FamiliAR Feedback: Investigating Feedback Modality and Familiarity in Classroom Settings Using Spatial Augmented Reality

Nick Wittig
University of Duisburg-Essen
Essen, Germany
nick.wittig@uni-due.de

Yannick Dohmen
University of Duisburg-Essen
Essen, Germany
yannick.dohmen@stud.uni-due.de

Jonathan Liebers
University of Duisburg-Essen
Essen, Germany
jonathan.liebers@uni-due.de

Donald Degraen
University of Canterbury
Christchurch, New Zealand
donald.degraen@canterbury.ac.nz

David Goedicke
University of Duisburg-Essen
Essen, Germany
david.goedicke@uni-due.de

Stefan Schneegass
University of Duisburg-Essen
Essen, Germany
stefan.schneegass@uni-due.de

Figure 1: Left: Staged setup, the projector-camera system that can project feedback onto the paper. Right: The three familiarity levels with image feedback. Images (c–f) are AI-generated for anonymity; in the study, real faces were shown—unfamiliar (c–d) and familiar (e–f, the participant's teacher). Conditions: (a) neutral–correct, (b) neutral–wrong, (c) unfamiliar–correct, (d) unfamiliar–wrong, (e) familiar–correct, (f) familiar–wrong.

Abstract

Spatial Augmented Reality (SAR) can enhance learning experiences through interactive, real-time digital information overlays. Using SAR, content can be projected directly onto physical paper to provide students with in situ task feedback. Our work explores how students perceive different types of SAR feedback. We first identified feedback methods and dimensions from a literature review. We then conducted an expert focus group (N = 5) of professionals who had backgrounds in education and teaching experience. With the focus group, we aimed to expand on the literature review results to identify feedback modalities and dimensions commonly used in classrooms today. Next, we performed a field study (N = 16) with high school students in which we compared the perception of different feedback modalities (*text, image, video*) and familiarity (*neutral, unfamiliar, familiar*) in a classroom setting. Our results revealed that perceived user distraction and novelty are significantly affected by feedback modality through a large effect, with videos being perceived as more distracting and more novel. Familiar, trusted individuals best deliver positive feedback, whereas negative feedback from people should be avoided. We discuss the usage of feedback modalities in various contexts, providing a foundation for future use of SAR feedback in education.

MUM '25, Enna, Italy

ACM ISBN 979-8-4007-2015-4/25/12
https://doi.org/10.1145/3771882.3771883

CCS Concepts

• **Human-centered computing** → *Empirical studies in HCI*; **Field studies**; **Mixed / augmented reality**.

Keywords

spatial augmented reality, corrective feedback, education

ACM Reference Format:
Nick Wittig, Yannick Dohmen, Jonathan Liebers, Donald Degraen, David Goedicke, and Stefan Schneegass. 2025. FamiliAR Feedback: Investigating Feedback Modality and Familiarity in Classroom Settings Using Spatial Augmented Reality. In *24th International Conference on Mobile and Ubiquitous Multimedia (MUM '25), December 01–04, 2025, Enna, Italy*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3771882.3771883

1 Introduction

Augmented Reality (AR) has emerged as a promising tool for enhancing learning experiences and offering engaging and interactive environments. It is a technology that superimposes computer-generated visual information onto the real world, enhancing the user's perception and interaction with their environment [1]. AR devices, such as tablets, smartphones, and head-mounted displays (HMDs), enable in situ overlays that can provide additional context, data, or interactive features [3]. Spatial Augmented Reality (SAR) takes this one step further, projecting digital information directly onto physical surfaces in the environment, thereby eliminating the need for personal devices such as HMDs or screens (e.g., smartphones, tablets, monitors) [21, 23]. This technology uses projector-camera setups to map digital content onto real-world objects, allowing for interactive and immersive experiences that are visible to multiple users simultaneously. One promising application of AR and SAR is in the field of education [54].

The high demands placed on teachers, including the need for pedagogical expertise, individualized attention, and management of increasing class sizes, have led to teacher shortages worldwide [12, 32, 41]. The integration of digital elements into teaching environments has been put forward as a solution to some of these issues. However, improper integration of technology can negatively impact learning outcomes [7]. Therefore, although we know that integrating technology can be beneficial, it is important to understand how this integration can be optimized. One critical aspect of education that can be assisted by technology is feedback. In this work, feedback is defined as evaluation provided for any completed task, which can manifest as either a correction (e.g., indicating whether an answer is correct or incorrect)–referred to as *corrective feedback*–or a hint (e.g., suggesting an alternative approach or providing additional information) [16]. Feedback can be delivered through various mediums, including visual cues, auditory signals, gestures, or facial expressions, and it is typically provided by a qualified individual, such as a teacher or tutor [14, 31]. The perception of feedback is influenced by multiple dimensions, including the modality (e.g., image, text, video, audio), the familiarity with the feedback provider (e.g., whether a familiar teacher or a stranger provides the feedback), and the timing and location of the feedback delivery [43].

Leveraging SAR in educational settings has been shown to have potential in improving learning- and user experiences [49, 51]. However, the selection of feedback modalities in AR learning contexts has been arbitrary and is thus in need of more empirical grounding, especially using representative design [13]. Exploring the dimensions of feedback that impact learning is essential for optimizing the educational use of AR technology. One research need is to explore how familiarity with feedback providers influences how learners receive and respond to feedback. This work investigates these dimensions, mainly focusing on how the familiarity towards the feedback provider and feedback modality affect learners' perception of feedback and learning experiences. While feedback has been widely studied in traditional educational settings, a research gap remains in understanding its use and impact within AR environments. Specifically, there is a lack of systematic investigation into how different feedback modalities function in AR-based education, including their trade-offs and perceived effectiveness from both pedagogical and technological perspectives. Specifically, our research investigates the following research question: (**RQ**) Which spatial augmented reality feedback modality – *text*, *image*, *video* – and which level of familiarity – *neutral*, *unfamiliar*, *familiar* – are educationally appropriate in classroom settings?

Contribution Statement. The contribution of our work is twofold. First, we contribute an empirical investigation of feedback methods in classrooms through an expert focus group with education professionals (N = 5) that presents findings on commonly used feedback types, ideas for SAR feedback scenarios, and ideas to utilize SAR for independent learning. Second, we evaluate these feedback methods in a school-based field study with young students (N = 16) to provide insights on both preferred feedback modalities for students as well as preferred familiarity levels with the feedback provider. Finally, we discuss the broader implications, specifically highlighting the found tradeoffs of the explored feedback types, the familiarity preferences, and guidance on which feedback modality should be used.

2 Related Work

This section provides an overview of both the theoretical and practical aspects of feedback in education. In addition, we examine corrective feedback and conduct a literature review on SAR within educational contexts. The synthesis of these feedback methodologies, together with insights derived from our expert focus group, informs the design of our field study.

2.1 Feedback in Education

Feedback is widely recognized as an essential component for effective learning in educational settings. However, research on assessment feedback often struggles to demonstrate not only what works but also how and why it works. Feedback research is fragmented across various disciplines, which has led to a call for more scientific and cohesive studies on the topic [48].

One issue in feedback research is the feedback gap, which refers to the discrepancy between the potential impact of feedback and how students actually use it. This gap is especially evident in large classrooms and diverse student populations. Iraj et al. show that digital tools can help bridge this gap, but educators often lack a deep understanding of how students engage with feedback [19]. Their study found that students who engaged with feedback earlier in the course were more likely to succeed, which underscores the need for personalized and actionable feedback to improve outcomes. This emphasizes the importance of not only delivering feedback but also ensuring that students interact with and apply it effectively [19].

Feedback interventions aim to improve performance but have produced mixed results. In their meta-analysis, Kluger and DeNisi found that although feedback can enhance performance, it negatively affected more than a third of cases [24]. Their feedback intervention theory proposes that effectiveness depends on whether feedback directs the learner's attention to task learning, task motivation, or self-related processes. When feedback shifts from the task to the self, its effectiveness decreases. This raises questions about how to structure feedback to maintain a task-focused orientation [24].

While feedback has been extensively studied in traditional educational contexts, less is known about how feedback operates in AR systems. Given the increasing use of AR in education, it is important to explore how this technology might affect feedback delivery and reception. Thus, in this work, we explore how feedback in AR is perceived and which feedback dimensions are influential.

2.2 Corrective Feedback and Familiarity

Feedback is widely recognized as an important factor in learning, and its effectiveness depends on several dimensions, including modality and familiarity. Previous work showed that timing and the balance between positive and negative feedback influence student outcomes and reveal the complexity of feedback's role in education [16]. Corrective feedback focuses on task performance and helps distinguish between correct and incorrect answers while guiding students toward additional information. It usually addresses surface-level knowledge and is common in classrooms, where about 90% of teachers' questions target this level. Studies show that corrective feedback has a significant positive impact on learning [16].

Peer feedback, in particular, can be shaped by the trust students place in their peers, which has been shown to enhance engagement and learning [2]. Thus, one aspect of feedback is the relationship or level of familiarity between the feedback provider and receiver. Peer assessment is further complicated by interpersonal and intrapersonal factors such as motivation, trust, and psychological safety [33]. Familiarity has several aspects, like familiarity with regard to an object, task, or system [40], and familiarity with other people and faces [35, 44]. Familiarity plays a role in designing feedback systems, as demonstrated by Saplacan and Herstad's work on feedback for the elderly, suggesting that familiarity with systems improves user comfort [40]. When anonymity in peer feedback is not possible, familiarity among peers influences the honesty and comfort of communication [44]. One part of familiarity is in regards to familiar faces. Different levels of familiarity with faces affect factors like person knowledge, personal relationships, and emotions. These levels were introduced as personally familiar, visually familiar, famous, and experimentally learned faces [35].

Through the review of existing literature, a clear need emerges to investigate the role of feedback in education, with particular attention to its dimensions, such as modality and familiarity. By concentrating on corrective feedback, we aim to identify factors influencing the perception of SAR feedback in classroom environments. Although previous studies have explored feedback modalities, further examination of current feedback practices employed in educational settings could offer valuable insights. We see a research need in the systematic exploration of the tradeoffs of feedback modalities and familiarity.

2.3 (Spatial) AR in Education

We explored commonly occurring feedback modalities in AR. We found the following modalities: *text*, *images and shapes*, *videos and animations*, and *audio*. Table 1 shows several works that utilized AR as a tool for feedback.

Images were used in studies to serve as visual cues in AR and SAR settings. For instance, Freitas and Campos [15] projected images of animals and vehicles using SAR for educational purposes, while Rzayev et al. [38] employed familiar icons to enhance AR notifications. Animations, defined as changes in object properties, were found in multiple studies [37, 45–47, 55]. These animations ranged from guiding assembly tasks [47] to visualizing rehabilitation progress [46]. Text feedback was commonly used in AR [20, 28, 30, 38, 42, 45–47, 53]. Text was often displayed with backgrounds, such as rectangles or speech bubbles, to improve readability. Auditory feedback was employed in five studies [15, 20, 45–47], where information was delivered via speakers or headphones. Shapes were the most frequently used feedback method [8, 25, 28–30, 42, 45, 47, 53, 55]. These included 2D and 3D objects like circles, arrows, and triangles, which were used to convey information. For example, shapes were utilized to highlight joints in exercise routines [30] or to guide users through assembly tasks [29]. Prior studies have systematically investigated integrating audio and visual modalities in XR notification components, demonstrating that context-aware multimodal designs that convey full message content, such as Text-Ping, enhance noticeability, reduce response time, and align with user preferences [10].

Our literature review reveals that (corrective) feedback is a valuable tool in educational settings, with several commonly used methods for AR use-cases already identified. However, the effects of each feedback modality in educational settings remain underexplored. Thus, we conducted an expert focus group with educational professionals aimed at deepening the understanding of feedback usage in contemporary classroom environments.

3 Focus Group

In addition to the findings of our literature review, we conducted a focus group to develop a design space for feedback used in AR applications. The design, participants, procedure, and results are detailed in this section.

Table 1: An overview of the identified feedback types. An "✓" indicates the presence of this type in the source, while a "–" denotes its absence.

Source	Images	Animations	Text	Audio	Shapes
Intelligent Augmented Reality Training [47]	–	✓	✓	✓	✓
SMART [15]	✓	–	–	✓	–
AR-Fit [30]	–	–	✓	–	✓
AR-Maze [20]	✓	–	✓	✓	–
LigthUp [8]	–	–	–	–	✓
Exploring Real-time Precision Feedback [42]	–	–	✓	–	✓
Exploring Tangible Interaction and Diegetic Feedback [28]	–	–	✓	–	✓
Effects of Position and Alignment of Notifications [38]	✓	–	✓	–	–
AR - Enhanced Workouts [55]	–	✓	–	–	✓
Exploring Bi-modal Feedback [45]	–	✓	✓	✓	✓
Augmented Reality for Rehabilitation [46]	–	✓	✓	✓	–
piARno [37]	✓	✓	–	–	–
Evaluating the benefits of real-time feedback [29]	–	–	–	–	✓
Real-time in-situ visual feedback [25]	–	–	–	–	✓
Dropping Hints [50]	✓	✓	✓	–	✓
Designing Textual Information in AR Headsets [53]	–	–	✓	–	✓
Spatial augmented reality: a tool for operator guidance and training evaluated in five industrial case studies [5]	✓	✓	✓	–	–

3.1 Goal and Design

The goal of the focus group was to establish a foundation for identifying feedback modalities and dimensions commonly used in the classroom that warrant further investigation. The focus group design follows the definition and eight characteristics introduced by Hennink [17]. To mitigate potential challenges like *production blocking* and *evaluation apprehension* during brainstorming, the Affinity Diagram (AD) method was employed, which facilitates organizing ideas by grouping similar concepts and promoting a structured problem-solving process [22]. This method helped enable smooth interaction within the focus group.

Tasks. We conceptualized three focus group tasks. The first task aimed to identify the types of feedback professionals currently provide to students after exercise completion, specifically distinguishing between feedback for correctly and incorrectly solved exercises. Following the initial task, two additional tasks expanded on the feedback types identified in Task 1. Task 2 introduced the concept of using SAR for feedback, prompting participants to consider how this technology could enhance feedback delivery. Finally, Task 3 involved brainstorming potential real-world applications of SAR, where participants developed scenarios for using SAR to provide feedback in educational contexts. The specific focus group tasks are outlined below:

> **T_1** Which kind of feedback do you provide to learners after task completion?
> **T_2** Which feedback would you provide after task completion leveraging SAR technology?
> **T_3** What could a possible scenario be in which learners receive feedback with the help of SAR?

3.2 Participants

For focus group participant recruitment, we defined two selection criteria to ensure they shared adequate expertise in teaching. First, all participants were required to have prior experience in practical teaching. Additionally, they were required to be currently enrolled in or have completed a teacher education program in the German educational system. Three of the participants have obtained a Bachelor of Education, and all five had prior teaching experience. As a result, the group consisted of female participants (N = 5) aged between 24 and 28 years (M = 25.32, SD = 2.1).

3.3 Procedure

The focus group followed the established design, beginning with introductions and proceeding to three tasks. A five-minute presentation introduced participants to the technologies and focus group objectives. After the introduction, participants completed the first task, which involved identifying feedback types using the AD method. Each participant recorded responses individually, after which one was selected to present their results. The group then organized all responses into clusters, iteratively repeating the process until all contributions were represented in a visual map of feedback modalities. For tasks two and three, participants received worksheets with task descriptions and response spaces, shown one at a time to prevent bias. As in Task 1, they had five minutes to respond individually before the group discussion.

3.4 Results and Implications

Our focus group revealed the following results.

T_1: Feedback modalities in classrooms. $\mathbf{T}_1$ focused on identifying feedback modalities, resulting in three primary categories: *visual*, *audio*, and *gestural*. Additionally, participants defined an *emotional* category and a *miscellaneous* category for feedback types not aligned with the primary groups.

Visual feedback was further split into graphics, shapes, and text. Examples of these graphics were *stamps*, which could be used to mark achievement or progress, *stickers*, often serving as fun and encouraging markers of success, and *emojis*, which offered a quick, emotionally resonant way to express approval or disapproval. For shapes, *ticks*, *marks*, and *arrows* were mentioned, while text included *notes* of encouraging messages such as "You improved!" and *side notes* for correction.

Auditory feedback involved verbal interactions such as *vocal praise* and *corrective guidance*. This category included spoken *words of encouragement* such as "great job" or "excellent work," as well as *constructive verbal corrections* aimed at helping the recipient improve their performance. Gestural feedback included *high five* or *thumbs-up* gestures, as well as facial expressions that conveyed approval and encouragement, such as *smiling*.

Emotional feedback deploys positive reinforcement, emphasizing encouragement, praise, and recognition. Typical expressions included affirmations such as "well done", "awesome", or "correct", often paired with the explicit use of the child's name to personalize the feedback. Participants also highlighted encouraging children to revise their work, encouraging repeated attempts for difficult exercises, and offering supportive comments on formal aspects such as neatness or structure. Feedback could be delivered verbally or as written notes.

Finally, the miscellaneous feedback group included any forms of feedback that the participants could not place into one of the aforementioned categories. This included ideas such as *leaderboards*, *sample solutions*, *consolation*, and using appropriate language. Moreover, the participants discussed ways to guide students in correcting their errors themselves.

T_2: SAR Feedback modalities. The focus group results for $\mathbf{T}_2$, which explored the potential of SAR technology for providing feedback after student exercise completion, revealed a strong preference for a variety of digital, interactive, and personalized feedback mechanisms. The participants highlighted the effectiveness of displaying images alongside text, as well as the ability to click on text to reveal related visuals, creating a more engaging and responsive learning experience. Additionally, gamified avatars (e.g., mascots representing each class) and digital score systems were suggested as motivational tools to increase learner engagement. The group also emphasized the value of incorporating tutorial and guide videos, detailed mistake analysis, and animations, which provide comprehensive and dynamic feedback not feasible in traditional educational settings. Furthermore, personalized features (e.g., reading texts aloud), mentioning the student's name, and providing real-time corrective feedback were identified as key advantages.

T_3: Independent Learning through SAR. Finally, for $\mathbf{T}_3$, the participants pointed to several applications that enhance independent learning. They suggested that SAR could be utilized to provide individualized feedback during self-guided learning sessions, enabling students to work autonomously without the need for constant teacher supervision. SAR was also seen as a tool to correct final responses, offering immediate feedback upon exercise completion, even in the absence of a teacher. Furthermore, the system could be particularly beneficial in situations where teacher presence is limited or unavailable, allowing for continuous learning and assessment. These scenarios highlight the role of SAR in fostering independent learning environments, where personalized, real-time feedback can be delivered seamlessly beyond what is typically possible in traditional classroom settings.

3.5 Implications for the Experiment Design

Based on the ideas collected from the focus group and related work, we derived our experimental design.

Through $\mathbf{T}_1$, we learned about a variety of possible feedback modalities that are commonly used in classrooms. For this work, we focused on visual, gestural, and emotional feedback. We used the ideas of *emojis*, *encouraging text messages*, and *affirmations*, and gestures like *thumbs-up*. Regarding $\mathbf{T}_2$, we learned that personalized feedback is an interesting aspect of using an SAR system. We also focused on showing videos, individualized messages, and images, which were highlighted as an essential advantage compared to traditional feedback methods. In $\mathbf{T}_3$, we found that the utility of self-supervised learning enabled by an SAR system, particularly the provision of immediate feedback in the absence of a teacher, represents a key advantage. Consequently, we implemented immediate feedback in our study.

In summary, our study's feedback modalities were derived from the literature review and focus group findings. We identified key educational benefits of SAR systems, particularly personalized video and image feedback and the immediacy of responses. From this, we defined three modalities—*text*, *image*, *video*. To integrate emotional feedback, we adopted three familiarity levels based on literature: visually familiar (*neutral*), experimentally learned (*unfamiliar*), and personally familiar (*familiar*) [35].

4 Field Study

To investigate preferences for SAR feedback in the classroom and the influence of familiarity with the feedback provider, we conducted a field study. The study aimed to address the following research question:

(**RQ**) Which spatial augmented reality feedback modality – *text*, *image*, *video* – and which level of familiarity – *neutral*, *unfamiliar*, *familiar* – are educationally appropriate in classroom settings?

The primary goal of this work is to determine which SAR feedback modality should be used in the school context. Therefore, we decided to investigate the clarity, distraction, and novelty of these modalities. Feedback needs to be clearly understandable and recognizable by students to be valuable. Additionally, we want to minimize the distraction that students perceive from the feedback.

4.1 Hypotheses

We hypothesize that video feedback is generally preferred but may be more distracting than other types. We also hypothesize that familiarity with the feedback provider increases the perceived trustworthiness and helpfulness of the feedback. Additionally, we expect that feedback, even when negative, will not negatively impact students' mood. Based on these considerations, we formulated several hypotheses.

> **H_1** (a) Clarity, (b) distractedness, (c) novelty are higher in *video* than in *image* and *text*.
> **H_2** *Familiar* feedback is prefered more often than (a) *neutral* or (b) *unfamiliar* feedback.
> **H_3** Fun is higher in *video* than *image* and *text*.
> **H_4** (a) Boredness and (b) frustration are lower in *video* than *image* and *text*.

We posit hypothesis **H_1(b)** because the video feedback might be more distracting than other feedback modalities, as the moving content attracts the user's focus. Borup et al. found no significant difference in video and textual feedback quality or delivery [4]. In contrast, other work explored the effectiveness of textual and video feedback in computer-assisted learning [26]. Here, the video showed significantly higher learning and retention scores. This leads us to posit **H_1(a)** and **H_1(c)**. The novelty of our approach lies in the investigation of these methods in a classroom-based context using SAR technology. This classroom-based context allows for a representative design [13].

Previous work showed that student-teacher familiarity positively affects achievement, as students matched with the same teacher in consecutive years show improved test scores [18]. To analyze this aspect, we adapted the familiarity levels defined by Ramon and Gobbini [35]. We considered three levels of familiarity: personally familiar (*familiar*), visually familiar (*neutral*), and experimentally learned (*unfamiliar*). Furthermore, by applying the mere-exposure paradigm [56], we categorized familiarity into zero exposure (*unfamiliar*) and repeated encounters (*familiar*). We posit **H_2** based on this framework.

Because we assume that video feedback is more motivating than text or image feedback, we posit **H_3** and **H_4**. This is reflected by findings showing that video-based learning increases student satisfaction, attracts students' attention, motivates them, and thus increases their in-class participation [39].

4.2 Study Design

We designed a within-subject field study. Our first independent variable is the feedback modality with three levels: *text*, *image*, *video*. Our second independent variable is the correctness of the task with two levels: *correct*, *wrong*. Thus, we explored the following six study conditions: *text+correct*, *text+wrong*, *image+correct*, *image+wrong*, *video+correct*, *video+wrong*. The first variable – modality – is chosen based on the previous literature review and focus group. These three levels were selected as they cover the visual feedback spectrum. We decided not to use any audio feedback as we focused on visual AR feedback. A Latin square design ensured that the order of feedback presentation (*text*, *image*, or *video*) was counterbalanced across participants. We decided to use correctness as a secondary independent variable as our professionals in the focus group pointed out that feedback should be individual and based on task correctness. Here, our dependent variables are clarity, distraction level, novelty and mood.

In the second part of our study, we introduced another independent variable – familiarity – with three levels: *unfamiliar*, *neutral*, *familiar*. Here, we investigated these six study conditions: *neutral+correct*, *neutral+wrong*, *unfamiliar+correct*, *unfamiliar+wrong*, *familiar+correct*, *familiar+wrong*. The abstraction of the familiarity levels for the image conditions is illustrated on the right side of Figure 1. For the video conditions, feedback was provided through an entity with varying levels of familiarity, analogous to the image conditions, but presented in video format. In the case of text conditions, familiarity levels were conveyed through signatures associated with the entity, corresponding to each level of familiarity. Through our focus group and literature review, we determined that feedback has several hidden dimensions that could also affect the perception of given feedback. We decided to use an explorative approach to study the familiarity with the feedback provider as a secondary part of our study to investigate how it affects the perceived feedback.

We employed a mixed methods approach, collecting both quantitative and qualitative data. To determine the clarity, distraction level, and novelty of the feedback modalities, we used three custom seven-point Likert items. We created the following statements: "This feedback is clear." (**$L_{clarity}$**), "This feedback is distracting." (**$L_{distraction}$**), "This feedback is novel." (**$L_{novelty}$**).

Additionally, we asked participants to select a preferred modality and familiarity. Pekrun et al. [34] showed that the emotions of fun, boredom, and frustration are related to the students' motivation and academic achievement. We, therefore, recorded this variable to understand if different AR feedback types promote changes in student mood. We used a questionnaire adapted from Riemer and Schrader [36] and used by Lauer et al. [27]. The questionnaire captured the participants' current levels of three emotions (*fun*, *boredom*, *frustration*) using a 5-point Likert scale. Finally, we conducted semi-structured interviews with participants, collecting qualitative data.

4.3 Apparatus

In this study, we employed the ProLamp, a projector-camera setup integrated into a single apparatus, to provide feedback [51]. The system utilizes a small projector and camera to project 2D feedback directly onto physical surfaces. This setup allows for the flexible projection of visual feedback, including images and text, onto any desired location in real time, making it a suitable tool for delivering interactive and contextual information to participants. The system is shown in Figure 2 to illustrate its configuration and function. A replica of the apparatus can be seen in Figure 1 on the left. Participants were seated at a desk. On the desk, the projector-camera setup was deployed.

4.4 Power Analysis

To determine our required sample size, we conducted an a priori power analysis for a repeated-measures within-factors analysis of variance (ANOVA), given α = .05, β = .8, and effect size f = .4. Following our experimental study design, we set the number of groups to one and the number of measurements to six. G*Power[1] 3.1.9.7 suggested a total sample size of 8 (λ = 15.36, F = 2.49).

[1] https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower, last retrieved on October 21, 2025

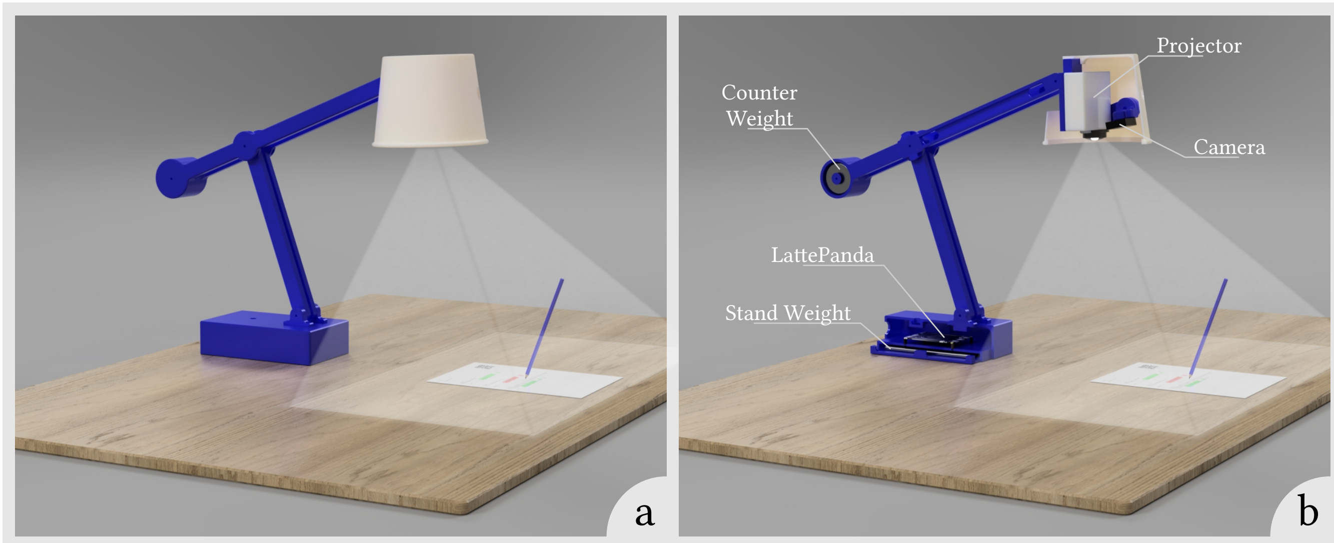


Figure 2: The ProLamp: A projector-camera system enabling spatial AR Feedback (a) and an intersection view of the device, showing its internal structure and components (*b*) [51]

4.5 Participants

The study was conducted with a group of 16 students aged between 14 and 16 years (M = 15.07, SD = .68). All of the participants were attending the ninth grade at high school. The group comprised nine boys and seven girls. The participants were students from a single ninth-grade class.

4.6 Ethics

Before conducting the study, we received approval from our institution's ethics committee. In addition, we obtained informed consent from the legal guardians of the students with student assent, ensuring that they understood their right to withdraw from the study at any point without any detriments. To further safeguard participants' privacy, we assigned randomized IDs to each individual and destroyed the mapping from individual to ID afterward.

4.7 Procedure

The study followed a within-subjects design to examine the impact of feedback modality, task correctness, and familiarity with the feedback provider on students' perceptions. Before the study began, the participants were asked to have their legal guardian sign an informed consent form, ensuring they were fully aware of the study's goals, purpose, and procedure, as well as their rights as participants. This process emphasized the confidentiality and voluntary nature of the study, and only participants with signed consent were allowed to proceed.

Once consent was obtained, the participants were introduced to the study. The introduction outlined the goals, explaining that the research focused on investigating how different feedback modalities—*text*, *image*, and *video*—affected their understanding and engagement, especially when linked to the correctness of their answers and the familiarity of the feedback provider. After they had given assent, the participants were given a worksheet containing simple pre-answered math tasks with a mixture of correct and incorrect responses. The purpose of this worksheet was to simulate a real-world scenario where feedback would be provided without making the students solve the tasks themselves. We did this to ensure that the task correctness could be controlled and that there were no emotions from solving tasks either correctly or incorrectly, which could influence the perception of the feedback itself. To foster a sense of attachment to the task, the experimenter discussed the correctness of each answer with the participants, helping them engage more deeply with the material. Following this discussion, the experimenter guided each participant to the correct solutions to ensure the same task correctness for all participants.

The system then presented the first feedback modality. The feedback was based on the correctness of the task, with participants receiving feedback on both correct and incorrect answers. After they had viewed the feedback, the participants were asked to rate it using a Likert scale, assessing its clarity, novelty, and level of distraction. This process was repeated until each participant had experienced all three feedback modalities for correct and incorrect answers. After all feedback modalities had been experienced, the participants were asked to choose their preferred modality. This step aimed to determine which modality they found most suitable for educational contexts.

In the second phase of the study, the participants were introduced to feedback with an additional independent variable: familiarity with the feedback provider. Feedback was presented for both correct and incorrect tasks, with the participants seeing text, image, and video modalities from feedback providers with varying levels of familiarity (*neutral*, *unfamiliar*, *familiar*). Once the participants had received feedback from all three levels of familiarity across all modalities, they were asked to select their most preferred level of familiarity. Finally, each participant partook in a semi-structured interview, during which they discussed their preferences and provided insights into how familiarity influenced their feedback experience. The study procedure took approximately 40 minutes on average.

4.8 Results

First, we present the quantitative results related to the answered questionnaires and Likert items. Second, we provide insights derived from the conducted interviews (qualitative results).

To examine the effect of the independent variables *modality* and *correctness* on the students' ratings of *clarity*, *distraction*, and *novelty*, we conducted three separate two-way factorial repeated-measures analyses of variance (RM-ANOVAs). The *modality* variable had three levels (*image*, *text*, and *video*), while *correctness* had two levels (*correct* and *wrong*). Students answered these Likert items after each condition on a scale of -3 to +3, where -3 corresponds to "Strongly Disagree" and +3 corresponds to "Strongly Agree".

Assumptions. We first checked the assumptions for all three RM-ANOVAs. Shapiro-Wilk tests indicated that none of the dependent variables (i.e., students' Likert responses) followed a normal distribution (*novelty*: $W = 0.9244$, $p < .0001$; *distraction*: $W = 0.8813$, $p < .0001$; *clarity*: $W = 0.7780$, $p < .0001$). Therefore, we applied the Aligned Rank Transform (ART) procedure before running one RM-ANOVA per dependent variable [52].

4.8.1 Main Effects and Post-hoc Tests.

We mostly found statistically significant main effects for *modality* for all three dependent variables. For brevity, we only report significant contrast tests ($p < .05$). Figure 3 provides an overview.

Ratings for Clarity. For *clarity*, we found a significant main effect of *modality* with F(2, 75) = 3.7266, p = .0286, η_p^2 = .0904. However, the post-hoc t-tests did not indicate statistical significance, as *text* (Med. = 3.00, IQR = 1.00) receiving only marginally higher ratings than *image* (Med. = 2.00, IQR = 1.25) with t(75) = -2.3625, p = .0617, and *text* rated marginally higher than *video* (Med. = 2.00, IQR = 1.25), t(75) = 2.3661, p = .0617.

Ratings for Distraction. For *distraction*, another significant main effect of *modality* was observed F(2, 75) = 10.0696, p = .0001, η_p^2 = .2117. Post-hoc tests showed that *video* (Med. = -1.00, IQR = 2.00) received significantly higher ratings than both *image* (Med. = -2.00, IQR = 2.00) with t(75) = -2.9442, p = .0086 and *text* (Med. = -2.00, IQR = 1.00), t(75) = -4.4052, p = .0001. In the pair-wise contrast tests, it was particularly found that *text+correct* (Med. = -2.00, IQR = 1.00) provided significantly lower ratings than *video+wrong* (Med. = -1.00, IQR = 1.25), t(75) = -3.8893, p = .0032. In addition, *text+wrong* (Med. = -2.00, IQR = 1.25) provided significantly lower ratings than *video+wrong*, t(75) = -3.6546, p = .0067.

Ratings for Novelty. For *novelty*, again, a significant main effect of *modality* was found, F(2, 75) = 11.8398, p < .0001, η_p^2 = .2400. Post-hoc pairwise comparisons with Holm's correction revealed that *video* (Med. = 1.00, IQR = 2.00) was rated significantly higher than both *image* (Med. = 0.00, IQR = 2.00) with t(75) = -4.2629, p = .0002 and *text* (Med. = -1.00, IQR = 2.25), t(75) = -4.1639, p = .0002. Additionally, the following contrasts for *modality* and *correctness* were significant: *image+wrong* (Med. = -1.50, IQR = 2.00) was rated lower than *video+correct* (Med. = 1.00, IQR = 1.25) with t(75) = -3.2189, p = .0285 and lower than *video+wrong* (Med. = 1.00, IQR = 1.25), t(75) = -3.0383, p = .0425. Similarly, *text+wrong* (Med. = -1.00, IQR = 2.25) was rated lower than *video+correct* with t(75) = -3.1660, p = .0313 and lower than *video+wrong*, t(75) = -2.9855, p = .0459.

Main Effect for Correctness and Interaction Effects. No significant main effects were observed for *correctness* across any of the dependent variables: i) for *clarity* with F(1, 75) = 1.3544, p = .2482, η_p^2 = .0177, ii) for *distraction* with F(1, 75) = .7560, p = .3874, η_p^2 = .0100, and iii) for *novelty* with F(1, 75) = .0021, p = .9639, η_p^2< .0001. Also, we did not observe any significant interaction effect.

Summary. To summarize, we find statistically significant support for hypotheses $\mathbf{H_1(b)}$ and $\mathbf{H_1(c)}$, where *video* lead to consistently higher scores than *text* or *image* in *novelty* and *distraction*. For $\mathbf{H_1(a)}$, which is related *clarity*, we found a significant effect of group, but the corresponding post-hoc t-tests failed to support that there exists a significant difference. Also, here, *text* was deemed to yield the highest *clarity* values. Overall, we see by the η_p^2 for *modality* that its influence is deemed a large effect in *distraction* (η_p^2 = .2117) and *novelty* (η_p^2 = .2400) [11, p. 79 ff.] In addition, we did not observe statistical differences that *novelty*, *clarity*, or *distraction* were influenced by *correctness*. Thus, we have not found any quantitative evidence that the actual correctness of the feedback posits any influence on these three variables.

4.8.2 Familiarity.

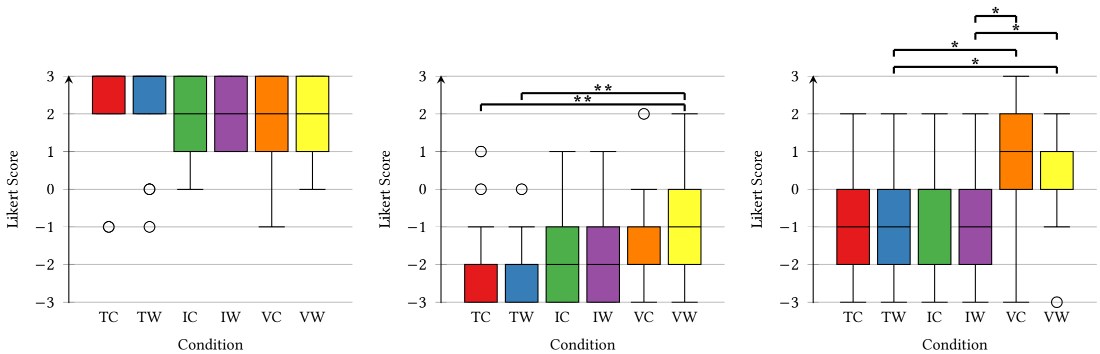
To analyze the data collected concerning the preference and ranking for familiarity in feedback, we first conducted a Cochran's Q test for all six combinations of feedback in regard to modality and task correctness. The tests for all combinations – text correct ($\chi^2(2)$ = 6.13, p = 0.047), text wrong ($\chi^2(2)$ = 9.88, p = 0.0072), image correct ($\chi^2(2)$ = 8.93, p = 0.011), image wrong ($\chi^2(2)$ = 9.73, p = 0.0077), video correct ($\chi^2(2)$ = 7.13, p = 0.028), video wrong ($\chi^2(2)$ = 7.13, p = 0.028) – indicated a statistically significant difference between proportions of preference for the three familiarity conditions (*neutral*, *unfamiliar*, *familiar*).

A pairwise comparison using continuity-corrected McNemar's tests with Bonferroni-Holm correction revealed that significantly more participants preferred neutral than unfamiliar (p<0.1, ϕ = 0.47) for image correct and familiar than unfamiliar (p<0.1, ϕ = 0.40). We also found this for image wrong – neutral over unfamiliar (p<0.1, ϕ = 0.39) and familiar over unfamiliar (p<0.1, ϕ = 0.36), video correct – neutral over unfamiliar (p<0.1, ϕ = 0.39) and familiar over unfamiliar (p<0.1, ϕ = 0.40) – and video wrong – neutral over unfamiliar (p<0.1, ϕ = 0.35) and familiar over unfamiliar (p<0.1, ϕ = 0.44).

This supports our hypothesis $\mathbf{H_2(b)}$ but does not support $\mathbf{H_2(a)}$.

4.8.3 Mood.

To analyze our results regarding the mood variable, we first used a Friedman test. The results indicate significant effects after feedback modality usage regarding fun ($\chi^2(2) = 11.07$, $p = 0.011$, $N = 16$). For boredom ($\chi^2(2) = 5.37$, $p = 0.15$, $N = 16$) or frustration, we did not find any significant shifts in mood. For the frustration, all participants answered with "Strongly Disagree" for all conditions. Thus, we did not conduct any further tests to find differences. This fails to support hypothesis $\mathbf{H_4(a)}$ and $\mathbf{H_4(b)}$. As a post hoc analysis, we used a Wilcoxon test with Bonferroni-Holm correction. The Wilcoxon Signed-rank test shows that there is a significant effect of feedback modality (W = 1, Z = 2.65, p < 0.05, r = 0.66). This supports $\mathbf{H_3}$.



(a) Clarity (b) Distraction (c) Novelty

Figure 3: Boxplots for pair-wise contrast tests for the three Likert items – (a) clarity, (b) distraction, and (c) novelty. Split for each of the six conditions: *text+correct* (TC), *text+wrong* (TW), *image+correct* (IC), *image+wrong* (IW), *video+correct* (VC), *video+wrong* (VW). The observed significant differences are indicated with * ($p < .05$) and ** ($p < .01$).

Qualitative Results. In our study, we also asked participants to explain why they chose the feedback modality and familiarity level to get further insights into the explored feedback dimensions. During the study, the experimenters documented participants' statements in the form of summarized notes. These notes were then analyzed through an inductive open coding process, during which one researcher assigned codes to each statement.

Familiar Feedback and Task Correctness. Participants highlighted the importance of task correctness in their interactions, especially in relation to the familiarity of the feedback source. Participants (6 statements) remarked that receiving negative feedback from a known person, particularly when a task was incorrect, left them feeling disappointed and demotivated. For instance, one participant stated, "*When a person tells you that it's wrong, you get demotivated and disappointed with yourself*" (P13), emphasizing how personal feedback affects emotional response. Another participant echoed this sentiment, saying, "*Negative feedback is not something you like to hear from people*" (P14), underlining the emotional burden of receiving corrective feedback from a human source rather than a neutral one. This discomfort was particularly pronounced when participants knew the individual providing the feedback. As one remarked, "*If you know the person, it affects you more, which is why no names should be used*" (P10), indicating that familiarity intensified the emotional impact of the feedback. Participants (2 statements) also noted that receiving feedback from a neutral source felt less emotionally charged, as one commented, "*Neutral feedback is just information. Personal [feedback] is better*" (P13).

Another participant remarked that they liked: "*to get the feeling of being praised by a person*" (P14). This highlights the complex interplay between task correctness and the feedback mechanism, with familiar sources amplifying the emotional response to errors.

Trust. Participants expressed a clear preference for receiving feedback from individuals they trust, noting that such feedback felt more meaningful and constructive. E.g., one participant stated, "*Knowing the person for a longer time, [you] know their competence and understand their evaluation [better]*" (P07), emphasizing the value of familiarity and trust in the feedback process. This sense of trust appeared to heighten the participants' receptiveness to feedback, with another noting, "*[It's] more personal, you know who evaluated it*" (P04). The personal connection thus played a crucial role in how feedback was perceived, making it more impactful and relevant when it came from trusted sources. E.g., one participant expressed, "*A neutral person would be more disappointing [because] you make a bad impression*" (P14), demonstrating that it might be uncomfortable to receive feedback from strangers. Additionally, participants (4 statements) remarked that receiving feedback from people felt worse when there was no personal connection. This distinction between trusted and unfamiliar feedback sources highlights the participants' desire for a personal connection in evaluative interactions, reinforcing the importance of trust in the feedback process.

5 Discussion

In this section, we will discuss our findings in terms of modality trade-offs and the relation between positive and negative feedback regarding feedback provider familiarity. We situate our discussion within the context of previous research. Our results support hypotheses $\mathbf{H_1(b)}$, $\mathbf{H_1(c)}$, $\mathbf{H_2(b)}$, and $\mathbf{H_3}$.

5.1 Modality Trade-offs

Feedback can be delivered through multiple modalities, each with its own advantages and potential drawbacks. In educational contexts, choosing the right modality is crucial for maximizing student engagement, comprehension, and comfort. We discuss these trade-offs according to the results of our quantitative analysis.

5.1.1 Video feedback is distracting. Video feedback introduces additional cognitive load, as it requires one to focus on visual and auditory elements simultaneously. The presence of video may divert attention from the actual content of the feedback to the speaker's facial expressions, body language, or even background. This can particularly affect students who struggle with focus or are sensitive to multi-sensory input. This is in line with the results of our quantitative analysis $\mathbf{H_1(b)}$. In addition, previous foundational work has shown that split-source information can generate high cognitive load [9]. The added personal dimension may also cause the feedback to feel overwhelming or overly intrusive, negatively impacting the learning experience.

5.1.2 Video feedback is novel and fun. On the other hand, video feedback can provide a more personal and engaging experience. It allows instructors to communicate nuance through tone, facial expression, and gestures, making the feedback feel more genuine and relatable. This can increase the perceived authenticity of the feedback and foster a stronger connection between students and instructors, especially in remote or online learning environments [49]. Video may also be able to clarify complex ideas that would be difficult to convey through text alone, providing students with a richer, more holistic learning experience. We base this on our supported hypothesis $\mathbf{H_1(c)}$ and $\mathbf{H_3}$. We did not observe significant differences in clarity between modalities. This does not suggest that there is no effect, but we were not able to find a large effect. Overall, clarity ratings were high across all conditions, suggesting that all modalities were perceived as clear by participants.

5.2 Familiarity and Task Correctness

The accuracy or correctness of a student's performance plays a role in how they perceive and prefer feedback. Students often adjust their preferences for feedback based on whether the feedback was positive or negative. In our quantitative results, we found no evidence that task correctness influenced clarity, distraction, or novelty. Yet, our qualitative results showed that task correctness influences feedback perception when affected by variance in familiarity level with the feedback provider.

5.2.1 Students want to receive positive feedback from trustworthy people. Students generally seek validation from individuals they trust, especially when the feedback is positive. Positive feedback from a familiar, reliable figure—such as a known teacher or mentor—tends to reinforce their learning and boost their self-confidence. As derived from our qualitative analysis, when the source of the feedback is perceived as credible, the positive remarks feel more meaningful and are more likely to motivate continued effort and improvement. This is in line with related work, specifically stating the importance of trust [6].

5.2.2 Students do not want to receive negative feedback from people. Negative feedback can be difficult for students to accept, particularly when it comes from individuals they know personally or respect. Students prefer to distance themselves from negative feedback, as it can feel more emotionally charged or personally critical. This often leads to a preference for receiving critical feedback through less personal modalities, such as text, where the feedback carries less emotional weight. Generally, previous work has shown that the impact of negative feedback can elevate student learning experiences for some but worsen them for others, depending on the individual [16]. Thus, providing negative feedback through digital methods that are seen as less personal might improve learning experiences, which is in line with our findings.

5.2.3 Feedback from strangers is uncomfortable. When feedback comes from someone unfamiliar, such as a stranger or an anonymous figure, it can feel awkward or disconnected ($\mathbf{H_2(b)}$). Students might question the credibility of the feedback and feel less inclined to engage with it, especially when it's delivered in personal modalities such as video. The lack of a prior relationship between the student and the feedback provider may also lead students to perceive the feedback as less relevant or out of context for their specific learning journey.

5.3 Which feedback modality should be used?

Our qualitative results highlight that the choice of feedback modality depends on the nature of the feedback and the student's familiarity with the feedback provider. Text-based feedback is often sufficient and efficient for routine, impersonal feedback, such as clarifications on assignments. However, for more personalized or motivational feedback, especially when delivered by a trusted figure, video or image feedback can offer added value through personal connection and emotional reinforcement. A balanced approach might involve using video or image feedback at more significant points in a student's learning process–such as at the end of a term or during important projects–while relying on text-based feedback for everyday assignments and corrections. In addition, giving students some control over how they receive feedback could enhance their comfort and engagement, tailoring the process to individual preferences and sensitivities.

6 Limitations and Future Work

This study has several limitations. Students did not complete tasks themselves but evaluated pre-filled worksheets. We deliberately chose this approach to minimize emotional attachment to exercise outcomes, ensuring that evaluations of negative feedback were less influenced by participants' emotional responses to errors, while recognizing that this design choice may limit the ecological validity of the results. We measured clarity, distraction, and novelty using custom Likert items. Employing a standardized questionnaire could have strengthened measurements. Furthermore, we conducted the study in a single ninth-grade classroom (N = 16) within one cultural and educational context, which restricts the generalizability of the results to other age groups or school systems. Finally, we acknowledge that our power analysis was based on a large effect size, which limits our ability to detect smaller effects. Therefore, non-significant results should not be interpreted as evidence for the absence of small to medium effects.

Future work can explore dynamic and contextual feedback through SAR. Researchers can customize and adapt these systems to address student needs, classroom configurations, and ongoing tasks, thereby enhancing their applicability in authentic educational settings. The integration of AI-based techniques to detect when students require assistance or feedback can be a step toward the development of more responsive and adaptive systems.

7 Conclusion

This work examined how feedback modality and familiarity influence users' perception and reception of feedback in educational contexts. Through a literature review and a focus group study

(N = 5), we identified key feedback modalities (*text*, *image*, *video*) and their respective strengths and weaknesses. A subsequent field study (N = 16) further investigated how participants perceived these modalities in relation to feedback provider familiarity and task correctness.

Results showed that video feedback was perceived as novel and engaging but also distracting, likely due to its cognitive and emotional load. Participants valued the richer communication of video but noted that it could shift attention away from the feedback itself. In contrast, text-based feedback was considered more neutral and less intrusive, particularly when the feedback was corrective rather than affirming. Task correctness also shaped how feedback was received. Positive feedback was preferred from familiar sources, while negative or corrective feedback produced more varied preferences, with some participants favoring less personal modes. However, quantitatively, task correctness did not affect clarity, distraction level, or novelty.

Based on these findings, we recommend selecting feedback modalities according to context. For positive feedback in familiar settings, video can strengthen trust and connection. For corrective feedback or when addressing task performance, neutral modalities such as text may be more suitable, offering clarity without emotional distraction. When familiarity with the feedback giver is low, text can provide a more comfortable and digestible experience.

Acknowledgments

This work is partially funded by the German Federal Ministry of Education and Research (16SV8663).

References

[1] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (1997), 355–385. doi:10.1162/pres.1997.6.4.355

[2] S Bharuthram and M van Heerden and. 2023. The affective effect: Exploring undergraduate students' emotions in giving and receiving peer feedback. *Innovations in Education and Teaching International* 60, 3 (2023), 379–389. doi:10.1080/14703297.2022.2040567

[3] Oliver Bimber and Ramesh Raskar. 2006. Modern Approaches to Augmented Reality. In *ACM SIGGRAPH 2006 Courses on - SIGGRAPH '06*. ACM Press, Boston, Massachusetts, 1. doi:10.1145/1185657.1185796

[4] Jered Borup, Richard E. West, and Rebecca Thomas. 2015. The Impact of Text versus Video Communication on Instructor Feedback in Blended Courses. *Educational Technology Research and Development* 63, 2 (April 2015), 161–184. doi:10.1007/s11423-015-9367-8

[5] Tim Bosch, Gu van Rhijn, Frank Krause, Reinier Könemann, Ellen S. Wilschut, and Michiel de Looze. 2020. Spatial augmented reality: a tool for operator guidance and training evaluated in five industrial case studies. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu, Greece) *(PETRA '20)*. Association for Computing Machinery, New York, NY, USA, Article 40, 7 pages. doi:10.1145/3389189.3397975

[6] David Boud and Elizabeth Molloy and. 2013. Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education* 38, 6 (2013), 698–712. doi:10.1080/02602938.2012.691462

[7] Matt Bower, Cathie Howe, Nerida McCredie, Austin Robinson, and David Grover. 2014. Augmented Reality in education – cases, places and potentials. *Educational Media International* 51, 1 (2014), 1–15. doi:10.1080/09523987.2014.889400

[8] Joshua Chan, Tarun Pondicherry, and Paulo Blikstein. 2013. LightUp: an augmented, learning platform for electronics. In *Proceedings of the 12th International Conference on Interaction Design and Children*. ACM, New York New York USA, 491–494. doi:10.1145/2485760.2485812

[9] Paul Chandler and John Sweller. 1991. Cognitive load theory and the format of instruction. *Cognition and instruction* 8, 4 (1991), 293–332. doi:10.1207/s1532690xci0804_2

[10] Hyunsung Cho, Drew Edgar, David Lindlbauer, and Joseph O'Hagan. 2025. Evaluating Dynamic Delivery of Audio+Visual Message Notifications in XR. In *2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 277–287. doi:10.1109/VR59515.2025.00052

[11] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Routledge. doi:10.4324/9780203771587

[12] European Commission, Sport Directorate-General for Education, Youth, and Culture. 2023. *Education and training monitor 2023 – Comparative report*. Publications Office of the European Union. https://data.europa.eu/doi/10.2766/810689

[13] Mandeep K Dhami, Ralph Hertwig, and Ulrich Hoffrage. 2004. The role of representative design in an ecological approach to cognition. *Psychological bulletin* 130, 6 (2004), 959. doi:10.1037/0033-2909.130.6.959

[14] Hilal Ergül. 2023. The Case for Smiling? Nonverbal Behavior and Oral Corrective Feedback. *Journal of Psycholinguistic Research* 52, 1 (Feb. 2023), 17–32. doi:10.1007/s10936-021-09807-x

[15] Rubina Freitas and Pedro Campos. 2008. SMART: a SysteM of Augmented Reality for Teaching 2nd grade students. In *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 2* (Liverpool, United Kingdom) *(BCS-HCI '08)*. BCS Learning & Development Ltd., Swindon, GBR, 27–30. doi:10.14236/ewic/HCI2008.26

[16] John Hattie and Helen Timperley. 2007. The Power of Feedback. *Review of Educational Research* 77, 1 (2007), 81–112. doi:10.3102/003465430298487

[17] M.M. Hennink. 2013. *Focus Group Discussions*. Oxford University Press. https://books.google.de/books?id=5DLLAgAAQBAJ

[18] NaYoung Hwang, Brian Kisida, and Cory Koedel. 2021. A familiar face: Student-teacher rematches and student achievement. *Economics of Education Review* 85 (2021), 102194. doi:10.1016/j.econedurev.2021.102194

[19] Hamideh Iraj, Anthea Fudge, Huda Khan, Margaret Faulkner, Abelardo Pardo, and Vitomir Kovanović. 2021. Narrowing the Feedback Gap: Examining Student Engagement with Personalized and Actionable Feedback Messages. *Journal of Learning Analytics* 8, 3 (Nov. 2021), 101–116. doi:10.18608/jla.2021.7184

[20] Qiao Jin, Danli Wang, Xiaozhou Deng, Nan Zheng, and Steve Chiu. 2018. AR-maze: a tangible programming tool for children based on AR technology. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. ACM, Trondheim Norway, 611–616. doi:10.1145/3202185.3210784

[21] Sasa Junuzovic, Kori Inkpen, Tom Blank, and Anoop Gupta. 2012. IllumiShare: sharing any surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1919–1928. doi:10.1145/2207676.2208333

[22] Jiro Kawakita. 1991. The original KJ method. *Tokyo: Kawakita Research Institute* 5 (1991), 1991.

[23] Jeongyun Kim, Jonghoon Seo, and Tack-Don Han. 2014. AR Lamp: Interactions on Projection-Based Augmented Reality for Interactive Learning. In *Proceedings of the 19th International Conference on Intelligent User Interfaces* (Haifa, Israel) *(IUI '14)*. Association for Computing Machinery, New York, NY, USA, 353–358. doi:10.1145/2557500.2557505

[24] Avraham N Kluger and Angelo DeNisi. 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin* 119, 2 (1996), 254. doi:10.1037/0033-2909.119.2.254

[25] Aaron Kotranza, D. Scott Lind, Carla M. Pugh, and Benjamin Lok. 2009. Real-time in-situ visual feedback of task performance in mixed environments for learning joint psychomotor-cognitive tasks. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. IEEE, Orlando, FL, USA, 125–134. doi:10.1109/ISMAR.2009.5336485

[26] James P. Lalley. 1998. Comparison of Text and Video as Forms of Feedback during Computer Assisted Learning. *Journal of Educational Computing Research* 18, 4 (June 1998), 323–338. doi:10.2190/LXNP-WAPB-VH9A-HFRW

[27] Luisa Lauer, Kristin Altmeyer, Sarah Malone, Michael Barz, Roland Brünken, Daniel Sonntag, and Markus Peschel. 2021. Investigating the Usability of a Head-Mounted Display Augmented Reality Device in Elementary School Children. *Sensors* 21, 19 (Jan. 2021), 6623. doi:10.3390/s21196623

[28] Jingya Li, Erik Van Der Spek, Jun Hu, and Loe Feijs. 2019. Exploring Tangible Interaction and Diegetic Feedback in an AR Math Game for Children. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. ACM, Boise ID USA, 580–585. doi:10.1145/3311927.3325333

[29] Can Liu, Stephane Huot, Jonathan Diehl, Wendy Mackay, and Michel Beaudouin-Lafon. 2012. Evaluating the benefits of real-time feedback in mobile augmented reality with hand-held devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 2973–2976. doi:10.1145/2207676.2208706

[30] Sara Mandic, Rhys Tracy, and Misha Sra. 2023. ARFit: Pose-based Exercise Feedback with Mobile AR. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. ACM, Sydney NSW Australia, 1–3. doi:10.1145/3607822.3618008

[31] H. Nassaji and E. Kartchava. 2021. *The Cambridge Handbook of Corrective Feedback in Second Language Learning and Teaching*. Cambridge University Press. https://books.google.de/books?id=b7EfEAAAQBAJ

[32] Tuan Nguyen, Chanh Lam, and Paul Bruno. 2022. *Is there a national teacher shortage? A systematic examination of reports of teacher shortages in the United States*. WorkingPaper 22-631. Annenberg Institute at Brown University. doi:10.26300/76eq-hj32

[33] Ernesto Panadero, Maryam Alqassab, Javier Fernández Ruiz, and Jose Carlos Ocampo and. 2023. A systematic review on peer assessment: intrapersonal and interpersonal factors. *Assessment & Evaluation in Higher Education* 48, 8 (2023), 1053–1075. doi:10.1080/02602938.2023.2164884

[34] Reinhard Pekrun, Thomas Goetz, Wolfram Titz, and Raymond P. Perry and. 2002. Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. *Educational Psychologist* 37, 2 (2002), 91–105. doi:10.1207/S15326985EP3702_4

[35] Meike Ramon and Maria Ida Gobbini. 2018. Familiarity matters: A review on prioritized processing of personally familiar faces. *Visual Cognition* 26, 3 (2018), 179–195. doi:10.1080/13506285.2017.1405134

[36] Valentin Riemer and Claudia Schrader. 2019. Mental Model Development in Multimedia Learning: Interrelated Effects of Emotions and Self-Monitoring. *Frontiers in Psychology* Volume 10 - 2019 (2019). doi:10.3389/fpsyg.2019.00899

[37] Liam Rigby, Burkhard C. Wünsche, and Alex Shaw. 2020. piARno - An Augmented Reality Piano Tutor. In *32nd Australian Conference on Human-Computer Interaction*. ACM, Sydney NSW Australia, 481–491. doi:10.1145/3441000.3441039

[38] Rufat Rzayev, Susanne Korbely, Milena Maul, Alina Schark, Valentin Schwind, and Niels Henze. 2020. Effects of Position and Alignment of Notifications on AR Glasses during Social Interaction. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. ACM, Tallinn Estonia, 1–11. doi:10.1145/3419249.3420095

[39] Marija Sablić, Ana Mirosavljević, and Alma Škugor. 2021. Video-Based Learning (VBL)—Past, Present and Future: an Overview of the Research Published from 2008 to 2019. *Technology, Knowledge and Learning* 26, 4 (Dec. 2021), 1061–1077. doi:10.1007/s10758-020-09455-5

[40] Diana Saplacan and Jo Herstad. 2018. Fear, Feedback, Familiarity… How are These Connected?–Can familiarity as a design concept applied to digital feedback reduce fear?. In *International Conferences on Advances in Computer-Human Interactions ACHI*. ThinkMind, Technische Informationsbibliothek (TIB)-German National Library …, 171–179. http://urn.nb.no/URN:NBN:no-75127

[41] Leib Sutcher, Linda Darling-Hammond, and Desiree Carver-Thomas. 2019. Understanding teacher shortages: An analysis of teacher supply and demand in the United States. *Education Policy Analysis Archives* 27 (April 2019), 35. doi:10.14507/epaa.27.3696

[42] Xingyue Tang, Zhuang Chang, Weiping He, Mark Billinghurst, and Xiaotian Zhang. 2023. Exploring Real-time Precision Feedback for AR-assisted Manual Adjustment in Mechanical Assembly. In *29th ACM Symposium on Virtual Reality Software and Technology*. ACM, Christchurch New Zealand, 1–11. doi:10.1145/3611659.3615712

[43] Fabienne M. Van der Kleij and Anastasiya A. Lipnevich. 2021. Student Perceptions of Assessment Feedback: A Critical Scoping Review and Call for Research. *Educational Assessment, Evaluation and Accountability* 33, 2 (May 2021), 345–373. doi:10.1007/s11092-020-09331-x

[44] Martina van Heerden and Sharita Bharuthram and. 2021. Knowing me, knowing you: the effects of peer familiarity on receiving peer feedback for undergraduate student writers. *Assessment & Evaluation in Higher Education* 46, 8 (2021), 1191–1201. doi:10.1080/02602938.2020.1863910

[45] Arathi Varghese. 2020. Exploring Bi-modal Feedback in Augmented Reality. In *IndiaHCI '20: Proceedings of the 11th Indian Conference on Human-Computer Interaction*. ACM, Online India, 55–61. doi:10.1145/3429290.3429296

[46] João Vieira, Maurício Sousa, Artur Arsénio, and Joaquim Jorge. 2015. Augmented Reality for Rehabilitation Using Multimodal Feedback. In *Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques*. ACM, Lisbon Portugal, 38–41. doi:10.1145/2838944.2838954

[47] Giles Westerfield, Antonija Mitrovic, and Mark Billinghurst. 2015. Intelligent Augmented Reality Training for Motherboard Assembly. *International Journal of Artificial Intelligence in Education* 25, 1 (March 2015), 157–172. doi:10.1007/s40593-014-0032-x

[48] Naomi E. Winstone and Robert A. Nash and. 2023. Toward a cohesive psychological science of effective feedback. *Educational Psychologist* 58, 3 (2023), 111–129. doi:10.1080/00461520.2023.2224444

[49] Nick Wittig, Tobias Drey, Theresa Wettig, Jonas Auda, Marion Koelle, David Goedicke, and Stefan Schneegass. 2024. LeARn at Home: Comparing Augmented Reality and Video Conferencing Remote Tutoring. In *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia (MUM '24)*. Association for Computing Machinery, New York, NY, USA, 255–263. doi:10.1145/3701571.3701577

[50] Nick Wittig, Uwe Gruenefeld, Lukas Glaser, Mak Krvavac, Florian Rademaker, Johannes Waltmann, Donald Degraen, and Stefan Schneegass. 2024. Dropping Hints: Visual Hints for Improving Learning using Mobile Augmented Reality. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction* (Trier, Germany) *(SUI '24)*. Association for Computing Machinery, New York, NY, USA, Article 35, 3 pages. doi:10.1145/3677386.3688900

[51] Nick Wittig, Noro Schlorke, Roman Heger, Theresa Wettig, Marion Koelle, Uwe Gruenefeld, David Goedicke, Donald Degraen, Ricarda Steinmayr, and Stefan Schneegass. 2025. You ARe Correct! Comparing Augmented Reality Displays for Individual Feedback in Classroom Settings. In *Proceedings of the 24th Interaction Design and Children*. Association for Computing Machinery, New York, NY, USA, 622–635. https://doi.org/10.1145/3713043.3728864

[52] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. doi:10.1145/1978942.1978963

[53] Julia Woodward and Jaime Ruiz. 2023. Designing Textual Information in AR Headsets to Aid in Adults' and Children's Task Performance. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. ACM, Chicago IL USA, 27–39. doi:10.1145/3585088.3589373

[54] Hsin-Kai Wu, Silvia Wen-Yu Lee, Hsin-Yi Chang, and Jyh-Chong Liang. 2013. Current status, opportunities and challenges of augmented reality in education. *Computers & education* 62 (2013), 41–49. doi:10.1016/j.compedu.2012.10.024

[55] Yihong Wu, Lingyun Yu, Jie Xu, Dazhen Deng, Jiachen Wang, Xiao Xie, Hui Zhang, and Yingcai Wu. 2023. AR-Enhanced Workouts: Exploring Visual Cues for At-Home Workout Videos in AR Environment. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, San Francisco CA USA, 1–15. doi:10.1145/3586183.3606796

[56] R.B. Zajonc. 2001. Mere Exposure: A Gateway to the Subliminal. *Current Directions in Psychological Science* 10, 6 (2001), 224–228. doi:10.1111/1467-8721.00154