University of Stuttgart

Institute for Artificial Intelligence
Analytic Computing

Universitätsstraße 32
70569 Stuttgart

Bachelor Thesis

Enhancing Online Lecture Engagement Through Gaze and Emotion Feedback Integration

Frederik Horn

Study program:	B.Sc. Softwaretechnik
1. Examiner:	Prof. Dr. Steffen Staab
2. Examiner:	Prof. Dr. Mathias Niepert
Advisors:	Dr. Chandan Kumar, Bhupender Kumar Saini
start date:	15.08.2024
end date:	14.02.2025

Statement

I hereby certify that this thesis has been composed by me and is based on my own work, that I did not use any further resources than specified – in particular no references unmentioned in the reference section – and that I did not submit this thesis to another examination before. The paper submission is identical to the submitted electronic version.

	Yes	No
I agree to have this thesis published in the library.	☐	☐
I agree to have this thesis published on the Web.	☐	☐
The thesis text is available under a Creative Commons License (CC BY-SA 4.0).	☐	☐
The source code is available under a GNU General Public License (GPLv3).	☐	☐
The collected data is available under a Creative Commons License (CC BY-SA 4.0).	☐	☐

. . . .

(Place, Date) (Signature)

Note

- If you would like us to contact you for the graduation ceremony, please provide your personal E-mail address: ....
- If you would like us to send you an invite to join the AC Alumni and Members group on LinkedIn, please provide your LinkedIn ID : ....

Acknowledgements

I would like to thank everyone who contributed to the development of my bachelor's thesis. First, a huge thank you to all the participants who voluntarily invested their time—whether by completing the survey, helping with planning, or attending the lecture experiments for hours on end. Your involvement was invaluable. I am also incredibly grateful to the lecturers who used our tool during their presentations and provided honest feedback, insightful suggestions, and constructive criticism to help improve it for the future.

A special thanks goes to everyone at the University of Stuttgart and the Institute for Artificial Intelligence, especially Prof. Steffen Staab, for their support in structuring, organizing, and refining this work. Finally, I want to express my deepest appreciation to my two advisors, Chandan and Bhupender, for sitting down with me every week, spending countless hours discussing design variations, refining the survey, and guiding me through this entire process.

Without each and every one of you, this thesis would not have been possible, and I am truly grateful for your help!

Abstract

The shift towards online meetings, seminars, and lectures has rapidly gained momentum in recent years, particularly accelerated by the events surrounding the COVID-19 pandemic. However, the absence of nonverbal cues such as eye contact and gestures in virtual settings presents significant challenges for effective communication. Existing frameworks for virtual meetings often fail to adequately address this issue, making it difficult for presenters to accurately assess audience engagement and comprehension. This thesis investigates these challenges by capturing participant attention and emotion in real-time and evaluating how presenters interact with live feedback. Through the implementation and study of a feedback system that visualizes gaze and emotional data, we aimed to bridge the gap between in-person and online lecture experiences. To ensure an informed approach, we conducted a requirement analysis prior to our experiment, identifying key factors for effective feedback integration. In the experiment, we collected data through our tool by logging meeting interactions and analyzing responses from a post-experiment questionnaire. Our findings provide insights into the impact of such feedback on presenter motivation, delivery adjustments, and engagement levels, ultimately contributing to the improvement of educational quality in virtual environments.

Zusammenfassung

Die Verlagerung hin zu Online-Meetings, Seminaren und Vorlesungen hat in den letzten Jahren rasant an Bedeutung gewonnen, insbesondere beschleunigt durch die Ereignisse rund um die COVID-19-Pandemie. Allerdings stellt das Fehlen nonverbaler Hinweise wie Blickkontakt und Gesten in virtuellen Umgebungen eine erhebliche Herausforderung für eine effektive Kommunikation dar. Bestehende Frameworks für virtuelle Meetings greifen dieses Problem oft unzureichend auf, wodurch es für Vortragende schwierig wird, das Engagement und das Verständnis ihres Publikums genau einzuschätzen. Diese Arbeit untersucht diese Herausforderungen, indem sie die Aufmerksamkeit und Emotionen der Teilnehmenden in Echtzeit erfasst und analysiert, wie Vortragende mit Live-Feedback interagieren. Durch die Implementierung und Erforschung eines Feedback-Systems, das Blick- und Emotionsdaten visualisiert, soll die Lücke zwischen Präsenz- und Online-Vorlesungen verringert werden. Um eine fundierte Herangehensweise zu gewährleisten, führten wir vor dem Experiment eine Anforderungsanalyse durch, um zentrale Faktoren für eine effektive Feedback-Integration zu identifizieren. Während des Experiments sammelten wir Daten durch unser Tool, indem wir Interaktionen während der Meetings aufzeichneten und die Antworten aus einem nachgelagerten Fragebogen auswerteten. Unsere Ergebnisse liefern Einblicke in die Auswirkungen eines solchen Feedbacks auf die Motivation der Vortragenden, deren Anpassungen in der Vortragsweise und das Engagement-Niveau, wodurch letztendlich die Qualität der Lehre in virtuellen Umgebungen verbessert werden kann.

Contents

1 Introduction

Imagine yourself standing in front of an audience and delivering a presentation. You have likely encountered those moments when you met confused or puzzled expressions in the crowd. The ability to gauge immediate reactions from the audience is crucial for effective communication. Presenters rely heavily on nonverbal cues such as emotions and eye contact to assess the comprehension and engagement of their listeners. They adjust their delivery based on these cues, deciding whether to elaborate or clarify certain points.

However, in the realm of online meetings and lectures, this vital aspect is often lost. Without the ability to see the faces of attendees, presenters are deprived of instant feedback. Instead of observing attention, facial expressions, and gestures, they are met with gray squares or numerous tiny video boxes that make it nearly impossible to discern emotions or reactions. The absence of nonverbal communication channels like eye contact and gestures significantly hampers the presenter's ability to effectively engage with their audience and tailor their delivery accordingly.

This thesis aimed to bridge the gap in nonverbal feedback between in-person and online lecture environments, ultimately enriching the educational experience in virtual settings. Specifically, the focus was on formal presentations and lectures involving a single presenter, enabling targeted adjustments to their delivery based on feedback received from all participants.

Building on the foundational work of a previous project conducted at the University of Stuttgart [1], which introduced a framework enabling gaze and emotion detection in online meetings, this research investigated how the integration of gaze and emotion (audience engagement) in online lectures could enhance the presenter's experience and how effectively it could replicate the natural feedback environment of a traditional classroom. The core focus of this thesis was to address the following key questions:

- **How does the presence or absence of nonverbal feedback (such as attention and emotion) in online lectures impact the flow, effectiveness, and motivation of presenters?** To understand how the absence of non-verbal feedback, such as attention and emotional responses, influences the flow and motivation of presenters during online lectures, we conducted a requirement analysis prior to the experiment. This included an online survey targeting presenters who frequently engage in online lectures and presentations. The survey gathered insights into their expectations from a non-verbal feedback framework, as well as concerns related to privacy and potential distractions.

- **What are the most effective design approaches for visualizing nonverbal feedback (such as attention and emotion) in online meetings, considering factors like information granularity, potential distractions, misinterpretations, and privacy?** Based on the requirement analysis, we designed and implemented different visualization strategies that provide real-time feedback on participants' attention levels and emotional responses. The goal was to develop intuitive and effective representations that minimize distractions, reduce the risk of misinterpretation, and respect participant privacy. Several variations of visualization techniques were created, focusing on different levels of data aggregation and information granularity.

- **How do different visualization designs for nonverbal feedback in online lectures influence the perceived usefulness, acceptance, and feasibility by presenters, and how closely do they replicate the nonverbal feedback typically observed in in-person meetings?** To assess the effectiveness, acceptance, and feasibility of the developed visualizations, we conducted an experiment in which 8 presenters led online sessions while using different feedback designs. Each presenter conducted a session with 7-11 participants, allowing us to compare how different visualization approaches influenced their experience. During the experiment, we collected real-time meeting data through our tool and followed up with a post-experiment questionnaire to gain deeper insights into how presenters perceived and utilized the feedback. The collected responses enabled a comparative analysis of the designs, focusing on their effectiveness in replicating non-verbal feedback from in-person meetings and enhancing the overall presentation experience.

The outcomes of this thesis include an advanced virtual meeting framework tailored for presenters, integrating eye gaze and emotion detection technologies to enhance communication and engagement. The findings provide insights into the effectiveness of real-time feedback based on eye gaze and emotional cues, offering recommendations for optimizing their integration into educational virtual meeting platforms. Furthermore, the results suggest implications for future virtual learning environments and instructional practices.

2 Related work

This work extends the project described in Section 2.1. Section 2.2 briefly discusses related efforts on incorporating gaze into virtual meetings, while Section 2.3 examines alternative modalities for studying engagement.

2.1 Detecting and Analyzing User Engagement in Virtual Meetings

This thesis builds on previous work [1, 2], which introduced a framework for recording and sharing nonverbal behavioral cues, such as gaze and emotion, among participants within a web-based virtual meeting platform. The study proposed two visualization schemes to convey gaze information and evaluated them through a comprehensive remote study simulating real-world meeting scenarios. Their framework also introduced a novel concept wherein users can share their attention information even when their camera feed is deactivated. However, the study did not explore emotion analysis in depth, nor its integration and synchronization with gaze, or its visualization to participants — gaps that this thesis investigates.

2.2 Eye gaze in online meetings and lectures

Considerable research has focused on gaze correction methods in video conferencing, helping individuals maintain eye contact with their fellow participants [3]. True-view [4] aims to correct gaze alignment for two meeting participants, creating the illusion of close proximity. D'Angelo and Gergle [5] used remote eye tracking to share gaze information, investigating how remote pairs utilize graphical representations of each other's eye gaze during tightly-coupled collaborative tasks. Burch [6] employed mobile eye tracking devices to continuously monitor students' eye movements as they paid visual attention to lecture slides. Langner et al. [7] conducted studies with student groups working remotely on course assignments, utilizing Tobii eye trackers to explore how eye-based joint attention enhances efficient collaboration.

2.3 Other Modalities to study engagement in online meetings

In addition to gaze awareness, prior studies have delved into examining the effects and utilization of various other modalities within the realm of online meetings. Research has also examined the role of spatial cues [8], such as the positioning of individuals within virtual spaces to convey information or establish social dynamics. Proximity cues [9] have also been scrutinized, exploring how the perception of personal space and distance influences communication and interaction in virtual environments.

Sharma et al. [10] examined the significance of emotion feedback for presenters, calculating three levels of engagement based on emotional data. Their research underscored the direct relationship between engagement and comprehension of meeting content. However, their study was conducted in a one-to-one setting without providing live feedback to the presenter. In contrast, this thesis implemented live feedback, combined with gaze tracking, within a group meeting scenario.

3 Requirement Analysis

Before designing the framework with its various functions and visual cues, we first assessed the relevance and necessity of nonverbal feedback in virtual meetings.

3.1 Survey Motivation and Objectives

The primary goal was to identify the types of feedback that participants felt were lacking in online meetings and to understand how this absence affected presenters' delivery and motivation. Additionally, we explored specific scenarios in lectures or meetings that had the most impact on presenters. For instance, one might assume that positive emotions and high attention levels create a sense of acknowledgment and encouragement for presenters. In contrast, negative emotions or consistently low attention levels could lead to feelings of disengagement, potentially lowering the presenter's motivation.

3.2 Survey Design and Key Questions

The survey began with a brief introduction, followed by questions regarding the demographic background of participants. Chapter 3 of the survey focused on gathering insights into participants' experiences in online group meetings. Questions addressed the frequency of participation, the frequency of taking on the role of a presenter, and the average number of cameras enabled during meetings.

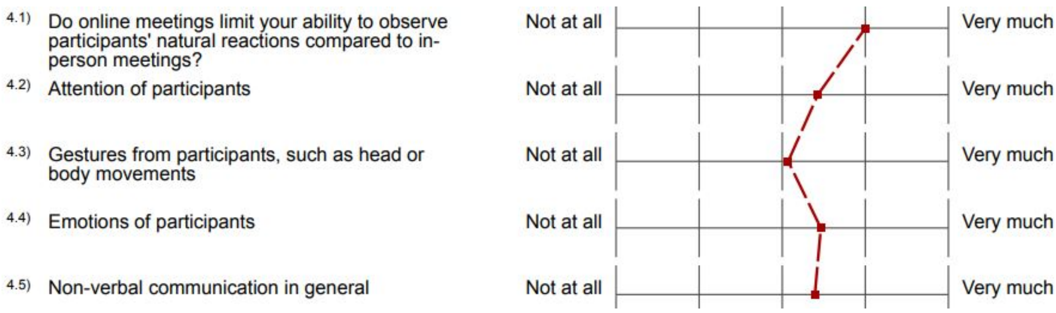

Figure 1: Impact of missing nonverbal cues — Responses to survey questions 4.2 - 4.5 regarding how the lack of emotional and attentional feedback affects presenters.

In Chapter 4, the survey shifted towards assessing the role of nonverbal feedback and its impact on presenters in online meetings. As shown in Figure 1, these questions aimed to evaluate how the lack of nonverbal cues influences the experience of presenters. A particularly striking result emerged from question 4.1, which asked about the ability to observe natural reactions from participants. Not a single respondent selected "Not at all" (1), and more than 78% rated their ability to observe reactions at (4) or (5). This clearly indicates a significant difference between in-person and virtual meetings in terms of assessing audience engagement and comprehension.

Questions 4.2 to 4.5 explored the extent to which the absence of various nonverbal cues affected presenters. The results showed that emotions and attention levels of participants had a stronger impact on presenters compared to gestures such as head nods or body movements. However, the general lack of nonverbal communication in online meetings was perceived as impactful overall. One possible explanation for the relatively lower influence of missing gestures could be the existing features of most virtual meeting platforms, which allow participants to raise their hands or react with visual indicators like a thumbs-up. Based on personal observations, these features are widely used and effectively compensate for the absence of physical gestures.

Chapter 5 of the survey focused on different meeting scenarios, each representing a combination of audience attention levels (Low, Medium, High) and emotional states (Negative, Neutral, Positive). Participants were asked to rank these scenarios based on the perceived impact on the presenter.

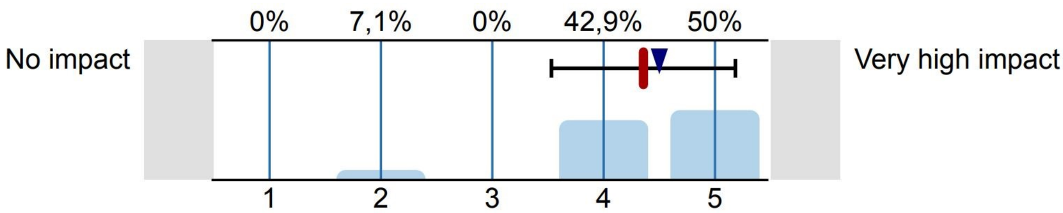

Figure 2: The most impactful meeting scenario: **Low Attention** and **Negative Emotions** — 50% of participants rated this as having a very high impact.

The scenario with the highest reported impact was one in which audience attention was low, and emotions were predominantly negative. As illustrated in Figure 2, 50% of participants indicated that this situation would have a very high impact on them. In general, scenarios featuring negative emotions were ranked among the most impactful. Scenarios dominated by neutral emotions were rated lower in impact, except when paired with low attention levels. Positive emotion scenarios were generally perceived as more influential, particularly when accompanied by high attention levels.

The significant impact of positive feedback, as in figure 3 can be attributed to fundamental human social and psychological responses. Research in social psychology suggests that smiling is often reinforcing social bonding and group cohesion [11]. Additionally, shared laughter fosters a sense of connection and collective engagement [12]. A positive environment has been shown to enhance motivation and reduce stress, as individuals tend to feel more relaxed and encouraged in supportive social settings [13]. The psychological benefits of positive reinforcement are well-documented, making its influence both logical and expected.

Similar to the questions in Chapter 4, we asked participants whether they would

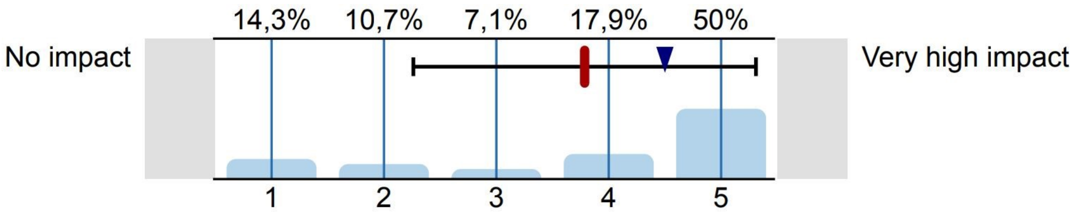

Figure 3: The second most impactful meeting scenario: **High Attention** and **Positive Emotions** — Participants found this scenario highly motivating.

find emotion and attention feedback beneficial when presenting in virtual meetings. The results were unanimous—every participant found such feedback useful in some way for assessing their audience's emotional state and attention levels.

When asked about the type of feedback they would like to see in the future, responses varied, indicating that preferences for feedback are highly individual, as shown in Figure 4. However, some general trends emerged. Real-time alerts seemed to be of interest to most participants, while live suggestions were perceived as potentially overwhelming or not particularly useful during a presentation. A clear preference emerged for post-meeting analytics: over 75% of participants agreed that a summarized analysis of emotions and gaze data, presented as statistical insights after the meeting, would be valuable.

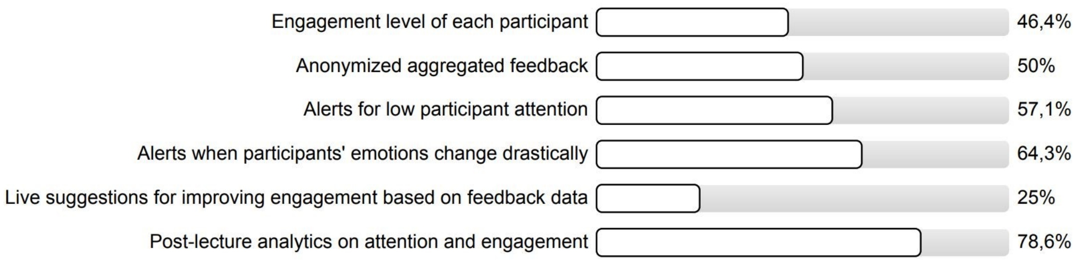

Figure 4: Preferred types of attention and emotional feedback features — Survey results showing what presenters would like to see implemented in online meeting platforms.

3.3 Key Findings in Open-Ended Questions

To gain deeper insights, we also collected qualitative data through open-ended survey questions. Several key themes emerged from participants' responses:

- **I can see if people can follow the technical level of my presentation. I can adapt the technical depth and explanation, especially with people whose prior knowledge I am unfamiliar with.** This response highlights the challenge of presenting to an audience with mixed or unknown levels of expertise.

In such cases, real-time feedback on comprehension could be crucial for adapting explanations dynamically.

- **In smaller groups with cameras on, you can recognize emotions in people's faces. In meetings with more than 10 participants, it is not possible.** This emphasizes the need for different feedback designs based on group size. While small meetings allow for natural recognition of facial expressions, larger meetings require automated support to assess audience reactions effectively.
- **The reactions you see can influence your confidence during a presentation.** This highlights an important psychological aspect—negative feedback may impact the presenter's confidence, potentially leading to a downward spiral where reduced confidence results in a weaker presentation, which in turn generates more negative feedback. Careful design considerations are needed to ensure that feedback is constructive rather than discouraging.
- **I think this data is interesting for professional presenters but less relevant for students or daily meetings.** This response suggests that real-time audience analytics may be more useful for experienced presenters who can adjust dynamically, whereas those less accustomed to public speaking may find the information overwhelming. Customization options could be necessary to tailor the system to different user needs.
- **Not every person is equally relevant over the course of the meeting. Oftentimes, it is acceptable for some listeners to focus on other tasks to remain productive overall.** This raises an essential point: inattentiveness should not automatically be treated as a negative factor. The framework should filter and present only relevant insights, ensuring that presenters receive meaningful and actionable feedback rather than an overwhelming amount of raw data.

4 Design and Implementation

4.1 Designing the Framework

The design process was structured into three key phases: Brainstorming, Narrowing Down and Decision, and Optimizing. This approach ensured a systematic exploration of potential solutions, allowing for a broad generation of ideas followed by a focused selection and refinement process. The brainstorming phase aimed to explore a wide range of design possibilities and critically discuss different approaches. In the narrowing down and decision phase, the most promising concepts were identified based on their alignment with the project goals and insights from the survey analysis. Finally, the optimization phase focused on refining the selected designs to enhance usability, address user feedback, and ensure the final implementation adhered to the established priorities.

4.1.1 Brainstorming

During this initial phase, the primary goal was to generate a wide range of visualization concepts without being constrained by specific technical limitations. We explored various ways to represent emotion and gaze feedback, considering both conventional and more abstract approaches. Beyond standard bars and charts, we experimented with creative designs that could intuitively convey information at a glance. Some sketches focused on highly detailed, feature-rich representations, while others prioritized simplicity and clarity. The complexity of the final design would ultimately determine how the underlying data would need to be processed — whether through real-time aggregation, averaging, or condensing information to ensure usability and readability.

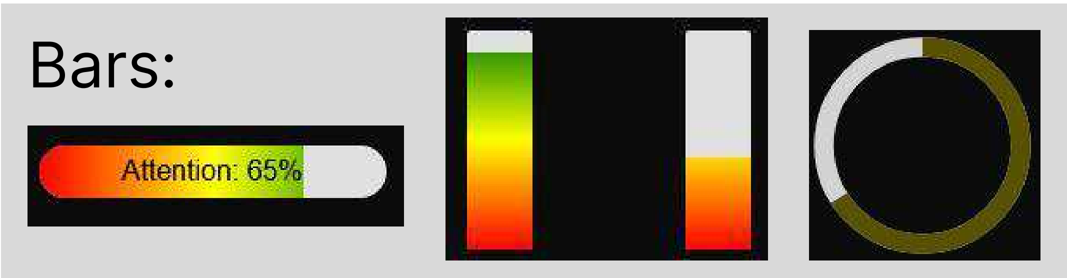

Figure 5: Various bar designs for visualizing attention levels and emotional states, with different levels of aggregation.

Another intuitive approach for visualization was the use of various types of bars to represent levels of attention or emotional states. These bars could be color-coded to distinguish between positive, neutral, and negative emotions, ensuring quick and effortless interpretation. To further streamline the design, a textual representation could be integrated alongside the bars, allowing for a more compact display. The most condensed version of this approach is illustrated on the right in Figure 5, where a radial progress bar is used in combination with color as the only indicators. While this design prioritizes simplicity and readability, it requires significant data aggregation, inevitably leading to a loss of detail.

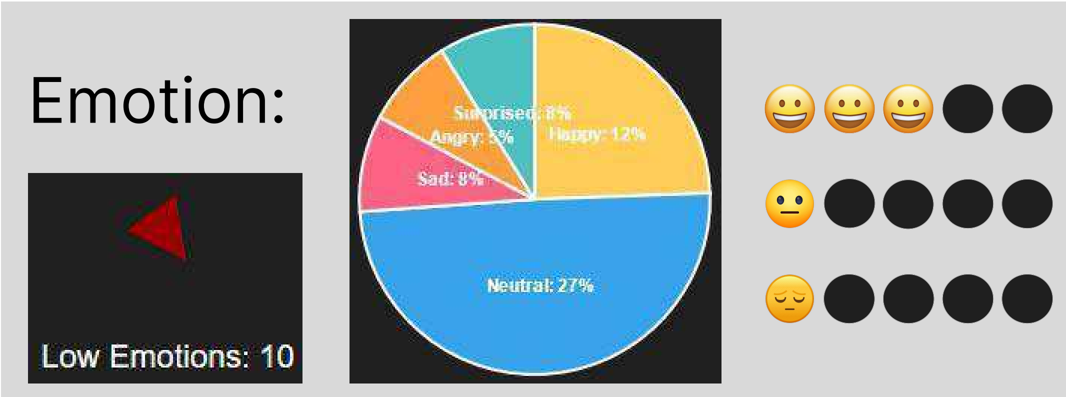

Figure 6: Different visualization approaches for emotional feedback, including trend arrows, pie charts, and emoji-based representations.

As the brainstorming phase progressed, it became increasingly clear that the most challenging aspect of the design process would be developing an effective and meaningful way to provide real-time feedback on emotional states in the meeting. Consequently, a major focus of Phase 1 was on exploring different methods of emotion visualization. If the primary interest was in identifying overall emotional trends rather than individual emotions, a simple up or down arrow could be used to indicate shifts in sentiment, complemented by color coding and brief explanatory text. However, for a more detailed approach, one could attempt to display all individual emotions and their aggregated levels among participants. Prior research [1], which this thesis builds upon, introduced a method of using dots of varying sizes to represent different levels of specific emotions. Another fine-grained alternative involved the use of a pie chart displaying the proportional distribution of emotions among participants. However, both of these approaches pose a significant drawback: they require considerable cognitive effort to interpret, which could distract presenters, causing disruptions or loss of focus during their delivery.

To simplify the process of interpreting emotions in real time, an alternative design choice was to group emotions into broader categories — positive, neutral, and negative — rather than displaying each emotion separately. One particularly effective way to implement this idea was through the use of emojis, which naturally convey emotional meaning without requiring textual labels. By replacing individual emotion names with expressive icons and consolidating emotional data into three key categories, the design became more intuitive. A straightforward yet informative visualization was then developed, consisting of three bars displaying the levels of positive, neutral, and negative emotions, accompanied by corresponding emojis.

This approach significantly reduced cognitive load while maintaining the essential aspects of emotional feedback.

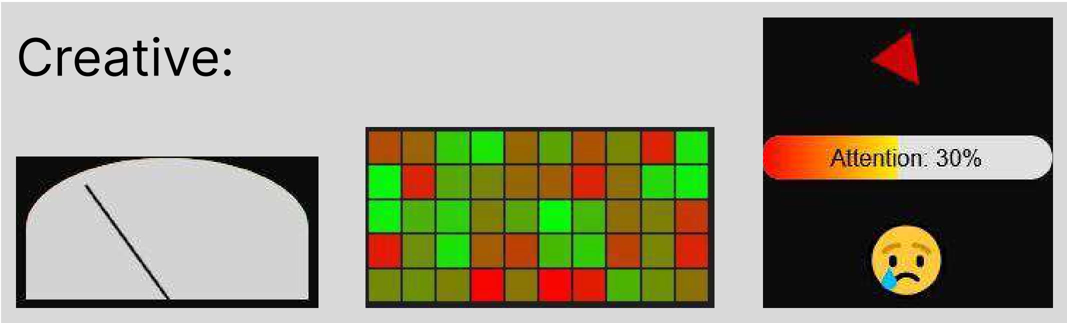

Figure 7: Exploring creative and abstract ways of visualizing engagement and emotional states using shape, color, and dynamic elements.

Lastly, we explored more creative and abstract ways of visualizing data, aiming to break away from conventional designs already used in multiple contexts. The goal was to experiment with unique approaches that could offer a fresh perspective on presenting feedback. One particularly intriguing concept was the heatmap shown in the middle of Figure 7. Although it initially appeared somewhat chaotic, the idea resonated with the team because each tile could potentially represent an individual participant in the meeting. This design had the potential to provide an intuitive and engaging way to visualize audience engagement and emotional states at a glance.

Once all initial designs were sketched and discussed in terms of usability and effectiveness, we moved on to the next phase: refining and improving the most promising concepts. At this stage, the focus shifted toward enhancing the clarity and intuitiveness of selected designs while also brainstorming additional ideas that could integrate both gaze and emotion data. The ultimate objective was to create a feedback system that closely mimics the experience of standing in front of a live audience. While elements like trend arrows or pie charts might be useful for post-lecture analysis, they lacked the immediacy and natural feel necessary for real-time feedback. For this reason, we prioritized designs that conveyed information in an intuitive and unobtrusive way, allowing presenters to stay focused while still gaining meaningful insights into their audience's reactions.

4.1.2 Narrowing Down

Before refining the existing designs, we first explored ways to deliver feedback in a subtle and unobtrusive manner by integrating visual elements into the presenter's screen. The goal was to ensure that the feedback remained in the presenter's peripheral vision rather than demanding active attention. This way, the information would be naturally perceptible—allowing the presenter to stay fully engaged in their talk while still maintaining a general awareness of the audience's overall mood and attentiveness. By designing feedback to blend seamlessly into the presenter's workflow, we aimed to create an experience that closely mirrors in-person presentations, where speakers intuitively pick up on audience reactions without consciously analyzing each individual's expression.

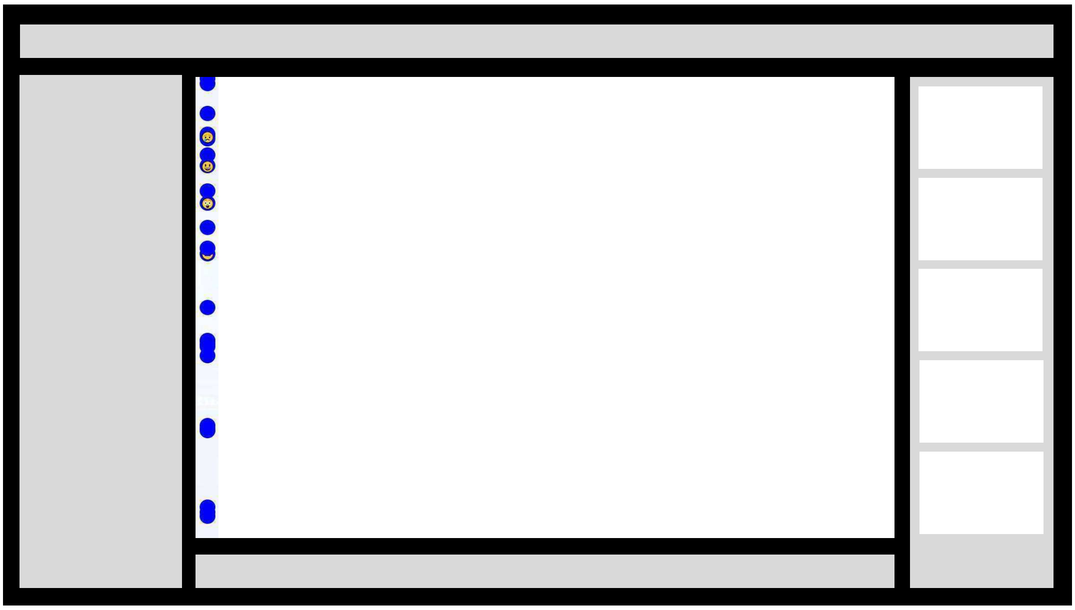

Figure 8: Participant circles shifting position based on attention and emotional state — early concept with high movement and unclear structure.

The first design, shown in Figure 8, visualized each participant as a circle that moved based on their level of attentiveness. If a participant remained focused, their circle would slide upwards; if they became distracted, their circle would move downwards. Additionally, sudden emotional changes were represented by an emoji appearing inside the respective circle. While the concept initially seemed promising, it quickly became apparent that the continuous movement and lack of clear structure made the design overwhelming. The high level of visual activity required too much attention from the presenter, making it difficult to extract relevant information efficiently.

Figure 9: Participants appearing or disappearing at the bottom of the presenter's screen based on attention level, mimicking head movement in a physical room.

To address these issues, we explored an alternative approach that aimed to mimic the experience of standing in front of a real audience. In this design, shown in Figure 9, the participant circles were positioned at the bottom of the presenter's screen, sliding up or down to simulate head movement in a physical classroom. If a participant was paying attention, their full circle was visible, while a distracted participant's circle would move downward, partially disappearing—similar to someone looking away or down at their notes. Emojis would still appear to highlight sudden emotional changes. This version felt more intuitive compared to the previous iteration, as it aligned with real-world behavior. However, the feedback elements remained quite small, making it difficult to quickly interpret the emotional reactions of all participants at a glance. Additionally, the presence of multiple moving elements still introduced a certain level of distraction.

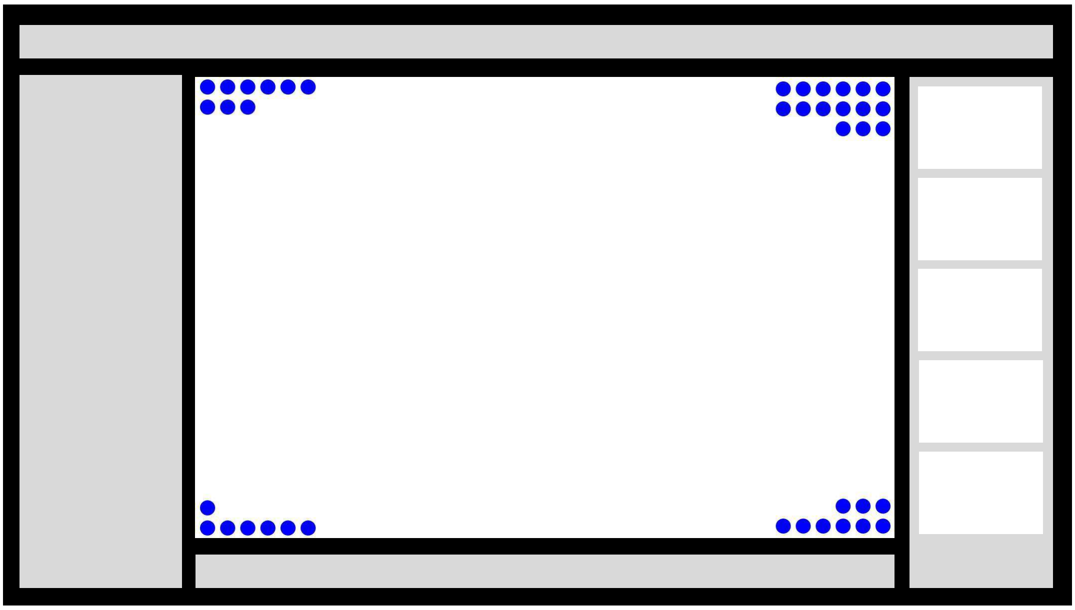

Figure 10: Concept of using screen corners to represent different emotional states or attention levels, keeping information in the presenter's peripheral vision.

To further refine the approach and make feedback even more effortless to perceive, we experimented with an idea that leveraged the presenter's peripheral vision. Instead of displaying individual circles in a fixed location, we assigned different corners of the screen to specific audience states. For instance, one corner could represent positive engagement, another could indicate inattentiveness, and participants' circles would appear in the corresponding corner based on their current state. A sketch of this idea is shown in Figure 10. This method minimized movement and allowed the presenter to get an overall impression of the audience's mood without consciously focusing on individual participants. However, the downside was that presenters would need to learn the meaning of each corner, which could introduce a learning curve. The first few uses of this system might feel unfamiliar or even confusing, potentially causing more distraction rather than reducing it.

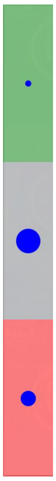

Figure 11: Refining the left-side visualization by introducing predefined areas to reduce movement and improve clarity.

To address the issues of excessive movement and overlapping elements in the designs shown in Figure 8 and Figure 9, we refined the concepts by introducing predefined areas for different feedback categories.

For the left-side design, one of the primary problems was the continuous motion of individual circles, making it difficult to track changes in emotions over time. Additionally, the emojis representing sudden emotional shifts were too small and could become overlapped by other moving elements. To mitigate this, we introduced the idea of three mainly static circles within the designated area. These circles would grow or shrink dynamically based on the number of participants displaying positive, negative, or neutral emotions. This adjustment significantly reduced movement while providing a clearer representation of the audience's overall emotional state.

However, a key limitation remained — this design only conveyed emotions and did not account for gaze. Without gaze information, it was impossible for the presenter to differentiate between negative emotions expressed by actively engaged listeners and those from participants who were distracted or looking away from the presentation. A more effective approach was needed to incorporate both gaze and emotion feedback.

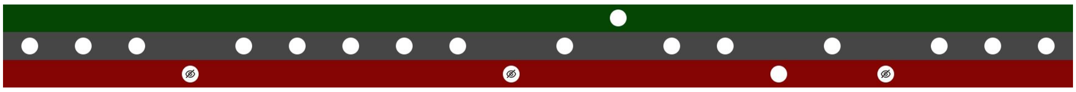

Figure 12: Splitting up the bottom feedback area while integrating icons for distracted participants.

Since the left-side approach did not seem ideal for integrating both aspects, we revisited the bottom-positioned design. As shown in Figure 12, we adapted it by dividing the area into predefined sections while still maintaining individual circles to represent each participant. This preserved the ability to include icons within the circles, indicating whether a participant was focused on the presentation or not. While this method successfully combined emotion and gaze feedback, it came at the cost of losing the initial concept of mimicking head movements through the appearance and disappearance of circles. Since circles now had to move to their respective designated areas based on emotional state and attentiveness, the natural feel of shifting gaze was diminished.

After multiple discussions, we concluded that a peripheral feedback design should minimize cognitive load and avoid excessive visual clutter. As a result, we decided to shift towards a more subtle yet effective solution — an adaptive border around the presenter's screen that provides a high-level indication of the audience's overall mood through color changes. This design significantly reduces distractions while still ensuring that the presenter remains aware of emotional trends. The border functions as a passive, non-intrusive alert system, subtly drawing attention when a significant change in mood occurs.

Based on the findings from the survey, alerts were identified as an important feature for many participants. To complement this high-level feedback, we decided that the presenter should also have access to a secondary, more detailed feedback tool. This tool could be actively consulted when the border indicates a relevant shift in audience engagement, allowing the presenter to investigate the change more thoroughly if needed.

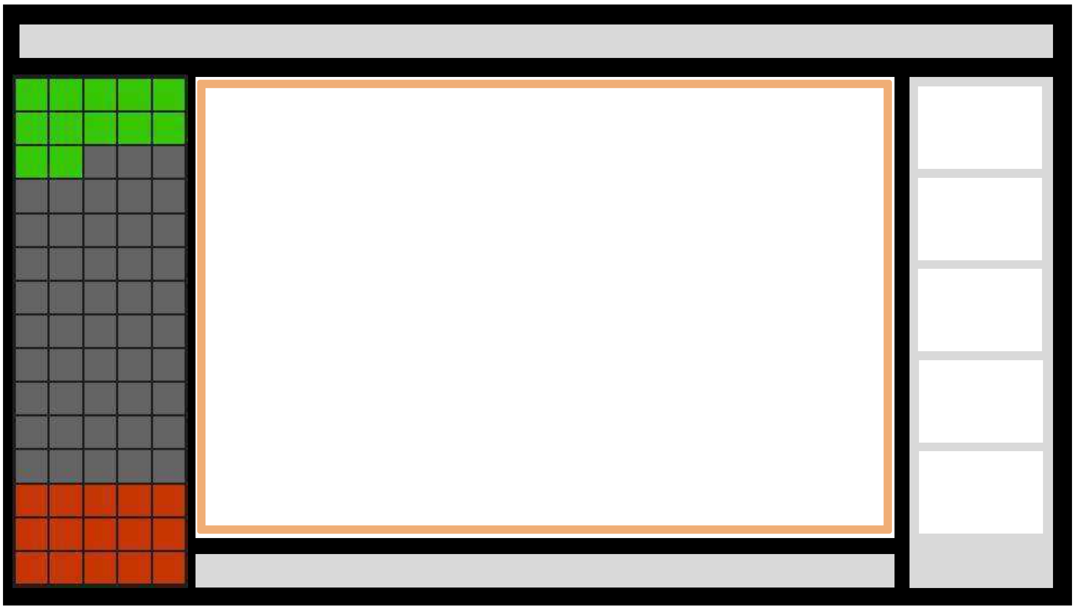

Figure 13: The final border design — showing overall audience mood through color, with an additional tool for detailed feedback when needed.

4.1.3 Optimizing

After finalizing the decision to use a border as a subtle indicator and providing additional feedback when needed, we shifted our focus to optimizing these supplementary features to ensure the highest possible usability. One of the primary elements we refined was the emoji-based emotion bars.

Initially, the design consisted of rows of circles filling up with emojis that represented the levels of positive, neutral, and negative emotions. However, this visualization lacked differentiation for gaze-related feedback. The most straightforward way to integrate both emotional states and attentiveness was to introduce a split

Figure 14: Optimization of emoji bars, exploring different layouts and color coding for clearer emotional feedback.

system: two sets of three rows — one for attentive participants and one for inattentive participants. This revised layout, illustrated on the left in Figure 14, allowed us to maintain the clarity of emotional trends while also indicating if participants were disengaged. This approach was both intuitive and effective, so we decided to proceed with it.

One major concern that arose was the effectiveness of peripheral perception. Since most emojis are predominantly yellow, the differences between positive, neutral, and negative emotions were not immediately distinguishable at a glance. To improve quick readability, we explored alternative designs that relied more on color coding rather than emoji shapes. The most intuitive choice was to assign distinct colors to each emotional category: green for positive emotions, gray for neutral, and red for negative.

Our first attempt at implementing this involved replacing the circles with squares that filled with the respective colors. However, this raised concerns about the granularity of feedback — since the design only contained five squares per row, the resolution of the feedback was relatively low. This meant that even small shifts in emotional distribution could result in noticeable jumps, such as the difference between one out of five squares being filled versus none at all.

To increase the level of detail, we experimented with vertical bar structures composed of smaller squares that filled progressively as emotional intensity increased. While this improved the visualization slightly, it still lacked a smooth, continuous representation of emotional trends. Ultimately, we settled on a more refined approach: instead of individual squares, we reverted to a single continuous bar for each row. The bar's length would dynamically represent the emotional distribution, and its color would provide an immediate, at-a-glance indication of the dominant emotion. To further enhance usability, we retained the emoji labels next to the bars, reinforcing their meaning without requiring detailed reading.

With the emotion bars optimized, we moved on to refining the heatmap design. While the heatmap effectively displayed emotional states through color-coded tiles, it still lacked an integration of gaze feedback, which was essential for providing a comprehensive overview of audience engagement.

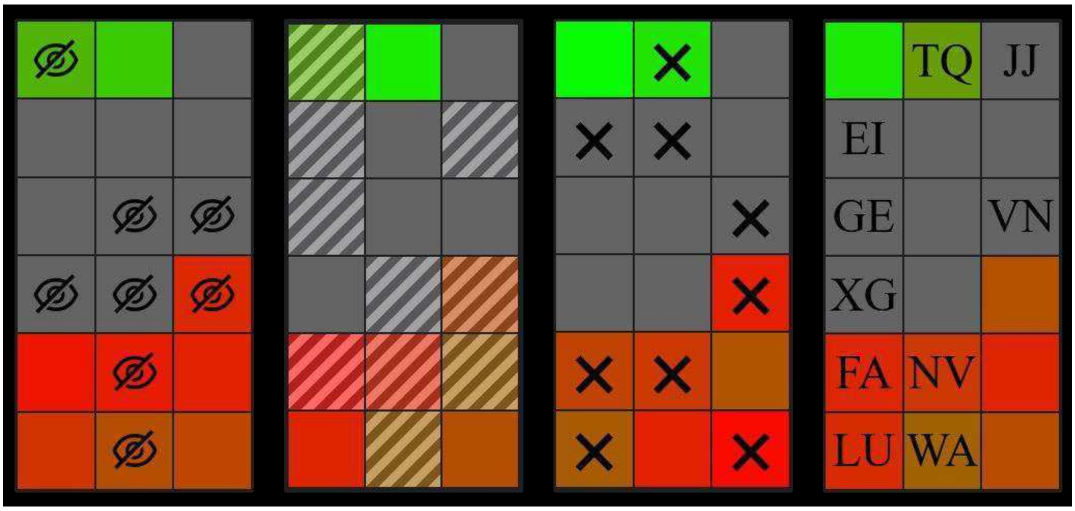

Figure 15: Testing different ways to integrate gaze feedback into the heatmap by adding content inside the cells.

Our first approach aimed to incorporate gaze feedback by adding visual elements inside each tile to indicate whether the corresponding participant was attentive or distracted. We explored several variations of this idea: initially, we tested inserting symbols or icons within the squares whenever a participant became inattentive. However, during early evaluations, we noticed that drawing attention to distracted participants was counterproductive. Instead of allowing the presenter to focus on the engaged audience, the additional symbols unintentionally highlighted disengaged individuals, leading to potential distractions. To counteract this effect, we reversed the logic: rather than adding something to indicate inattention, we experimented with showing the initials of attentive participants while making them disappear when they became inattentive. This subtly encouraged the presenter to focus on the present audience rather than those who had momentarily disengaged.

Our second approach moved away from adding content inside the cells and instead focused on modifying the tiles themselves in terms of position, size, or opacity. The first idea within this approach was to divide the heatmap into two sections—one for attentive participants and one for inattentive ones—mirroring the

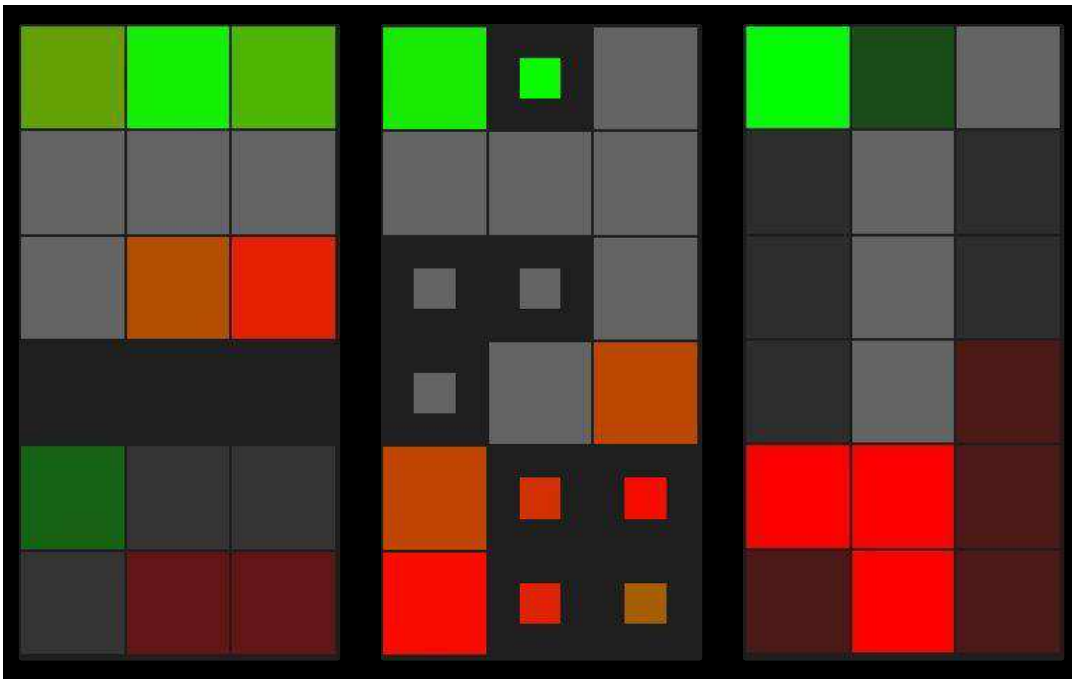

Figure 16: Exploring different approaches for gaze visualization in the heatmap, including predefined areas, size variation, and opacity changes.

concept used in the emoji bars. However, this introduced excessive movement, as tiles would have to shift between areas whenever a participant's engagement changed, leading to potential distractions for the presenter. A similar issue arose when experimenting with dynamic tile sizes, where inattentive participants' tiles would shrink. While this concept provided an intuitive visual cue, the required size difference had to be drastic to be noticeable at a glance, which once again resulted in unnecessary movement.

The final approach was to adjust the opacity of the tiles based on gaze behavior — making tiles fade when a participant became inattentive. This method showed promise due to its subtlety and minimal distractions. However, in practice, it proved difficult to interpret effectively. The reduced opacity weakened the visibility of the emotional colors, making it challenging to distinguish between different engagement states quickly, especially when relying on peripheral vision.

AAfter considering all approaches and engaging in extended discussions, we ultimately decided to use the design in which participants' initials appeared when they were attentive and faded when they were not. This approach closely resembles the design shown on the right in Figure 15, with the key difference that the squares are not sorted by emotional state, but instead, each tile is randomly assigned to a participant at the beginning. This results in a slightly less structured heatmap while ensuring that the initials remain static, even when a participant's emotional state changes.

While this approach raised minor concerns regarding privacy and ethical implications, it struck the right balance between detail and subtlety. This decision aligned with our overall strategy: the heatmap would serve as the most fine-granular feedback element, complementing the highly aggregated border and the moderately condensed emoji bars.

4.2 Implementation

Implementing most of the designs did not pose significant technical challenges, as many had already been prototyped using HTML, CSS, and JavaScript with example data to simulate movement and dynamic updates. The primary difference in the final implementation was the integration of real-time data and appropriate calculations for each feature of the framework. Each visualization required a distinct approach to processing and displaying data.

The border color, representing the overall emotional state of attentive participants, was determined by aggregating the positive, neutral, and negative emotion levels and computing an average index. Based on this index, the default neutral gray (corresponding to 100% neutral emotions) was adjusted towards green for positive emotions or red for negative emotions. This provided a subtle but effective way to reflect the audience's overall mood without overwhelming the presenter.

The emoji bars followed a similar data aggregation method but were designed to distinguish between attentive and inattentive participants. Instead of modifying the border color, this feature adjusted the height of bars representing the proportion of each emotional category.

To prevent excessive flickering due to minor emotional shifts, a threshold-based update mechanism was implemented. Changes were only reflected when the emotional distribution fluctuated by more than 10%. For example, if the proportion of neutral participants increased from 70% to 72%, the bars remained unchanged, but a shift from 71% to 83% would trigger a visible adjustment. This approach ensured stability in the visualization while still capturing significant emotional trends.

The heatmap implementation was relatively straightforward. Each participant was assigned a unique ID linked to a specific tile. The color of each tile was determined using the same calculation as the border, ensuring consistency in emotional representation across different visualizations. Additionally, to incorporate gaze feedback, participant initials were displayed within their assigned tile when they were attentive and removed when they were distracted.

4.2.1 Data Exchange and Server Infrastructure

The exchange of calculated data between the presenter and participants is managed via a Node.js server hosted on a cloud provider. The server collects and stores participant data, forwarding only the necessary information to the presenter's feedback tool. To facilitate further evaluation, all non-personal data is retained for later analysis.

To ensure precise data collection, the presenter's tool signals the server when the meeting officially starts. This prevents irrelevant data from being recorded during the calibration phase when participants are adjusting their frameworks.

4.2.2 Building Process

Since the participant framework is implemented as a browser extension with overlaying elements and the server functionalities are executed as a JavaScript file on a Node.js server, the primary development focus was on creating the presenter's feedback application. Unlike the browser-based participant tool, which overlays elements onto the screen, the presenter's tool needed to function alongside full-screen presentation software.

After evaluating various approaches for overlaying UI elements on full-screen applications, we decided to develop the presenter's tool as an Electron application. Electron provides a flexible environment that allows for HTML, CSS, and JavaScript development, minimizing the need for extensive adjustments.

The primary implementation involved configuring the application settings to specify:

- **Window behavior:** Full-screen mode, adjustable positioning, or resizable layout.
- **Click-through functionality:** Certain UI elements ignore mouse interactions to prevent interference with the presenter's primary application.

During the testing phase, we optimized the tool for variable screen sizes to ensure compatibility across different display setups. However, feedback from study participants revealed that a significant number of lecturers used macOS, requiring an additional version of the application. This led to further development and adaptation to ensure cross-platform functionality.

4.2.3 Challenges and macOS Compatibility

The most significant challenge during implementation was adapting the application for macOS. Unlike Windows, macOS applications must be built specifically for the target machine's architecture, requiring access to macOS hardware for development, testing, and compilation. Coordinating this workflow added complexity to the process.

Additionally, macOS has a more restrictive application packaging and distribution system. To work around this limitation, instead of distributing a pre-built application, we provided raw application files that lecturers could install and run manually. Participants using macOS were instructed to execute the following commands in the terminal:

- **npm install –save-dev electron**
- followed by **npm start**

This approach ensured compatibility across different macOS configurations while allowing lecturers to independently set up the tool on their machines.

5 Evaluation

The primary goal of this evaluation is to assess the effectiveness, usability, and presenter experience of the gaze + emotion feedback system. Specifically, it aims to:

- Compare the two visualization designs (individual feedback vs. aggregated feedback) in terms of presenter experience.
- Evaluate the presenter's perception of feedback usefulness, distraction, and impact on audience engagement.
- Compare the two designs against the conventional setup (no feedback) to understand perceived improvements.
- Gather qualitative feedback on how the tool could be improved for future iterations.

5.1 Study Design

We conducted eight experiments, each consisting of a 12-minute lecture followed by a questionnaire completed by the presenters. Prior to participation, presenters were required to download and test the tool according to their device's operating system to ensure functionality. The entire process, including setup, lecture, and questionnaire completion, took approximately 30 minutes per presenter. Each experiment covered a different topic chosen by the presenter. While each presenter participated only once, some passive participants — who attended as listeners and used the analysis tool — took part in multiple sessions.

To maintain consistency, all presenters followed the same setup. The border visualization was displayed throughout the entire meeting, while an additional feedback tool in the bottom left corner was accessible at different times. To evaluate both feedback designs, the available tool changed midway through the presentation: for the first six minutes, presenters had access to the aggregated emoji bars, after which this feature was disabled and replaced with the heatmap.

To counterbalance potential order effects, four presenters started with the aggregated feedback widget, while the remaining four began with the individual feedback widget. Additionally, to minimize distraction, presenters had the option to manually toggle the feedback tool's visibility using a dedicated button.

All collected data was stored on a secure server, ensuring anonymity through unique participant IDs. No personally identifiable information was recorded. The dataset included individual gaze and emotion data, as well as the aggregated emotional levels across all participants throughout the meeting.

5.2 Participants

The eight presenters, aged between 23 and 39, included six males and two females. They were researchers, PhD students, and experts, each delivering a 12-minute lecture. All presenters had prior experience with online presentations; however, only one had previously used a tool that provided real-time audience engagement or emotion feedback.

The audience comprised 18 unique participants, ranging in age from 22 to 45, with a gender distribution of 16 males and two females. Most of them attended multiple meetings, which was no problem for the quality of the data since all presentations were unique and had pauses in between. All participants had either completed or were currently pursuing a university degree. They had prior experience attending online lectures and were proficient in using computers. The setup process for the emotion and gaze analysis tool was generally straightforward, with most participants encountering no difficulties.

5.3 Results

After completing all experiments, we gathered and processed the data from both the server and the questionnaire.

5.3.1 Quantitative Data

After gathering demographic data about the lecturers in Chapter 1 of the follow-up questionnaire, we shifted our focus to comparing the feedback tool with a conventional setup that lacks emotional or gaze feedback.

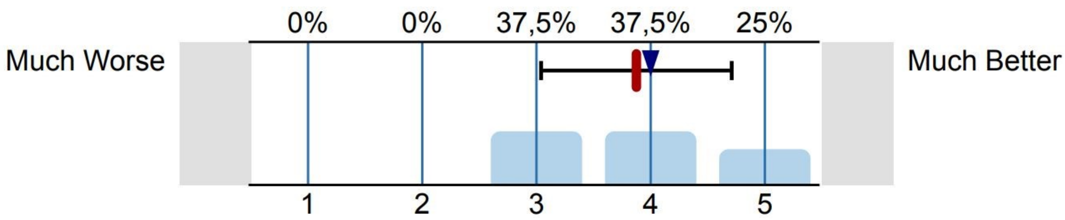

Figure 17: Comparison of feedback tool to conventional setup with no feedback.

Most lecturers responded positively to our application with live feedback, as illustrated in Figure 17. In follow-up questions about their motivation during the meeting and any adjustments they made based on the feedback, responses varied slightly. One lecturer reported a "slightly decreased motivation," two reported "no change," and five experienced a "slightly increased motivation" while holding their lecture.

The adjustments made in response to the feedback were consistent with common presentational adaptations. Some lecturers adjusted their speaking pace, while others made a conscious effort to smile more and sound less monotone in response to the audience's reactions.

In the first question of Chapter 3, we received a decisive answer to a central question of this thesis: **Which visualization design is preferred?** With unanimous votes,

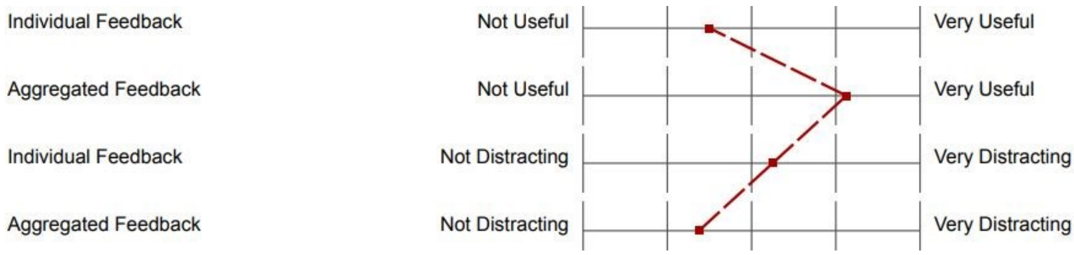

Figure 18: Comparison of usefulness and distraction levels for both feedback designs.

all lecturers favored the aggregated feedback tool over the individual feedback visualization. The rationale for this clear preference becomes evident in the following questions, shown in Figure 18, and through open-ended responses.

When asked about the usefulness and distraction levels of both feedback tools, the questionnaire results were strikingly clear. Most lecturers found the aggregated tool to be quite useful, while the individual feedback visualization was perceived as only slightly useful — if at all.

Similarly, opinions on distraction levels followed the same trend. The aggregated feedback tool was considered only slightly distracting, although still more distracting than having no feedback at all. However, it was notably less distracting than the individual feedback design, which contained more elements changing in real time and provided a significantly higher level of detail. While the increased distraction was expected due to the density of information in the individual feedback visualization, its limited usefulness was not as apparent before conducting the study.

Since both tools could potentially be integrated into online lectures of various sizes, we asked lecturers about their suitability for different group sizes in online meetings. The results were largely predictable. The aggregated feedback tool was seen as suitable for most group sizes. Although none of the lecturers considered it to be the most suitable for very small groups (fewer than 10 participants), half of them believed it was equally effective for all group sizes.

Conversely, the individual feedback tool was perceived as only suitable for very small group meetings. With fewer participants, there would be fewer cells in the heatmap, which reduces distractions for the lecturer. However, one of the eight lecturers even stated that the individual feedback tool is not suitable in any context. These results strongly suggest that beyond a certain number of participants, data must be aggregated for it to be effectively interpreted by the lecturer.

Before answering open-ended questions, lecturers were asked to rate the clarity of the different visualizations within the feedback tool. While the aggregated version was generally well understood, major issues arose in interpreting the border color and especially the heatmap. According to these responses, nearly all visualizations could benefit from further refinement to improve their intuitiveness. Ideally, the tool should be understandable with minimal explanation or introduction, allowing

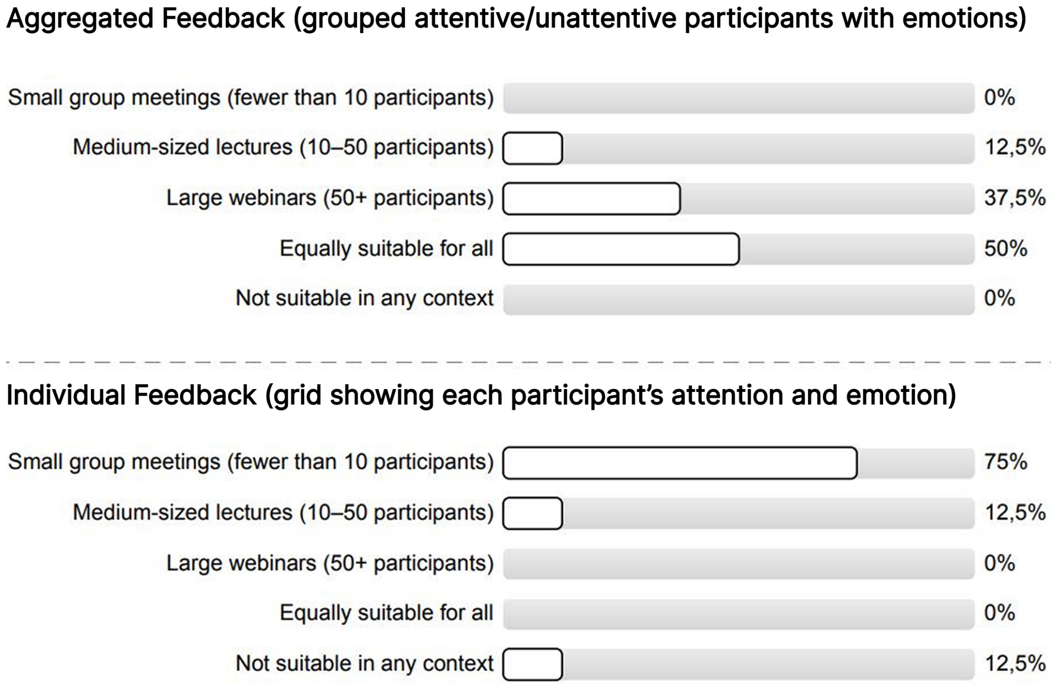

Figure 19: Lecturer opinions on the most suitable group sizes for each feedback design.

lecturers to interpret feedback naturally during their presentations.

5.3.2 Qualitative Feedback

These insights from the lecturers provide valuable feedback for refining the tool further. Here's a structured way to present these points in a more polished form:

- **"It was useful to have a feedback from the audience at all in contrast to not being able to see the audience when presenting."** Despite some distractions and usability concerns, all lecturers agreed that having any form of live feedback was significantly better than presenting without audience visibility. This reinforces the relevance and importance of live-feedback tools in online meetings.

- Lecturers mentioned issues such as **"changing colors," "rapid changes," "looking in the corner," "unusual form of feedback," "six bars total," "big size of feedback window," and confusion about specific elements ("...don't know what that means")**. Many of these concerns were raised multiple times, highlighting the need for refinement. A more condensed version of the border, alongside an extremely simplified emoji bar with possibly only three bars, could be a promising direction.

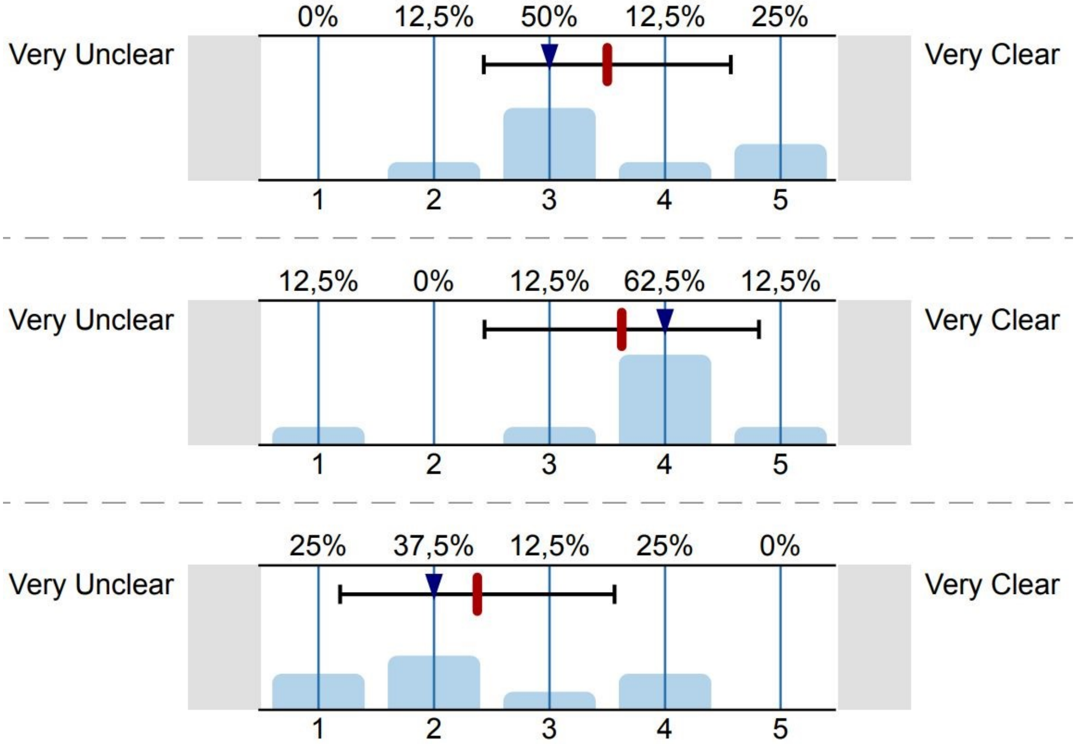

Figure 20: Lecturer ratings on the clarity of different feedback visualizations (top to bottom: Border, Individual Grid, Aggregated Emoji Bar).

- Some lecturers noted that **slow changes reflecting the overall mood, a clean and simple design, and a well-defined spectrum between happy, neutral, and sad** would be most effective. While they found the idea of the border useful, some felt it needed to be clearer or more intuitive rather than redesigned entirely.

- Some lecturers suggested reducing the number of emoji bars to three total for the whole group, as six bars seemed excessive for an aggregated view. Additionally, larger shifts in audience mood should be highlighted with a clearer moment of change, possibly through a subtle mini-alert or emphasis.

- One lecturer suggested having an **aggregated live-feedback tool during the lecture, combined with a detailed summary of audience engagement and emotional trends after the meeting**. Since the data is already collected and processed, adding this post-lecture analysis would require minimal effort while providing valuable insights without distracting the presenter during the session. This suggestion matches with a suggestion of the first survey prior to the design process.

These points make it clear that while the concept of live feedback was appreciated, simplifying and fine-tuning the tool could significantly enhance its usability.

5.3.3 Meeting Data Analysis

The first step involved generating plots for each experiment separately to identify any compromised or anomalous data points. Any values that appeared unrealistic or were the result of technical errors were excluded. During this process, we identified two meetings where emotion data was incorrectly recorded due to unexpected changes in user IDs. To ensure the reliability of our findings, we excluded these two meetings from all calculations related to emotional feedback.

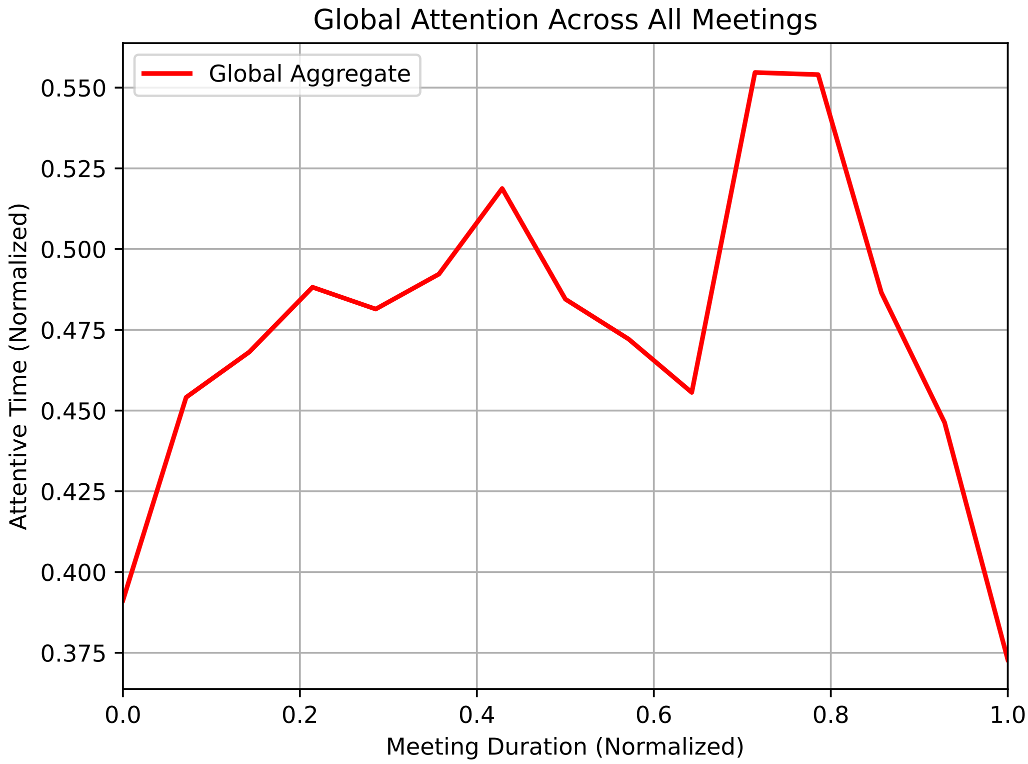

Figure 21: Average percentage of attention across all participants, aggregated over all experiments.

Even though some gaze data appeared unrealistic — such as participants being inattentive for over 80% of the meeting—we decided to keep it in the evaluation. While this could potentially be a technical issue, we could not determine it with certainty. By keeping this data, the overall average attention level is slightly lower, but the trend of gaze changes remains unaffected. Additionally, it reflects a realistic scenario in which some participants may be distracted for the entire 10–12-minute session, further reinforcing the importance of real-time feedback for presenters.

Figure 21 presents a plot of participants' gaze data, aggregated across all meetings. The plot divides each user's attentiveness over time into 20 intervals, calculating the percentage of time spent being attentive within each segment. These individual distributions were then averaged across all meetings to generate an overall trend.

The results reveal that attention levels were lowest at the beginning and end of the sessions. A likely explanation for this trend is that participants may have been engaged in other activities before the experiment began, requiring a few seconds to focus on the lecture. Similarly, towards the end of the sessions, some participants may have experienced fatigue or lost focus, leading to a gradual decline in attentiveness.

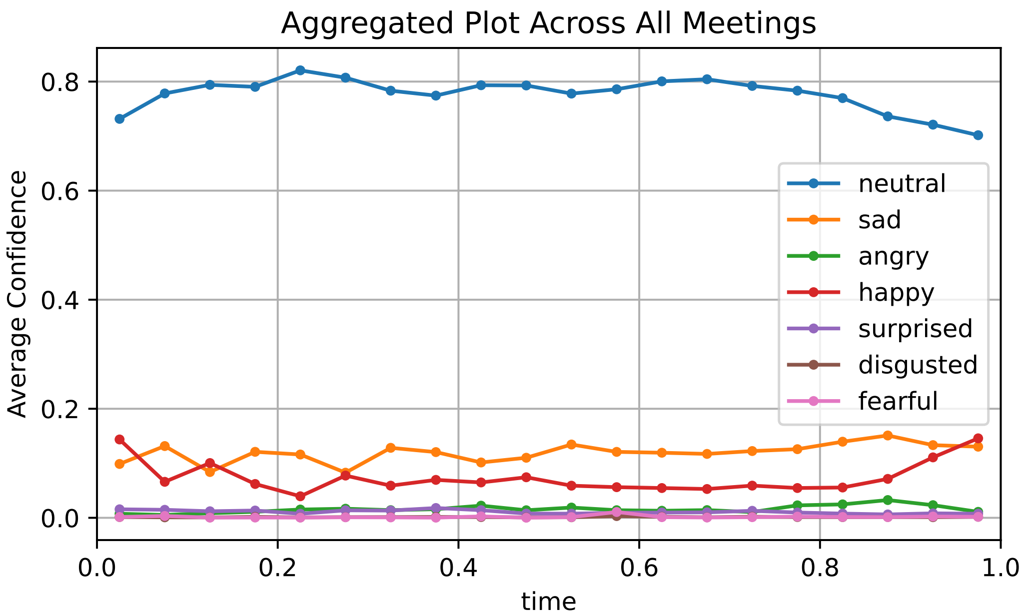

Figure 22: Distribution of all emotions with their average percentage, aggregated across all meetings.

After removing some corrupted data, as previously mentioned, we proceeded to calculate the aggregated emotion levels for all experiments and participants in a single plot. As shown clearly in Figure 22, some negative emotions played an insignificant role in the overall calculations. The three most dominant emotions were "neutral," "happy," and "sad."

Notably, we observed peaks in happy emotions at both the beginning and end of the meetings. The most logical explanation for this trend is the typical structure of a presentation, which often starts with an engaging element — such as an eye-catching statement or relatable story — and concludes with a discussion or meaningful takeaway.

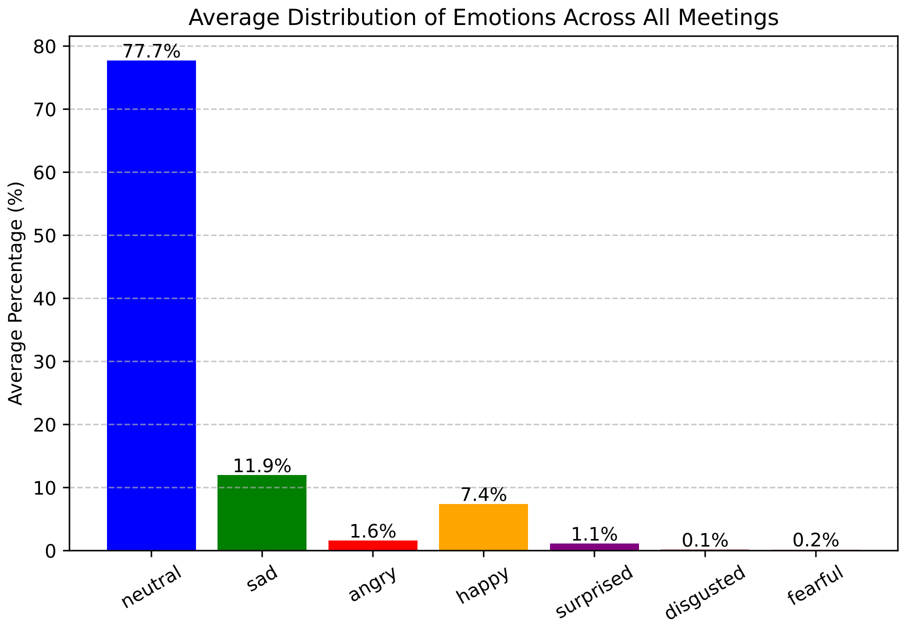

Figure 23: Distribution of all emotions throughout all meetings.

Although the level of sad emotions was relatively high, averaging around 12%, it remained fairly stable throughout all meetings. One possible explanation is the natural appearance of a relaxed face, where slightly downturned lips may be misinterpreted as sadness by the emotion recognition system. Given that all presentations were educational in nature, the overwhelmingly dominant emotion was "neutral," accounting for nearly 80% of the data. If the meetings had involved more interaction, such as open discussions or debates, the proportion of happy emotions would likely have been higher.

The most interesting insights emerged from analyzing how emotions changed throughout a meeting. In Figure 24, we calculated the rate of emotional shifts over time, aggregated across all meetings. There was a clear trend: at the beginning of each meeting, emotions fluctuated frequently, likely due to participants settling in, adjusting their focus, or reacting to the introduction. As the meeting progressed, participants became more engaged, leading to fewer emotion changes. Towards the end, there was once again a noticeable increase in emotional shifts, particularly a rise in happy emotions, possibly in response to concluding remarks, discussions, or interactions.

Finally, we calculated the variance of all emotions and identified the most significant emotion shifts between consecutive data points in Table 1.

When examining the variance values, we can clearly observe relatively high variance for the three dominant emotions. As previously analyzed, happy emotions

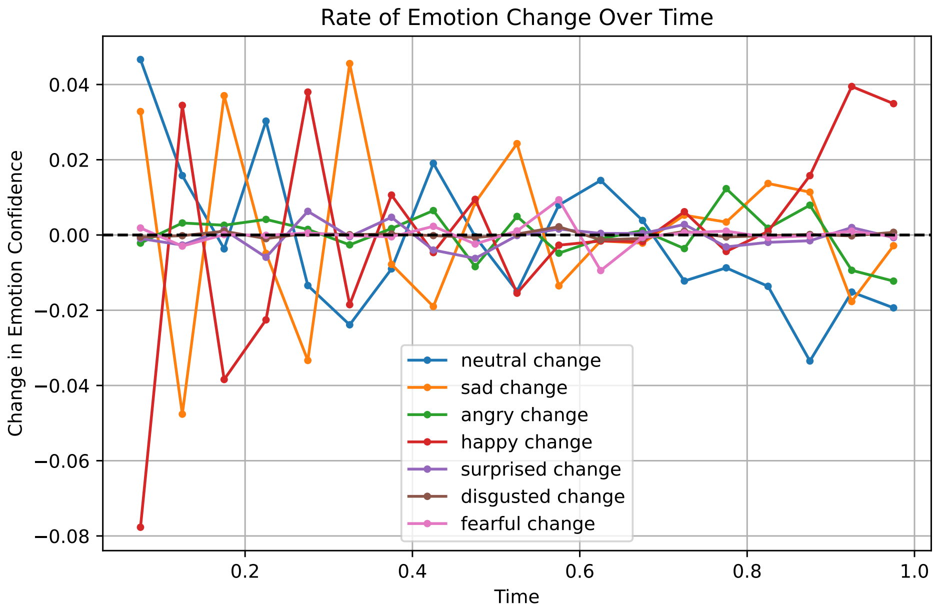

Figure 24: Change rate of emotions during individual meetings, aggregated across all experiments.

	Variance	Biggest emotional shift
neutral	0.0161	0.0466
sad	0.0175	0.0476
happy	0.0199	0.0777
angry	0.0048	0.0123
disgusted	0.0006	0.0022
fearful	0.0019	0.0094
surprised	0.0025	0.0063

Table 1: Variance of all emotions and biggest emotion shift between consecutive data points.

exhibit the highest variance. The same trend is reflected in the largest individual emotion shifts, where the most significant increase in happy emotions is noticeably higher than that of any other emotion.

A logical explanation for this pattern lies in the differing sources of emotional shifts. Happy emotions are often triggered by a shared experience — such as a humorous remark or a relatable moment — which multiple participants tend to react to simultaneously. In contrast, shifts toward sad emotions are more individual and do not necessarily stem from a specific event in the presentation. Instead, they might be influenced by external factors unrelated to the meeting. This collective behavior likely explains why changes in happy emotions tend to be more pronounced and occur in more synchronized patterns across consecutive data points.

6 Conclusion

After conducting all experiments and analyzing the collected data, we can confidently conclude that all participants valued and engaged with the framework. While some responded very positively, others simply appreciated having any form of audience feedback, especially when compared to the conventional setup of online meetings, which typically lack real-time insights into audience engagement and mood.

The evaluation also highlighted a clear preference for the aggregated feedback tool. While a few participants expressed a particular liking for the border visualization, the general consensus favored aggregating gaze and emotion data to provide an overall mood representation. Several optimization opportunities emerged from the post-experiment questionnaire, offering valuable insights for refining the tool further.

As online meetings continue to play a crucial role in education and professional collaboration, addressing the lack of social interaction remains essential in developing effective virtual communication tools. The need for real-time audience feedback is greater than ever, presenting a significant opportunity to bridge the gap between traditional in-person interactions and virtual environments.

6.1 Discussion, Limitations, and Future Improvements

While all participants in the survey, questionnaire, and experiment agreed on the value of emotion and gaze feedback, its implementation presented certain challenges. Most of the concerns raised by lecturers in the questionnaire revolved around design optimization, but additional ethical considerations and suggestions for expanding the tool's use cases were also highlighted.

6.1.1 Use Cases and Applicability

One key challenge is that the tool requires some familiarization before it can be used effectively. Presenters must analyze the feedback visualizations, interpret the data, and adjust their delivery accordingly. This process requires both practice with the tool and confidence in the presentation material. In lectures where the presenter is well-prepared and familiar with their content, integrating live feedback is more manageable. However, for less confident lecturers, a post-lecture summary — as suggested in the questionnaire and pre-experiment survey — could be a useful alternative, allowing them to reflect on audience engagement without real-time distractions.

Another significant advantage of the aggregated feedback visualization is its scalability. Unlike individual feedback, which becomes overwhelming in larger meetings, the aggregated tool remains effective regardless of audience size, making it adaptable for small discussions as well as large lectures.

6.1.2 Ethical and Privacy Considerations

To address privacy concerns, one potential improvement is to set a minimum participant threshold before feedback is displayed. In small meetings (e.g., 3-5 participants), an individual's engagement level could disproportionately impact the overall visualization, potentially making them feel singled out. This could create pressure on attendees, making them feel watched rather than naturally engaged.

However, in larger meetings, the aggregated nature of the feedback ensures that no single participant's emotions or gaze behavior can be traced back to them. Additionally, since the system works even with cameras off, it prevents lecturers from directly associating feedback with specific individuals, enhancing privacy and reducing discomfort.

6.1.3 System Improvements

Future iterations of the tool should offer customizable settings, allowing lecturers to tailor the feedback visualization to their personal presentation style and needs. The questionnaire revealed that each presenter has individual preferences in how they perceive and utilize feedback. Some key improvements could include:

- **Adjustable feedback granularity:** Users could choose between three emoji bars (simplified) or six emoji bars (differentiating between gaze and emotions).
- **Minimalist feedback mode:** An option to display only the border visualization with subtle alerts for major shifts in audience engagement.
- **Personalized engagement thresholds:** Presenters could define what constitutes "high engagement" based on their teaching style — for example, one lecturer may expect 30% engagement, while another may aim for near-total attention.

A more individualized approach would help address multiple concerns simultaneously, rather than attempting to optimize the tool universally for all users. Further testing and user feedback will be essential in refining the system and developing a reliable version suited for diverse online teaching and meeting environments.

References

[1] Oliver Ferch, Alexander Fischer, Jan Fuhrmann, Alexander Völker, and Chandan Kumar. Detecting and analysing user engagement in virtual meetings. In *Student Project Report, University of Stuttgart*, 2021.

[2] Chandan Kumar, Bhupender Saini, and Steffen Staab. Enhancing online meeting experience through shared gaze-attention. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024.

[3] Holger Regenbrecht and Tobias Langlotz. Mutual gaze support in videoconferencing reviewed. *Communications of the Association for Information Systems*, 37(1):45, 2015.

[4] L-Q Xu, A Loffler, PJ Sheppard, and D Machin. True-view videoconferencing system through 3-d impression of telepresence. *BT technology journal*, 17(1):59–68, 1999.

[5] Sarah D'Angelo and Darren Gergle. Gazed and confused: Understanding and designing shared gaze for remote collaboration. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pages 2492–2496, 2016.

[6] Michael Burch. Gaze-based monitoring in the classroom. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*, pages 1–3, 2023.

[7] Moritz Langner, Peyman Toreini, and Alexander Maedche. Eyemeet: A joint attention support system for remote meetings. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–7, 2022.

[8] Bill Buxton. Mediaspace–meaningspace–meetingspace. In *Media space 20+ years of mediated life*, pages 217–231. Springer, 2009.

[9] Erzhen Hu, Jens Emil Sloth Grønbæk, Austin Houck, and Seongkook Heo. Openmic: Utilizing proxemic metaphors for conversational floor transitions in multiparty video meetings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.

[10] Prabin Sharma, Shubham Joshi, Subash Gautam, Sneha Maharjan, Salik Ram Khanal, Manuel Cabral Reis, João Barroso, and Vítor Manuel de Jesus Filipe. Student engagement detection using emotion analysis, eye tracking and head movement with machine learning. In *International Conference on Technology and Innovation in Learning, Teaching and Education*, pages 52–68. Springer, 2022.

[11] Ulf Dimberg, Monika Thunberg, and Kurt Elmehed. Unconscious facial reactions to emotional facial expressions. *Psychological science*, 11(1):86–89, 2000.

[12] Robert R Provine. *Laughter: A scientific investigation*. Penguin, 2001.

[13] Barbara L Fredrickson. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American psychologist*, 56(3):218, 2001.