NFI, the Bayes factor exhibits great stability, ranging from 3 to 12 across a broad range of prior widths from 0.2 to 1.5. In the case of the second pair of hypotheses, the BF_{+0} value attained its peak at the outset (1.104 at r = 0.083) before gradually declining, eventually leaning towards evidence for the null alternative. Lastly, we investigate the parameter estimation for the first hypotheses pair on variable NFI, as the previous Bayesian t-test indicates strong evidence for H_{+}_NFI. By conducting a two-sided t-test, we observe a posterior distribution of δ = 0.99 with a 95% CI ranging from 0.128 to 1.946.

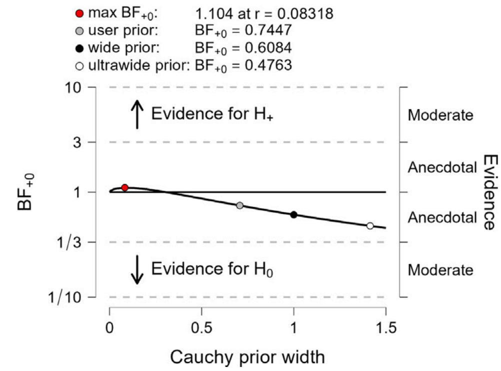

Fig. 14. The Bayes factor robustness plot for BF_{+0} concerning TNIF. The evidence for both the null and alternative hypothesis is rather weak across the range of prior distributions, with slightly more evidence for H_0 as the prior width increases.

7.2. Thematic analysis

For the collected qualitative data, we conducted a thematic analysis of the interview transcripts. 18 out of 22 participants provided consent for interview recordings, which were later transcribed for coding. The remaining four participants agreed to the interview but chose not to be recorded. Instead, we opted for taking notes during their interviews as an alternative to transcripts. The interviews had an average duration of 10.1 min (*SD* = 4.342), and the average length of each transcript is 1082.1 words (*SD* = 559.49) (see Fig. 15).

After collecting the qualitative data, we conducted a thematic analysis, resulting in a set of themes that represent participants' opinions on *AF-Mix*. In the process of coding and generating themes, we followed the guidelines proposed by Braun and Clarke (2021). In line with guidelines, we embraced the concept of recognizing themes as "stories we tell about our data" (Braun and Clarke, 2021). The inductive coding process, involving two independent coders (the first and second authors, both HCI researchers with experience in qualitative analysis), produced 92 codes. Subsequently, through collaborative efforts, these codes were consolidated into 27 overarching codes. A group discussion led to the identification of 10 sub-themes from the example codes. These sub-themes were then examined against the transcript data for accuracy and validity. Ultimately, four main themes emerged from the thematic analysis.

To ensure the reliability of our analysis, we cross-referenced the identified themes with the original codes from the transcripts. Additionally, Cohen's Kappa coefficient was calculated during the thematic analysis, with a result of 0.82 for the coding process and 0.88 for theme generation, indicating a satisfactory level of agreement between coders.

The resulting themes from our qualitative analysis are described in the following section. Quotes are presented with participant information enclosed in brackets, denoting their respective groups: C for the control group and *T* for the treatment group. For instance, PT1 indicates participant number one, who was in the treatment group.

Theme 1: Learning in MR results in difficulties to manage attention

Among the 18 participants who gave consent to the recording and transcribing of the interviews, 16 participants acknowledged the novelty of using MR for self-directed learning. One participant reported that using HoloLens 2 is "full of 'wow' factors" (PC2). Nonetheless, the learning experience in MR was somewhat overwhelming. As a result,

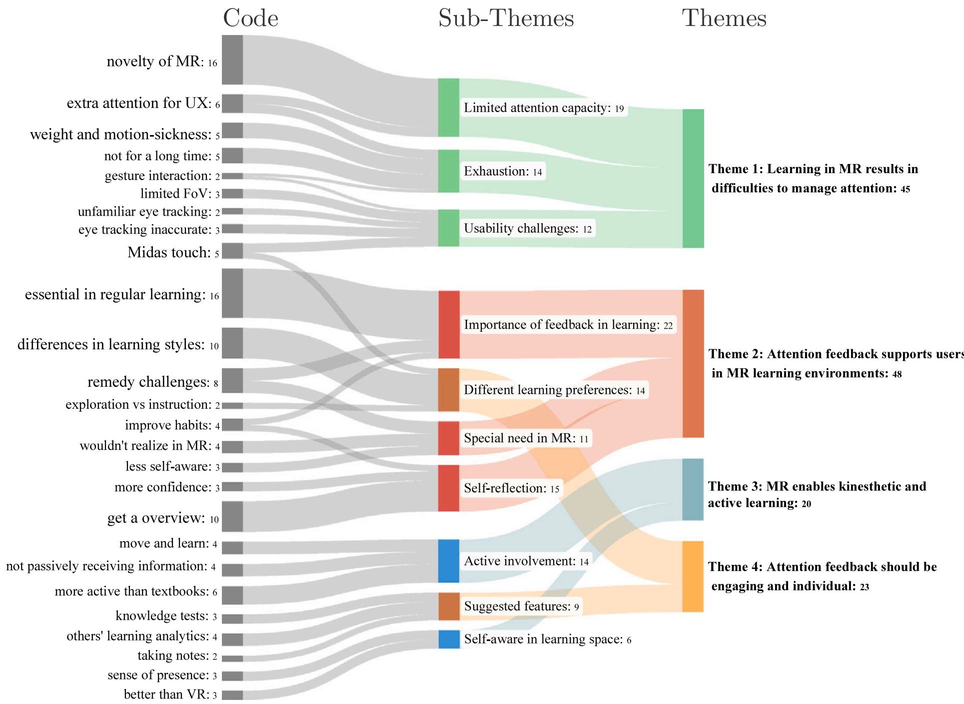

Fig. 15. Visualization of the thematic analysis with a Sankey diagram. The numerical values next to each code suggest its occurrence frequency within the transcripts.

participants could not manage to concentrate on the learning content as intended. Instead, they paid more attention to the proficiency of interaction methods because they "didmn't use anything like this before" (PT17).

Most of our participants (16 out of 22) had not used any MR devices before the evaluation study. For the six participants who had experience with MR, they had only used video-see-through (VST) devices such as HTC Vive. Therefore, all participants were first-time users of HoloLens 2. Common problems identified from the qualitative analysis include usability challenges and motion sickness. Three participants from both the control groups (PC2) and the treatment group (PT17, PT15) complained about the limited field of view (FoV) of the device. Two other participants (PC4, PT7) reported the difficulty of using their fingers to press the buttons in MR environments, while other participants suggested that the eye-gaze-based interaction was not easy to use, because "the gaze indicator is always 1–3 centimeters off from my real gaze position" (PT1). In the interviews, five participants explicitly suggest that the experience tends to "get very exhausting after a while" (PT15) and they "couldn't use it for a long time" (PC8). As a consequence, participants found it tricky to pay full attention to the learning content as they got accustomed to the learning environment.

Theme 2: Attention feedback supports users in MR learning environments

All participants in the treatment group expressed positive opinions on receiving the feedback information. Attention feedback in MR seems to remedy difficulties in the learning process and "optimize the whole learning process" (PT5). Many participants (16 out of 22) admit that receiving feedback is an essential process in their regular learning strategy, so integrating feedback into the learning process in MR made them more confident. Besides, the simple design and interaction method also contributed to supporting users in focusing on reviewing their learning progress. Overall, participants who reviewed their learning progress with attention feedback suggested the following significant advantages of having attention feedback in MR: First, attention feedback helped them find the content they had missed. Reviewing the content with attention feedback helped participants to "**see the big picture**" (PT3). Furthermore, attention feedback supports users in reflecting on their learning journey. The feedback provokes participants to think about "how I (they) have read and learned the content" (PT1) and "improve my habits of learning" (PT5). Eventually, participants reported that *AF-Mix* helped reduce their cognitive workload and improved their confidence after the learning journey. One participant explained this as follows:

> I really liked that part (feedback scene). Because this is my first time using this device (HoloLens 2), I wasn't sure whether I was doing the right thing or whether I have checked out everything. With the feedback, I know what I missed and get a better sense of the topic. (PT9)

In the control group, participants also suggested the need for such feedback. They recognized the learning support provided in their group by the regular flashcards (not color-coded) because they "gave me (them) the opportunity to review everything" (PC18). However, five participants admitted that they needed more support in reviewing their learning process. Using regular flashcards presented in the control group, some participants (PC2, PC4, PC6, PC18) still indicated that "I (they) wouldn't have even realized I've missed something" (PC2). To those participants, we proposed the design of adding color coding based on eye tracking data to the flashcards. All of them agreed that such a design would support them even better. One participant claimed that this design can "help you make sure that you really learn the whole thing and not just some information" (PC6).

In both groups, even participants who do not integrate feedback as a necessary part of their learning routine also appreciated having attention feedback in MR. Participants in our evaluation have reported different attitudes towards feedback for their regular learning experience (e.g., exam preparation). Some participants revealed that "feedback is not part of my learning strategy at all" (PC14), while others claimed that it is a crucial part of exam preparation for them (PT1, PC6, PC16). One participant justified her neglect of feedback as follows: "When I study for exams, instead of finding a way to get feedback, I might as well just keep studying" (PC14). Nonetheless, we found that attention feedback in MR environments is desired by all participants. Even for participants who normally do not value feedback, attention feedback is also desired in MR for two reasons. First, some participants explained that receiving attention feedback, at last, helped them build confidence, as a new learning environment in MR makes them (PC2 and PT9) wonder: “Have I really learned everything?” Secondly, participants expressed a desire to “know more about myself” (PT7), and receiving attention feedback fulfills such a need.

Theme 3: MR enables kinesthetic and active learning

Despite the usability challenges faced in MR learning environments, participants still acknowledged the unique advantages of the learning experience. Four participants from different groups (PT1, PC2, PT5, PT13) pointed out that users need to “move and learn” in MR environments. Even though a few participants suggested that it can be “demanding” for specific learning contexts, others assumed such a learning method may help users retrieve information and foster knowledge by nudging them to create a connection between spatial contexts and the learning content. One participant gave a concrete example of writing an exam:

> I noticed different information is arranged at different places in the space. So you can imagine when you are taking a quiz (about the MHP concept), you may still remember where the ear was (in the MR environment). This can help you recall relevant information. (PT5)

In general, participants needed to actively explore the learning content in MR by moving in the room and interacting with different elements, instead of passively receiving information as students would do when they “learn something from a textbook” (PT5). Previous research also suggests the advantage of motivating kinesthetic learning in MR environments (Knierim et al., 2020), and our thematic analysis supports this assumption. Apart from comparing our system with traditional learning formats such as textbooks, several participants (PC2, PC6, PC16) also compared it with the learning experience using VR-HMDs:

> Especially when I’m in this room (learning space at the university), I prefer using HoloLens instead of VR, (because) when I see the surrounding environment, I realize this is a space for learning and I should concentrate. In VR, I’m totally immersed in a virtual space (PC6).

Theme 4: Attention feedback should be an engaging and individual

Participants in our study exhibited different learning skills and strategies. For some participants, exploring the learning content actively in the exploration scene is an engaging experience. For others, self-directed learning as well as content exploration can be confusing, and they would like to “see something like a table of content(s)” and learn everything “with instructions and perhaps with someone else” (PT15). Such comments reflect individual differences in learning styles and strategies. Therefore, diverse and sometimes contradicting suggestions for improvements were proposed by participants.

First, participants reported a contrast between the exploration scene and the feedback scene. The exploration scene requires users to actively activate content items with eye-gaze-based interaction and move around the physical space, while the feedback scene presents all items one by one to users. While some participants recognized this as an effort to reduce cognitive workload in the feedback scene, others found it less engaging than the exploration process. Another emerging issue is the so-called Midas touch problem (Jacob, 1995; Velichkovsky et al., 1997). Since the intents behind visual attention remain unclear to our gaze-aware system, participants sometimes find the feedback frustrating. One participant elaborated on this during the interview as follows:

> I already knew everything about certain parts, so I skipped it. However, in the end, the application suggested I need to pay more attention. On the other hand, I looked at one part because I was confused, and it was hard to understand; but in the end, I was told I already fully explored it. This inaccuracy demotivates me a little bit. (PC2)

Participants also proposed potential solutions and attempted to make the process of receiving attention feedback more engaging. Four participants (PC4, PT7, PT15, PC16) suggested that integrating knowledge tests in MR can be supplementary to the current design. The knowledge tests can also contribute to determining the reason behind visual attention. For example, if a user could answer all questions regarding a certain item correctly but had not paid attention to it during the learning process, *AF-Mix* could assume that users had skipped it because of their familiarity with the content. Another idea to make the feedback and review process more engaging is presenting statistics of fellow users. This idea was elaborated by one participant as follows:

> I don’t mean that I care about other people who are studying the same thing, but seeing the learning analytics of everybody else who is learning the same thing can really motivate me in reflecting on my own learning behavior. (PT1)

8. Discussion

In general, participants of the evaluation study reviewed our system as an effective approach to helping them better allocate their attention during the learning journey. The core functionality of presenting attention feedback with color-coded flashcards received positive feedback in the qualitative analysis. The analysis of the eye tracking data also suggests that providing attention feedback helps users discover previously overlooked content more efficiently. Notably, this did not require additional time within the feedback scene on the newly discovered items. This observation suggests that the implemented design features may not have induced cognitive overload, preventing a scenario where participants would require additional time to process supplementary information, such as color coding. This confirms that our designed learning system is aligned with what users suggested in the participatory study.

Overall, the eye tracking data seem to suggest the following pattern introduced by the attention feedback: While all participants exhibited similar learning progress and outcomes during the exploration scene (validated by the pre-analysis), and they all had the option to examine all 20 content items in the feedback scene, participants in the treatment group demonstrated in the feedback scene a different review strategy. They effectively directed more attention towards previously overlooked items, resulting in a higher count of NFI (it is worth noting that a minimum gaze duration of 5 s is required for an item to be counted as NFI). Simultaneously, aided by color coding, they examined the items they had missed more efficiently in the review process, spending less time on each item than the control group. This resulted in a TNFI value comparable to that of the control group. This nuanced shift in attention allocation and review strategy highlights the potential impact of the introduced design intervention on participants’ interaction with the content in the MR learning environment.

8.1. Design implications

Meanwhile, the analysis of qualitative data reveals diverse opinions among users regarding attention feedback. In summary, we outline the following design implications:

Facilitating Effective Learning Through Attention Feedback

Exploring learning content in MR was reported by participants to be an engaging, innovative, and unique experience, and it promotes active learning. Nonetheless, they often felt overwhelmed, as usability challenges (e.g., limited FoV, inaccurate calibration) disrupted their attention in the learning process. These challenges may affect users’ autonomy and self-confidence in using MR systems. Concerning the learning support presented in our system, participants in both groups reported that the flashcards had reduced their cognitive workload after the exploration, as presenting flashcards in the feedback scene allowed participants to see all learning content and reflect on their learning progress. Under these conditions, participants in the treatment group emphasized that the gaze-aware attention feedback allowed them to reflect on their learning journey and allocate their attention more effectively during the review process. Consequently, attention feedback in MR helped them "see the big picture" (PT3). In addition, participants in the control group suggested that adding such gaze-aware support could have supported them further in the review process. Therefore, when comparing the results of the two groups, we consider that the color-coded flashcards with attention feedback provided participants with more effective learning support.

Even participants who did not typically integrate feedback into their regular learning routine considered receiving attention feedback in MR environments as necessary and beneficial. We attempt to interpret this opinion as follows: When participants explore content presented in traditional learning formats (e.g., textbooks), they have an established strategy to approach the content from their accumulated experience. Having much experience in efficiently using the given formats, feedback may not be deemed as quintessential for learning success by some participants. However, if participants are exposed to a new learning environment such as MR, with no prior experience and success, they may need more learning support such as attention feedback. Further research can explore the hypothesis that there is a correlation between the level of prior experience in MR and the need for attention feedback as a form of learning support. Lastly, our results suggest that attention feedback in MR should be designed compactly, assisting users in reducing cognitive overload and managing their attention.

Tailoring Attention Feedback to Individual Learning Styles

While participants found *AF-Mix* to be a practical support mechanism in using the MR learning system, our interpretation of the evaluation results revealed variations in participants' learning styles and preferences concerning attention feedback. Due to this variation, alternative designs for attention feedback were proposed. These include two ideas: First, integrating learning analytics of fellow students to promote extrinsic motivation. Second, making attention feedback more interactive with a quiz based on users' previous attention.

The first alternative design may benefit users who prefer peer learning over self-directed learning environments, information on the learning progress of others may motivate them to review their own learning outcomes. In contrast, the second alternative design might appeal to users favoring inquiry-based learning, which may better provoke active learning in the feedback process.

Having alternative design ideas proposed by participants does not necessarily suggest the inferiority of our current design. Rather, these alternative ideas illustrate the diversity of learning styles among users. Both our findings of the evaluation study and existing studies suggest that people embrace different learning strategies (Schunk, 2012). As a result, it is impossible to suggest a universal method to support learning in MR with feedback. Besides, our relatively small sample size and predefined learning topic cannot represent all learning scenarios in MR. We argued against a one-size-fits-all design of attention feedback. Our work aims to showcase the potential of attention feedback for MR learning environments and propose one effective design. More research is needed to understand how different designs of attention feedback in MR can benefit learners with different mindsets and learning styles.

8.2. Limitation and future work

Currently, *AF-Mix* only incorporates one method of providing feedback: attention feedback based on eye tracking. This was inspired by the identified research gap in providing gaze-aware attention support in MR learning systems. However, MR learning environments may benefit from other types of feedback. For instance, giving feedback on whether a physical action is performed correctly is crucial for kinesthetic and spatial learning. Existing research integrates such feedback in MR by using external sensors to collect behavioral data and compare it to a predefined baseline (Shao et al., 2020). Additionally, motivating kinesthetic and spatial learning is a major benefit of MR environments, as recognized by our participants. However, participants mentioned that the current design does not effectively encourage kinesthetic learning during the feedback and review process. Therefore, future work should explore additional methods to enhance *AF-Mix* with feedback for promoting kinesthetic learning.

Furthermore, understanding the intent behind users' attention is missing in our system, which is commonly known as the Midas touch problem in eye tracking research (Jacob, 1995). Participants in our study reported a lack of extrinsic motivation, when *AF-Mix* constantly assumes the missing attention as a result of poor attention management. In reality, users may ignore certain content because they have already acquired the relevant knowledge before. Meanwhile, paying excessive attention to certain content does not always lead to a better understanding of the content. Instead, our qualitative analysis shows that it can be correlated with confusion or frustration. Therefore, tracing users' attention and inferring user intent based solely on eye tracking data may be insufficient. Future work could explore ways to interpret individual intent behind attention, possibly incorporating additional biosignal data for more accurate attention feedback presentation. Another limitation in understanding users' attention is the predefined threshold for the high attention level. While this threshold is based on the average reading rate for non-fictional English text, it may not apply to all users or learning contexts. To enhance the reliability and generalizability of attention feedback, future research could implement an adaptive threshold that adjusts itself based on the difficulty of the learning content and learners' individual reading speed.

Moreover, future research can enrich the Bayesian analysis with more data and longitudinal studies, shedding light on nuanced aspects of the introduced attention feedback. Specifically, it remains unclear whether the integration of attention feedback, including gaze-aware learning analytics and color-coded flashcards, requires learners to adapt their learning or review strategies. Presently, there is no strong evidence suggesting that attention feedback demands an increase in review time for previously missed items (see Section 7.1 for rival hypotheses H_0_TNFI and H_+_TNFI), a positive indication that the system does not introduce unforeseen challenges in attention management when presenting gaze-aware assistance, resulting in participants' feeling overwhelmed and needing more time. However, the absence of a significant increase in review time may also imply that the feedback does not always stimulate active self-reflection, a process that typically consumes more time. Due to the relatively low sample size for our evaluation, no conclusive interpretation can be made to clarify the results. Collecting data on cognitive load through a survey including the NASA-TLX questionnaire may also provide valuable insights into the mental processes involved. Exploring these nuances can enhance our understanding of the cognitive processes influenced by the feedback mechanism.

Lastly, our study aims to enhance the MR learning experience through attention feedback. We did not compare the advantages of MR learning systems with traditional formats, focusing instead on refining the current MR learning experience with an existing system. We somewhat presume the benefits and advantages of MR systems over traditional learning formats as inherent in some aspects, as various advantages have been highlighted in prior research (Mohammadhossein et al., 2022; Akçayır and Akçayır, 2017). Nonetheless, to what extent such advantages hold true and whether they apply to all contexts (including HCI education in our study), require further research. We acknowledge the need for future work to validate assumed benefits and explore the specific advantages of MR learning systems over other learning formats, particularly in under-researched areas such as HCI education. Additionally, participants in the design workshop may bring legacy biases influenced by their familiarity with existing analog technologies. While this can lead to design solutions that feel intuitive and easy to use for novice users, such as the flashcard approach in our system, it may also limit the exploration of the full innovative potential offered by immersive learning technologies.

9. Conclusion

In this paper, we present *AF-Mix*, a gaze-aware learning system that provides attention feedback in MR. The motivation for integrating attention feedback in MR is to assist users in managing their attention in the learning experience. We identified several challenges users typically encounter in MR learning environments, including attention management. These challenges highlight the growing need for effective support in the learning experience. Therefore, our goal is to provide learning support by assisting users in managing their attention and reviewing their learning outcomes with attention feedback in self-directed MR environments.

By adopting a human-centered design approach to develop *AF-Mix*, we first conducted a participatory design study. Through participatory design workshops, we adopted the idea of visualizing the learning content as color-coded flashcards based on users' attention. We delivered the final design as follows: *AF-Mix* first allows users to actively explore the content using eye-gaze-based interaction while collecting eye tracking data during the process. Next, *AF-Mix* notifies users of their attention distribution based on the collected data, including the number of items users have previously missed, briefly explored, and explored with great attention during the content exploration. Subsequently, *AF-Mix* displays all learning content as a deck of flashcards. Here, each flashcard is color-coded based on users' previous attention levels, allowing users to review their learning outcomes.

The results of the mixed-method evaluation show that *AF-Mix* assisted participants in figuring out what content items they had missed during the content exploration and motivated them to examine these items in the review process. Eventually, participants described *AF-Mix* as efficient in guiding them to retrieve missed information and reflect on their learning progress. Thereby, they could allocate their attention more efficiently in the review process.

In summary, our work contributes to the field of HCI by highlighting the challenges associated with attention management in MR learning environments and providing effective gaze-aware support through attention feedback. Our findings suggest that attention feedback serves as an effective gaze-aware support method to improve the overall learning experience in MR. This research also offers a foundation for further exploration: The design methods and solutions proposed in our paper can also be applied to other learning contexts in MR. Future research can extend the design knowledge with evaluation in different learning scenarios, contributing further implications and innovative methods for enhancing gaze-aware support in MR learning environments. We believe that this study can contribute to unleashing the full potential of MR and provide valuable insights for ongoing research in designing human-centered MR learning systems.

CRediT authorship contribution statement

Shi Liu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Peyman Toreini:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Formal analysis, Conceptualization. **Alexander Maedche:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge funding by the Federal Ministry of Education and Research of Germany (BMBF) and the Ministry of Science, Research and the Arts (MWK) of Baden-Württemberg (Grant No.: 16DHBKI004).

Data availability

Data will be made available on request.

References

Akçayır, M., Akçayır, G., 2017. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. Educ. Res. Rev. 20, 1–11. http://dx.doi.org/10.1016/j.edurev.2016.11.002.

Arici, F., Yildirim, P., Caliklar, c., Yilmaz, R.M., 2019. Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. Comput. Educ. 142, N.PAG–N.PAG.

Bianchi, I., Zanatta, A.L., Rieder, R., 2020. Augmented reality in medical teaching-learning process content: A systematic review. In: 2020 22nd Symposium on Virtual and Augmented Reality (SVR). IEEE, Porto de Galinhas, Brazil, pp. 129–133. http://dx.doi.org/10.1109/SVR51698.2020.00032.

Boy, G.A., 1997. The group elicitation method for participatory design and usability testing. Interactions 4 (2), 27–33. http://dx.doi.org/10.1145/245129.245132, URL https://dl.acm.org/doi/10.1145/245129.245132.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qual. Res. Psychol. 3 (2), 77–101. http://dx.doi.org/10.1191/1478088706qp063oa, URL http://dx.doi.org/10.1191/1478088706qp063oa.

Braun, V., Clarke, V., 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual. Res. Psychol. 18 (3), 328–352. http://dx.doi.org/10.1080/14780887.2020.1769238.

Brysbaert, M., 2019. How many words do we read per minute? A review and meta-analysis of reading rate. J. Mem. Lang. 109, 104047. http://dx.doi.org/10.1016/j.jml.2019.104047.

Butler, D.L., Winne, P.H., 1995. Feedback and self-regulated learning: A theoretical synthesis. Rev. Educ. Res. 65 (3), 245–281, URL http://www.jstor.org/stable/1170684.

Card, S.K., 1981. The Model Human Processor: A Model for Making Engineering Calculations of Human Performance. Proc. Hum. Factors Soc. Annu. Meet. 25 (1), 301–305. http://dx.doi.org/10.1177/107118138102500180.

D'Mello, S., Olney, A., Williams, C., Hays, P., 2012. Gaze tutor: A gaze-reactive intelligent tutoring system. Int. J. Hum.-Comput. Stud. 70 (5), 377–398. http://dx.doi.org/10.1016/j.ijhcs.2012.01.004.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N.J., Gronau, Q.F., Haaf, J.M., Hinne, M., Kucharský, v., Ly, A., Marsman, M., Matzke, D., Gupta, A.R.K.N., Sarafoglou, A., Stefan, A., Voelkel, J.G., Wagenmakers, E.-J., 2020. The JASP guidelines for conducting and reporting a Bayesian analysis. Psychon. Bull. Rev. 28 (3), 813–826. http://dx.doi.org/10.3758/s13423-020-01798-5.

Giraudeau, P., Olry, A., Roo, J.S., Fleck, S., Bertolo, D., Vivian, R., Hachet, M., 2019. CARDS: A mixed-reality system for collaborative learning at school. In: Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces. Association for Computing Machinery, New York, NY, USA, pp. 55–64. http://dx.doi.org/10.1145/3343055.3359721.

Grogorick, S., Magnor, M., 2020. Subtle Visual Attention Guidance in VR. In: Magnor, M., Sorkine-Hornung, A. (Eds.), Real VR – Immersive Digital Reality: how To Import the Real World Into Head-Mounted Immersive Displays. Springer International Publishing, Cham, pp. 272–284. http://dx.doi.org/10.1007/978-3-030-41816-8_11.

Han, Y., Miao, Y., Lu, J., Guo, M., Xiao, Y., 2022. Exploring intervention strategies for distracted students in VR classrooms. In: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems. In: CHI EA '22, Association for Computing Machinery, New York, NY, USA, pp. 1–7. http://dx.doi.org/10.1145/3491101.3519627.

Hanid, M.F.A., Said, M.N.H.M., Yahaya, N., 2020. Learning Strategies Using Augmented Reality Technology in Education: Meta-Analysis. Univers. J. Educ. Res. 8 (5A), 51–56. http://dx.doi.org/10.13189/ujer.2020.081908.

Hennerley, J., Dickinson, M., Jiang, M., 2017. Augmented reality enhanced human robot interaction for social robots as E-learning companions. In: Proceedings of the 31st British Computer Society Human Computer Interaction Conference. In: HCI '17, BCS Learning & Development Ltd., Swindon, GBR, pp. 98:1–6. http://dx.doi.org/10.14236/ewic/HCI2017.98.

Heslin, P.A., 2009. Better than brainstorming? Potential contextual boundary conditions to brainwriting for idea generation in organizations. J. Occup. Organ. Psychol. 82 (1), 129–145. http://dx.doi.org/10.1348/096317908X285642, Place: United Kingdom Publisher: British Psychological Society.

Huang, Y.-H., Chang, H.-Y., Yang, W.-l., Chiu, Y.-K., Yu, T.-C., Tsai, P.-H., Ouhyoung, M., 2018. CatAR: A novel stereoscopic augmented reality cataract surgery training system with dexterous instruments tracking technology. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–12.

Huang, T.-C., Chen, C.-C., Chou, Y.-W., 2016. Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. Comput. Educ. 96, 72–82. http://dx.doi.org/10.1016/j.compedu.2016.02.008.

Huang, G., Qian, X., Wang, T., Patel, F., Sreeram, M., Cao, Y., Ramani, K., Quinn, A.J., 2021. AdapTutAR: An adaptive tutoring system for machine tasks in augmented reality. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. In: CHI '21, Association for Computing Machinery, New York, NY, USA, pp. 1–15. http://dx.doi.org/10.1145/3411764.3445283.

Hutt, S., Krasich, K., R. Brockmole, J., K. D'Mello, S., 2021. Breaking out of the Lab: Mitigating mind wandering with gaze-based attention-aware technology in classrooms. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. In: CHI '21, Association for Computing Machinery, New York, NY, USA, pp. 1–14. http://dx.doi.org/10.1145/3411764.3445269.

Ibáñez, M.-B., Di-Serio, Á., Villarán-Molina, D., Delgado-Kloos, C., 2016. Support for Augmented Reality Simulation Systems: The Effects of Scaffolding on Learning Outcomes and Behavior Patterns. IEEE Trans. Learn. Technol. 9 (1), 46–56. http://dx.doi.org/10.1109/TLT.2015.2445761.

Iqbal, M.Z., Mangina, E., Campbell, A.G., 2019. Exploring the use of augmented reality in a kinesthetic learning application integrated with an intelligent virtual embodied agent. In: 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). pp. 12–16. http://dx.doi.org/10.1109/ISMAR-Adjunct.2019.00018.

Jacob, R.J.K., 1995. Eye tracking in advanced interface design. In: Virtual Environments and Advanced Interface Design. Oxford University Press, Inc., pp. 258–288.

Khan, T., Johnston, K., Ophoff, J., 2019. The impact of an augmented reality application on learning motivation of students. Adv. Human- Comput. Interact. 2019, 1–14. http://dx.doi.org/10.1155/2019/7208494.

Khan, M., Trujano, F., Choudhury, A., Maes, P., 2018. Mathland: Playful Mathematical Learning in Mixed Reality. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, Montreal QC Canada, pp. 1–4. http://dx.doi.org/10.1145/3170427.3186499.

Kirschner, P.A., Sweller, J., Clark, R.E., 2006. Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educ. Psychol. 41 (2), 75–86. http://dx.doi.org/10.1207/s15326985ep4102_1.

Knierim, P., Kiss, F., Rauh, M., Schmidt, A., 2020. Tangibility is overrated: Comparing learning experiences of physical setups and their virtual equivalent in augmented reality. In: 19th International Conference on Mobile and Ubiquitous Multimedia. In: MUM 2020, Association for Computing Machinery, New York, NY, USA, pp. 299–305. http://dx.doi.org/10.1145/3428361.3428379.

Knierim, P., Kosch, T., Hoppe, M., Schmidt, A., 2018. Challenges and opportunities of mixed reality systems in education. In: Mensch Und Computer 2018 - Workshopband. Gesellschaft für Informatik e.V., Bonn, pp. 325–330.

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S.W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., Tsai, C.-C., 2013. A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educ. Res. Rev. 10, 90–115. http://dx.doi.org/10.1016/j.edurev.2013.10.001.

Liu, S., Toreini, P., Maedche, A., 2022. Designing gaze-aware attention feedback for learning in mixed reality. In: Proceedings of Mensch Und Computer 2022. In: MuC '22, Association for Computing Machinery, New York, NY, USA, pp. 503–508. http://dx.doi.org/10.1145/3543758.3547565.

Maas, M.J., Hughes, J.M., 2020. Virtual, augmented and mixed reality in K–12 education: A review of the literature. Technol. Pedagog. Educ. 29 (2), 231–249. http://dx.doi.org/10.1080/1475939X.2020.1737210.

Mayer, R.E., 2004. Should There Be a Three-Strikes Rule Against Pure Discovery Learning? Am. Psychol. 59 (1), 14–19. http://dx.doi.org/10.1037/0003-066X.59.1.14.

Milgram, P., Fumio, K., 1994. A taxonomy of mixed reality visual displays. 77, IEICE Trans. Inf. Syst. 77, (12).16,

Mohammadhossein, N., Richter, A., Lukosch, S., 2022. Benefits of Using Augmented Reality in Learning Settings: A Systematic Literature Review. ICIS 2022 Proc..

Moro, C., Phelps, C., Redmond, P., Stromberga, Z., 2021. HoloLens and mobile augmented reality in medical and health science education: A randomised controlled trial. Br. J. Educ. Technol. 52 (2), 680–694.

Muller, M.J., Kuhn, S., 1993. Participatory design. Commun. ACM 36 (6), 24–28. http://dx.doi.org/10.1145/153571.255960.

Oh, S., So, H.-J., Gaydos, M., 2018. Hybrid augmented reality for participatory learning: The hidden efficacy of multi-user game-based simulation. IEEE Trans. Learn. Technol. 11 (1), 115–127. http://dx.doi.org/10.1109/TLT.2017.2750673.

Panadero, E., 2017. A review of self-regulated learning: Six models and four directions for research. Front. Psychol. 8.

Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R.G., Kyza, E., Edelson, D., Soloway, E., 2004. A scaffolding design framework for software to support science inquiry. J. Learn. Sci. 13 (3), 337–386, arXiv:1466941.

r, M.A., r, G.A., 2017. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. Educ. Res. Rev. 20, 1–11. http://dx.doi.org/10.1016/j.edurev.2016.11.002, URL https://www.sciencedirect.com/science/article/pii/S1747938X16300616.

Radu, I., Hv, V., Schneider, B., 2021. Unequal impacts of augmented reality on learning and collaboration during robot programming with peers. Proc. ACM Hum.- Comput. Interact. 4 (CSCW3), 245:1–23. http://dx.doi.org/10.1145/3432944.

Radu, I., Schneider, B., 2019. What can we learn from augmented reality (AR)? Benefits and drawbacks of AR for inquiry-based learning of physics. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–12.

Rayner, K., Slattery, T.J., Bélanger, N.N., 2010. Eye movements, the perceptual span, and reading speed. Psychon. Bull. Rev. 17 (6), 834–839. http://dx.doi.org/10.3758/pbr.17.6.834, URL http://dx.doi.org/10.3758/PBR.17.6.834.

Roda, C., Nabeth, T., 2007. Supporting attention in learning environments: Attention support services, and information management. In: Duval, E., Klamma, R., Wolpers, M. (Eds.), Creating New Learning Experiences on a Global Scale. vol. 4753, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 277–291. http://dx.doi.org/10.1007/978-3-540-75195-3_20.

Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G., 2009. Bayesian t tests for accepting and rejecting the null hypothesis. Psychon. Bull. Rev. 16 (2), 225–237. http://dx.doi.org/10.3758/pbr.16.2.225.

Schunk, D.H., 2012. Learning Theories: An Educational Perspective, Sixth ed. Pearson, Boston.

Shao, Q., Sniffen, A., Blanchet, J., Hillis, M.E., Shi, X., Haris, T.K., Liu, J., Lamberton, J., Malzkuhn, M., Quandt, L.C., Mahoney, J., Kraemer, D.J.M., Zhou, X., Balkcom, D., 2020. Teaching American sign language in mixed reality. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 4 (4), 152:1–27. http://dx.doi.org/10.1145/3432211.

Sharma, K., Alavi, H.S., Jermann, P., Dillenbourg, P., 2016. A gaze-based learning analytics model: In-video visual feedback to improve learner's attention in MOOCs. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. In: LAK '16, Association for Computing Machinery, New York, NY, USA, pp. 417–421. http://dx.doi.org/10.1145/2883851.2883902.

Speicher, M., Hall, B.D., Nebeling, M., 2019. What is Mixed Reality? In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM, Glasgow Scotland Uk, pp. 1–15. http://dx.doi.org/10.1145/3290605.3300767.

Thoravi Kumaravel, B., Anderson, F., Fitzmaurice, G., Hartmann, B., Grossman, T., 2019. Loki: Facilitating remote instruction of physical tasks using bi-directional mixed-reality telepresence. In: Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology. ACM, New Orleans LA USA, pp. 161–174. http://dx.doi.org/10.1145/3332165.3347872.

VanGundy, A.B., 1984. Brain writing for new product ideas: An alternative to brainstorming. J. Consum. Mark. 1 (2), 67–74. http://dx.doi.org/10.1108/eb008097.

Vazquez, C.D., Nyati, A.A., Luh, A., Fu, M., Aikawa, T., Maes, P., 2017. Serendipitous language learning in mixed reality. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. ACM, Denver Colorado USA, pp. 2172–2179. http://dx.doi.org/10.1145/3027063.3053098.

Velichkovsky, B., Sprenger, A., Unema, P., 1997. Towards gaze-mediated interaction: Collecting solutions of the "midas touch problem". In: Howard, S., Hammond, J., Lindgaard, G. (Eds.), Human-Computer Interaction INTERACT '97: IFIP TC13 International Conference on Human-Computer Interaction, 14th–18th July 1997, Sydney, Australia. In: IFIP — The International Federation for Information Processing, Springer US, Boston, MA, pp. 509–516. http://dx.doi.org/10.1007/978-0-387-35175-9_77.

Vergel, R.S., Morillo Tena, P., Casas Yrurzum, S., Cruz-Neira, C., 2020. A comparative evaluation of a virtual reality table and a hololens-based augmented reality system for anatomy training. IEEE Trans. Human- Mach. Syst. 50 (4), 337–348. http://dx.doi.org/10.1109/THMS.2020.2984746.

Vovk, A., Wild, F., Guest, W., Kuula, T., 2018. Simulator sickness in augmented reality training using the microsoft hololens. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. In: CHI '18, Association for Computing Machinery, New York, NY, USA, pp. 1–9. http://dx.doi.org/10.1145/3173574.3173783.

Wilson, C., 2013. Brainstorming and Beyond: A User-Centered Design Method, first ed. Morgan Kaufmann, Amsterdam.

Wish-Baratz, S., Gubatina, A.P., Enterline, R., Griswold, M.A., 2019. A new supplement to gross anatomy dissection: HoloAnatomy. Med. Educ. 53 (5), 522–523. http://dx.doi.org/10.1111/medu.13845, URL http://dx.doi.org/10.1111/medu.13845.

Yue, Y.-T., Yang, Y.-L., Ren, G., Wang, W., 2017. SceneCtrl: Mixed reality enhancement via efficient scene editing. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. ACM, Québec City QC Canada, pp. 427–436. http://dx.doi.org/10.1145/3126594.3126601.