Attention of Many Observers Visualized by Eye Movements

Teresa Hirzle*
teresa.hirzle@uni-ulm.de
Ulm University, Institute of Media Informatics
Ulm, Germany

Marian Sauter*
sauter.marian@gmail.com
Ulm University, Institute of Psychology and Education
Ulm, Germany

Tobias Wagner
tobias.wagner@uni-ulm.de
Ulm University, Institute of Media Informatics
Ulm, Germany

Susanne Hummel
susanne-1.hummel@uni-ulm.de
Ulm University, Institute of Psychology and Education
Ulm, Germany

Enrico Rukzio
enrico.rukzio@uni-ulm.de
Ulm University, Institute of Media Informatics
Ulm, Germany

Anke Huckauf
anke.huckauf@uni-ulm.de
Ulm University, Institute of Psychology and Education
Ulm, Germany

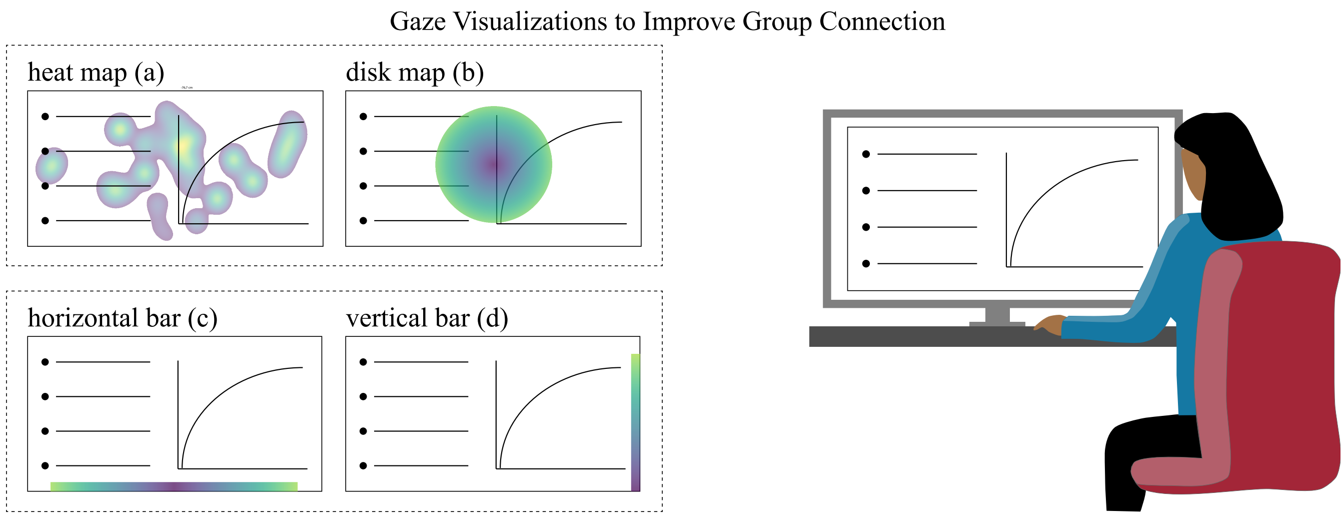

Figure 1: Eye movements of a group provide an important estimate of the attentional state of that group. To equip speakers with information of this crowd attention also during online interaction, we designed and implemented four versions visualizing different aspects of the gaze distributions (heat map (a), *disk map* (b), *horizontal bar* (c), and *vertical bar* (d)). In an evaluation study, 72 teachers preferred displays of location-specific gaze information and variability in terms of social connectedness and perceived usefulness, usability, and cognitive demand.

ABSTRACT

Interacting with a group of people requires to direct the attention of the whole group, thus requires feedback about the crowd's attention. In face-to-face interactions, head and eye movements serve as indicator for crowd attention. However, when interacting online, such indicators are not available. To substitute this information, gaze visualizations were adapted for a crowd scenario. We developed, implemented, and evaluated four types of visualizations of crowd attention in an online study with 72 participants using lecture videos enriched with audience's gazes. All participants reported increased connectedness to the audience, especially for visualizations depicting the whole distribution of gaze including spatial information. Visualizations avoiding spatial overlay by depicting only the variability were regarded as less helpful, for real-time as well as for retrospective analyses of lectures. Improving our visualizations of crowd attention has the potential for a broad variety of applications, in all kinds of social interaction and communication in groups.

*Both authors contributed equally to the paper

ETRA '22, June 8–11, 2022, Seattle, WA, USA

ACM ISBN 978-1-4503-9252-5/22/06.
https://doi.org/10.1145/3517031.3529235

CCS CONCEPTS

• **Applied computing** → **E-learning**; • **Human-centered computing** → *Empirical studies in visualization.*

KEYWORDS

gaze interaction, gaze visualization, gaze controlled UI, usability, gaze-based interfaces

ACM Reference Format:
Teresa Hirzle, Marian Sauter, Tobias Wagner, Susanne Hummel, Enrico Rukzio, and Anke Huckauf. 2022. Attention of Many Observers Visualized by Eye Movements. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22), June 8–11, 2022, Seattle, WA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3517031.3529235

1 INTRODUCTION

While interacting with a group, speakers have to be able to direct and thus also to observe the group's attention. This so-called crowd attention is perceived mainly via gaze [Sun et al. 2017], by rapidly pooling information from many members to estimate the direction of a group's collective gaze [Sweeny and Whitney 2014]. Understanding crowd attention has recently been suggested to play a central role in perceiving group intentions, orchestrating joint attention, and guiding behavior [Sun et al. 2017; Sweeny and Whitney 2014]. In classrooms, it has already been shown that this crowd attention is used to manage the audience's attention [Nizielski et al. 2012; Packard 1970; Sun et al. 2017], create an interpersonal connection to the students, and improve the quality of teaching and learning [Jarodzka et al. 2013; Korthagen et al. 2014; Pennings et al. 2018; Richmond 2002; Wubbels and Brekelmans 2012]. In addition, a lack of crowd attention cues can negatively affect teaching quality and the students' learning experience. As a result, teachers and students report missing *"social presence, interaction, satisfaction and overall quality"* in online teaching compared to face-to-face learning [Nambiar 2020, p. 791]. While in a co-located classroom setting, gaze cues allow establishing crowd attention relatively easily, in current video-conferencing platforms, such as Zoom[1] or Microsoft Teams[2], crowd attention cues are missing. With our work, we evaluate the potential to communicate crowd attention by visualizing the distribution of gazes during interaction.

For determining the attentional state of a group of observers, there are a lot of fixations to be depicted at each point in time. Combining all these scan paths produces a huge amount of information which can hardly be processed by presenters in real-time, especially as a secondary task during other primary interaction tasks. Therefore, visualizing crowd attention needs to be based on some pooling algorithm. As has been shown, not only the mean gaze position of a whole audience, but also the variability of students' gazes seems to be important to assess the current crowd attention state [Hassib et al. 2017; Sun et al. 2017].

We explore four different ways of visualizing crowd attention: Since heat maps are a well-known and intuitively understandable visualization of eye movements [Blascheck et al. 2017; Raschke et al. 2014], we developed a heat map-like overlay of eye movements on slides (Figure 1 (a)). However, such detailed information might provide too extensive information to be processed. The disk map was therefore developed as a simplified version (Figure 1 (b)). Both heat map and disk map produce spatial overlay and thus partially hide the content for the presenter. Therefore, we also implemented a version following Sun et al.'s [2019] and Hassib et al.'s [2017] feedback visualization and present the variability of students' gaze points as a vertical bar plot (Figure 1 (d)). This visualization abstains from spatial information and therefore requires additional mental processing steps to extract the information. To facilitate at least some of these processes a centralized horizontal bar plot version was additionally constructed (Figure 1 (c)). These visualizations can be expected to be less distracting but lead to more cognitive load on the presenter [Sun et al. 2019].

Our work aims to use gaze tracking for the direct assessment of crowd attention in virtual meetings, and to reveal an estimation of the necessity and the consequences of spatial information in visualizations communicating crowd attention in online group interaction. To that end, we implemented the four visualizations of crowd attention and presented them for evaluation. Seventy-two participants assessed subjective connectedness, perceived usefulness, usability, and how cognitively demanding it would be to use the visualizations.

2 RELATED WORK

Especially in video-conferencing platforms, it is difficult for a presenter to judge whether viewers pay attention to an intended location. To establish a connection between audience and presenter during live-stream lectures, Sun et al. [2019] present a system that visualizes the audience's aggregated flow-related states boredom, flow, and anxiety using a bar plot and a line chart. It was perceived positively by the presenters in an evaluation study, who noted that it helped them find problems in their lectures, and make adjustments accordingly. While some also mentioned that the feedback introduced additional load, as they had to shift their attention away from the content towards the feedback, they agreed that the feedback helped to make the online lecture *"more similar to traditional real-world teaching"*. With *AffectiveSpotlight* Murali et al. [2021] explore putting an artificial spotlight on selected members of the audience, following the idea of talk shows in which selected audience members are focused to convey a certain emotion to the viewer. In an evaluation, the authors found that the presenters were significantly more aware of the audience when using the AffectiveSpotlight system in comparison to a control condition.

There is a wide breadth of applied eye tracking research, including research to understand how the eyes inform mutual interaction (see [Valtakari et al. 2021] for a recent review). In particular, it was shown that eye movements can help to estimate the overall attentiveness [Kar et al. 2020], help predict whether students in class answer questions correctly, and help assess the question's difficulty [da Silva Soares et al. 2021]. A recent eye tracking study [Qvarfordt and Lee 2018] even found that the audience only looked at what the presenter talked about 53% of the time, indicating a need for direct feedback of the audience's attention to the presenter. Thus, measuring eye fixations of presentation viewers can inform about social, affective and cognitive factors, allowing the assessment of all viewers' attention and their understanding.

In in-person meetings, head and eye movements of many persons can be visually pooled. When pooling in online interaction, the fixation positions become available as a variety of fixation positions for a sequence of points in time. Visualizing all single values easily leads to an overload for the observer. So, how to optimally depict measures of central tendency and variance of a groups' fixation positions? For visualizing gaze, heat maps are a widespread

[1] https://zoom.us/, last accessed: February 17, 2022
[2] https://www.microsoft.com/microsoft-teams/group-chat-software, last accessed: February 17, 2022

technique [Bojko 2009; Burch et al. 2019; Pfeiffer and Memili 2016]. Heat maps are used as analytic tools to gain detailed information about the visual attention of viewers or an audience [Blascheck et al. 2017]. They have already been successfully used for retrospective analysis, e.g., to analyse students' gaze behavior in massive open online courses [Sharma et al. 2020] or for the analysis of the usability of learning material [Conley et al. 2020]. In heat maps, typically fixations over time are accumulated. Traditional heat maps thus cannot indicate differences between viewers or viewer groups. Isokoski et al. present several methods that focus on these inter-person differences, e.g., *difference maps* or *deviation maps* [Isokoski et al. 2018]. These heat maps show location-specific gaze variability overlaying the viewed content. This overlay impedes processing the content. Other works on visualizing physiological measures of a group used simple bar or line graphs to visualize audience feedback [Hassib et al. 2017; Sun et al. 2019]. These representations appear very compact and have shown to be easily comprehensible [Hassib et al. 2017; Sun et al. 2019]. Differing from heat maps, however, such bars only visualise the variability agnostic to the distributions across the content, and without showing a central tendency.

The omission of visualizing measures of central tendency might be less distracting (because visuals then do not necessarily overlay the content), but might also require more cognitive demand to process and interpret (because a spatial offset between the actual gaze and the visuals is created). Especially concerning crowd attention, one might argue that the heterogeneity - and thus, the deviation - of fixation positions might be most crucial for a speaker. Hence, to examine how important it is to also display location-specific gaze information for evaluating crowd attention, we compared visualizations including location-specific information and variability with those only depicting variability. To this end, we developed, implemented, and examined two versions for each of these kinds of visualizations: For visualizing central tendencies and variability, we designed a heat map that directly shows the distribution of a groups' eye movements on the slide (based on Isokoski et al.'s [2018] deviation map). Reducing the spatial overlay, we also designed a circle depicting the global mean and standard deviation of the groupsfixations. Visualizing only the standard deviation, we presented students' eye movements as a bar plot, moving horizontally in both versus vertically only in one direction, similar to the previous works of Sun et al. [2019] and Hassib et al. [2017].

3 EVALUATIONS OF CROWD ATTENTION VISUALIZATIONS

3.1 Participants

We recruited 100 participants whose job involved teaching (19-65 years, average 39.5 years; 50 male, 50 female) using the survey platform Prolific[3]. For analyses, we included data of all participants passing the attention check and indicating to be experienced in giving online lectures (n= 72; missing data are specified in the evaluation section).

[3]https://prolific.co/, last accessed: February 17, 2022

3.2 Implementation of Crowd Attention Visualizations

We present four gaze visualizations (see Figure 1). For the disk and both bar maps, crowd attention is implemented as the variability of the audience's eye movements, in particular the standard deviation (σ). It represents the degree to which individual gaze points deviate from the mean of all gaze points. The advantage of using the standard deviation over other variability measures, such as the variance, is that it is measured in the same unit as the original data, i.e., distance on a x/y-plane. High variability occurs when fixations differ between screen positions. Low variability is indicated by the majority of the audience members looking at a similar region on the slides (see our video Figure for a dynamic illustration of the visualizations).

To create the visualizations, for each two seconds in a video, a visualization frame was generated. Then each two visualization frames were interpolated to smooth the transition between the single frames, using linear interpolation over eight frames (see the video Figure for details).

For the first gaze visualization, we present a **heat map** (Figure 1 (a)) – a well-known and widely adopted technique to visualize eye movements [Blascheck et al. 2017; Raschke et al. 2014]. Our heat map is a two-dimensional overlay that indicates the pixel-wise standard deviation of the audience's gaze points from the mean of these gaze points at a specific location. It was implemented based on Isokoski et al.'s deviation map [Isokoski et al. 2018].

Figure 1 (b) shows the (**disk map**), a simplified version of the heat map. While the heat map displays the standard deviation of the gaze points at every pixel on the screen, the disk map displays the mean standard deviation of all gaze points to the disk map's radius. Further, instead of showing the variability of viewers' eye movements at every pixel at the screen, the disk map is positioned at the mean of all gaze points. First, the mean of the viewers' gaze points in a time window t is calculated $P(x,y)_{mean}$. Then, per viewer, the euclidean distance $d_{distance}$ of the gaze point to the mean point $P(x,y)_{mean}$ is calculated for each time window. Finally, the standard deviation of the per-viewer euclidean distances is calculated (d'_{σ}) and mapped to the disk's radius.

As both (heat map and disk map) spatially overlay the content on the slides, we also present two visualization variants that abstain from spatial information. First, following Sun et al.'s [2019] and Hassib et al.'s [2017] visualization technique, we implemented a **vertical bar**. This visualization maps the complex information about an audience's eye movement variability (i.e., mean of the standard deviation of individual gaze points to the mean of all gaze points) to the height of a bar plot, which might require additional mental processing power. Therefore, with the **horizontal bar**, we also present a centralized bar plot version. The horizontal bar is fixed to the center at the bottom of the screen (i.e. it does not 'move' with the mean gaze point). Here, the mean standard deviation of all individual gaze points to the mean of all gaze points is mapped to the width of the horizontal bar (Figure 1 (c) in both directions equally). This visualization can be understood as a position-fixed one-dimension fold of the disk map: the disk is folded by one dimension and compressed to a horizontal bar that increases and decreases bidirectionally. The bidirectional dynamics of the horizontal bar still give an impression about the spatial distribution of the audience members' eye movements, as it spreads into two directions. Therefore, it represents an intermediate step between disk map and vertical bar.

We implemented the gaze visualizations based on data that was retrieved in a webcam-based eye-tracking study [Sauter et al. 2022] using four 3-min conference talks with around 12 viewers per video, which represents the average size of a small group in a University seminar. For this first general evaluation of visualizations, we relied on pre-recorded talks as there is currently no live solution available.

3.3 Study Design and Procedure

A repeated-measures design with the independent variable **gaze visualization** with four levels: heat map, *disk map*, *horizontal bar*, and *vertical bar* was used with a 4x4 Latin square to randomize condition order and distribute participants evenly to the four groups. After giving informed consent, participants were told that they would imagine being a teacher giving four three-minute scientific presentations to a student audience while seeing four different visualizations indicating the students' eye movements. Participants were asked to evaluate how visualizations might help them. They started by indicating their personal experience giving online lectures in a pre-experience questionnaire. They then watched the videos in a counterbalanced order. After each video, participants answered an intermediate questionnaire about *subjective connectedness*, *perceived usefulness*, *perceived usability*, and *perceived cognitive demand* of the visualization. Finally, participants ranked the four visualizations with regard to five criteria and provided further qualitative feedback in a post-experience questionnaire. Mean study completion time was 28.5 minutes, and participants were rewarded with 4.5€.

3.4 Questionnaires

At the beginning of the study, the participants indicated their experience with giving online lectures on a 5-point scale. Intermediate questionnaires were rated on 7-point scales: Participants rated three questions on subjective connectedness ("not at all" to "very much") based on the questions of Parmar and Bickmore [2020]. In addition, we employed the *Inclusion of Other in the Self Scale* [Aron et al. 1992], which measures how close a person feels with another individual or group. The scale is a Venn diagram, i.e., the two roles "self" and "other" are shown as two circles. The overlap between the two circles indicates the relationship between "self" and "other". The scale reaches from two separated circles to two almost completely overlapping circles. *Perceived Usefulness.* There were ten questions adapted from the "perceived usefulness"-scale of the *"technology acceptance model for empirically testing new end-user information systems"* [Fred D. Davis 1985]. "Perceived usefulness" is defined as *"the degree to which an individual believes that using a particular system would enhance his or her job performance."* [Fred D. Davis 1985]. To evaluate the presenters' general perception of the usability of the visualizations, we used a usability survey consisting of seven questions. For this questionnaire, we adapted the questions provided by Murali et al. for the evaluation of a public speaking support interface [Murali et al. 2018] and the evaluation of the *AffectiveSpotlight* system [Murali et al. 2021]. Perceived cognitive demand was assessed with the *Rating Scale Mental Effort (RSME)* by Zijlstra and Van Doorn [1985]. This scale is a unidimensional rating scale that stretches from "0" to "150" and contains nine anchor points stretching from "absolutely no effort" to "extreme effort". Respondents indicated their current state of mental effort by positioning a slider to the perceived mental effort level. Finally, we asked the participants to rank the visualizations based on the following five criteria. *"Please rank the visualizations according to which you"* [C1] *"...prefer to use when giving online lectures."* (most preferred/least preferred), [C2] *"...prefer to use for retrospective analysis of recorded online lectures."* (most preferred/least preferred), [C3] *"...perceived as the most helpful for giving online lectures."* (most helpful/least helpful), [C4] *"...perceived as creating the closest connection to the audience."* (closest connection/least close connection), [C5] *"...perceived as the most distracting for giving online lectures"* (most distracting/least distracting).

3.5 Results

Individual questions were analyzed using the non-parametric Friedman test. If this test showed statistical significance, we proceeded calculating the non-parametric pairwise Wilcoxon signed-rank tests to statistically compare individual visualizations. To adjust the p-value for multiple comparisons, the common Holm-method was applied.

Perceived Subjective Connectedness. These measurements constituted of three questions on a scale from very much (7) to not at all (1). For question A1: 'How much of a personal connection did you feel with the audience?', means (± standard deviations) were 3.9 ± 1.9 (heat map), 4.1 ± 1.7 (disk map), 4.7 ± 1.8 (horizontal bar) and 5.0 ± 1.7 (vertical bar). A Friedman test indicated a significant difference, $\chi^2(3) = 15.52, p = .001, N = 68$. Post hoc Wilcoxon tests revealed this difference to be between the vertical bar and disk map condition ($p_{holm} = .015$) plus heat map condition ($p_{holm} = .003$) as well as between the heat map and horizontal bar condition ($p_{holm} = .024$). For question A2: 'How easy was it to see the non-verbal feedback from the audience?', means were 3.2 ± 1.9 (heat map), 3.4 ± 1.7 (disk map), 4.1 ± 1.8 (horizontal bar) and 4.2 ± 2.0 (vertical bar). A Friedman test indicated a significant difference, $\chi^2(3) = 19.02, p < .001, N = 68$. Post hoc Wilcoxon tests revealed this difference to be in all comparisons ($ps_{holm} < .017$) except between disk map and heat map ($p_{holm} = .84$). For question A3: 'How easy do you feel it would be to respond to the non-verbal feedback from the audience?', means were 4.1 ± 1.8 (heat map), 4.1 ± 1.7 (disk map), 4.5 ± 1.9 (horizontal bar) and 4.9 ± 1.7 (vertical bar) which did not differ $\chi^2(3) = 7.12, p = .068, N = 68$.

Inclusion of Others in the Self. This was rated in the style of a Venn diagram . Mean ratings were 3.7 ± 1.8 (heat map), 3.3 ± 1.5 (disk map), 2.9 ± 1.4 (horizontal bar) and 2.8 ± 1.4 (vertical bar). A Friedman test indicated a significant difference, post hoc Wilcoxon tests revealed this difference to be between the heat map and vertical bar condition ($p_{holm} = .008$) as well as the heat map and horizontal bar condition ($p_{holm} = .008$).

Perceived Usefulness. This was constituted of ten questions rated on a scale from very much (7) to not at all (1). None of the Friedman tests showed any significant differences, (all $ps = .21$).

Perceived Usability. These measurements constituted of seven questions rated on a scale from very much (7) to not at all (1). Significant differences (see Figure 2) in mean ratings (± SD) between the visualizations were only found for the second question (U2: 'How anxious do you think you would feel when using the visualization?'), $\chi^2(3) = 25.76, p < .001, N = 67$ and fourth question (U4: 'How distracting do you think the visualization would be when delivering a lecture?'), $\chi^2(3) = 54.59, p < .001, N = 67$. For the second question, reported means were 3.9 ± 2.0 (heat map), 4.1 ± 1.9 (disk), 4.7 ± 1.8 (horizontal), 4.7 ± 1.8 (vertical). Post hoc paired Wilcoxon tests indicate that the heat map leads to more anxious feelings as compared to the horizontal bar ($p_{holm} = .018$) and compared to the vertical bar ($p_{holm} = .029$). For the fourth question, reported means were 2.3 ± 1.6 (heat map), 2.9 ± 1.7 (disk), 4.4 ± 1.8 (horizontal), 4.5 ± 1.8 (vertical). Post hoc paired Wilcoxon tests indicate that except for the comparison vertical bar ≠ horizontal bar ($p_{holm} = .36$), all differences are significant (all $ps_{holm} = .039$). For the other questions, the Friedman tests did not yield significant differences ($ps > .092$).

Perceived Cognitive Demand. This was measured with the RSME scale for mental effort, stretching from "0" indicating "absolutely no effort" to "150" indicating "extreme effort". Although descriptively different (41 ± 32 for the heat map, 40 ± 31 for the disk map, 44 ± 29 ms for the horizontal bar, and 45 ± 34 for the vertical bar), there were no significant differences $\chi^2(3) = 6.99, p = .072, N = 68$.

After having watched all four visualizations, participants directly ranked them them according to five criteria. There were significant differences in mean ranks for preference for retrospective analysis $\chi^2(3) = 59.90, p < .001$, for helpfulness for giving online lectures $\chi^2(3) = 8.10, p = .044$, for the most distracting during online lectures $\chi^2(3) = 91.22, p < .001$, and for connection to the audience $\chi^2(3) = 40.12, p < .001$. Post hoc Wilcoxon comparisons are shown in Figure 3.

In the qualitative feedback section, six participants mentioned that they found the heat map distracting, two highlighted the visualizations' potential for retrospective analysis. Several comments highlight that a person would have to be trained using the visualizations, as they would be too difficult to apply right away.

4 DISCUSSION

The objective of this work was to support online gaze-based interaction in groups, and identify important measures to be included in visualizations of crowd attention. Therefore, we developed four versions of visualizations based on previous works depicting either only the variability of fixations or also location-specific information. These four versions were implemented and evaluated in an online study using a teaching scenario. Regarding subjective connectedness while watching the four versions of visualizations, our participants reported effects of the visualizations: There was highest connection with the heat map visualization, and lowest for the vertical bar, what is particularly well illustrated in questions 'How much of a personal connection did you feel with the audience' and 'How easy was it to see the non-verbal feedback from the audience?'. The results of *Inclusion of Others in the Self* scale confirm these findings. Both visualizations including location-specific information (heat map and disk map) show the students' eye positions on the slides, while the horizontal and the vertical bar do not present this spatial information. The according spatial proximity of the feedback on the slides seems to facilitate the understanding about the crowds' attention resulting in increased impressions of connectedness.

In terms of perceived usefulness, we could not observe differences between the four visualizations. Most of the scales were rated as neutrally useful. However, participants indicated that they would feel significantly less anxious when using the bar visualization compared to the heat maps. Also, they perceived both, the horizontal and the vertical bar as less distracting than the other visualizations and the heat map was rated as significantly more distracting than the disk map (see Figure 3 (C5)).

All visualizations achieved neutral and comparable values concerning the ease of usability. Also, there was no difference in the perceived cognitive demand, suggesting that the cognitive demands required to decode the bars would roughly balance out the cognitive demands created by the spatial overlay of the heat and disc map visualizations. Overall, the heat map produced the closest connections, was easy to understand, but was also the most distracting one, which can be summarized with P(1)'s words *"My favourite one was the heat map. Somewhat ironically I also felt this was probably one of the most distracting but at the same time I felt it was the easiest to view in real time as to what was capturing the attention of those in the audience"*.

Taken together, the versions including spatial information, although being more distracting, were preferred over the bar visualizations showing only variability. This suggests that teachers seem to accept their disturbance, as they were perceived to be more helpful in perceiving crowd attention (see Figure 3). Based on these results, one may conclude that amongst the current versions, disk map is the most promising one. The heat map might be more beneficial for retrospective analysis (see Figure 3 C2).

Visualizations can only be as good as the underlying data. Since these were collected through an online webcam-based study [Sauter et al. 2022], the temporal and spatial resolution can likely be improved. Although this probably does not severely affect the relative evaluations, it might have had consequences for the subjective impressions of the helpfulness and usability of the visualizations. In fact, one might assume that data of higher quality might lead to improved ratings. Nevertheless, if we think about live online lectures as a potential future application, data will be limited to webcam-based systems. Notably, this study relied on pre-recorded lectures (limiting ecological validity) and future research should, since we now know there is general potential in showing gaze-based visualizations, evaluate the possibility of a live system setup.

Overall, four visualizations, were evaluated by 72 participants concerning their perceived usefulness, usability, and cognitive demand in a study. Participants generally appreciated the added information about the attentional state of the audience. They preferred our location-specific visualizations over the abstract visualizations for retrospective analysis, despite them being more distracting. The disk map could thus constitute a suitable trade-off between the

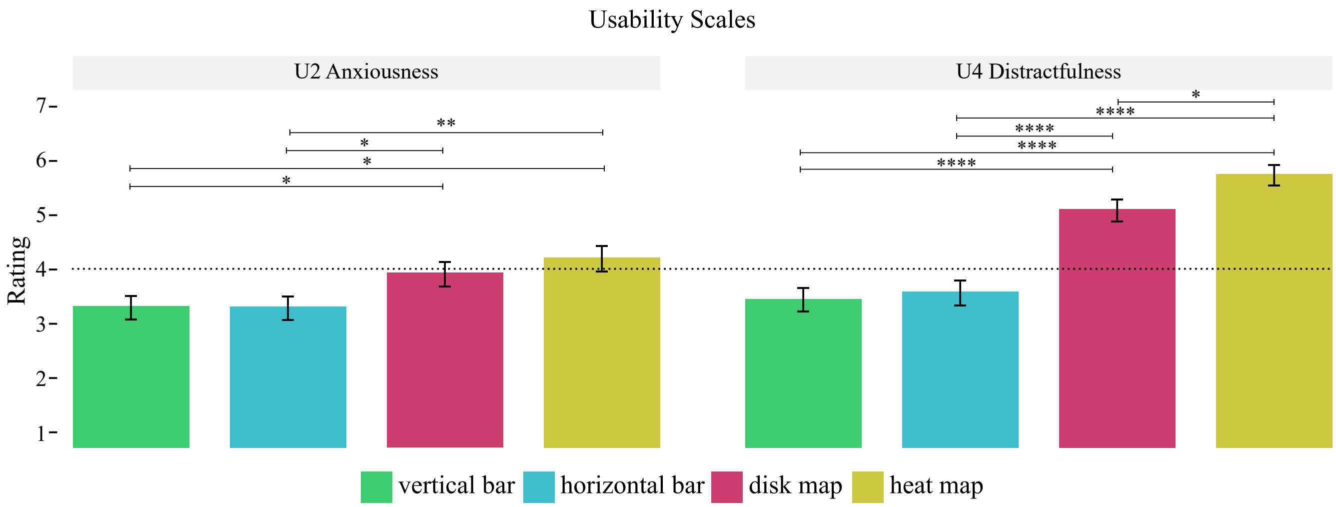

Figure 2: Results for the two questions of the usability scale for which we found statistically significant differences between the visualizations on a 7-point scale, stretching from "not at all (1)" to "very much (7)". U2: 'How anxious do you think you would feel when using the visualization?' and U4: 'How distracting do you think the visualization would be when delivering a lecture?' The horizontal dashed line indicates the middle of the rating scale.

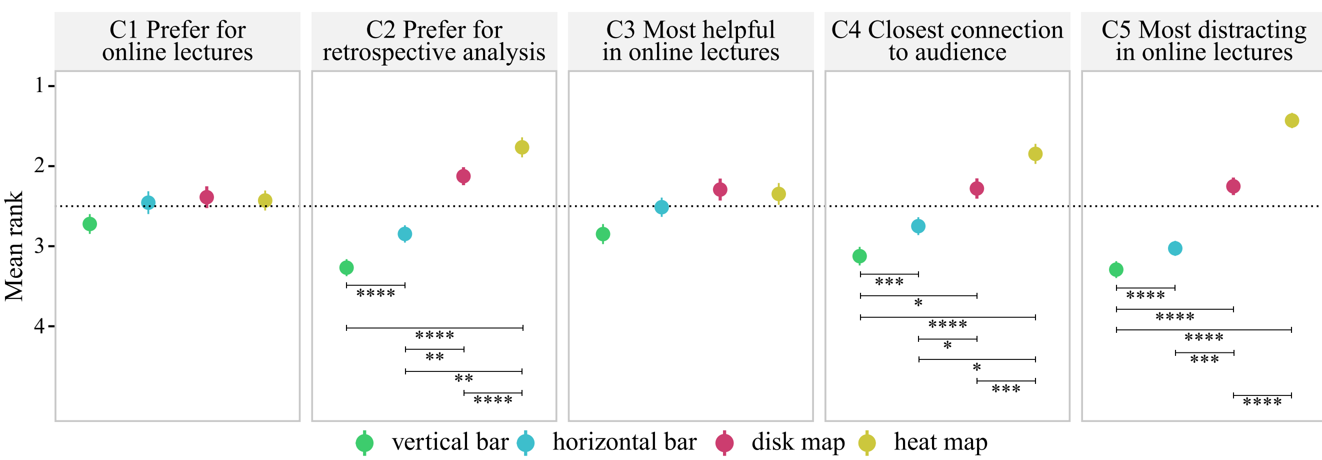

Figure 3: Mean ranks of the visualizations with regard to five criteria. Note that C1-C4 indicate that the desired outcome is on rank 1, while for C5 the desired outcome is rank 4.

spatial proximity of concrete and the unobtrusiveness of abstract visualizations.

The vision for our work is a realistic real-time visualization of crowd attention. This requires webcam-based eye tracking in distributed learning environments in real time, including real-time recording, real-time processing, real-time visualizing, and real-time feeding back the visualizations. With current technology, this comprises a great engineering effort. Nevertheless, such an environment would enable us to also investigate long-term effects on how to make optimal use of such visualizations of crowd attention in various settings.

ACKNOWLEDGMENTS

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, 425867974, to Anke Huckauf and Enrico Rukzio) and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.

REFERENCES

Arthur Aron, Elaine N. Aron, and Danny Smollan. 1992. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology* 63, 4 (1992), 596–612. https://doi.org/10.1037/0022-3514.63.4.596 Place: US Publisher: American Psychological Association.

T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. 2017. Visualization of Eye Tracking Data: A Taxonomy and Survey. *Computer Graphics Forum* 36, 8 (2017), 260–284. https://doi.org/10.1111/cgf.13079 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13079

Agnieszka (Aga) Bojko. 2009. Informative or Misleading? Heatmaps Deconstructed. In *Human-Computer Interaction. New Trends*, Julie A. Jacko (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 30–39.

Michael Burch, Ayush Kumar, and Neil Timmermans. 2019. An Interactive Web-Based Visual Analytics Tool for Detecting Strategic Eye Movement Patterns. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 93, 5 pages. https://doi.org/10.1145/3317960.3321615

Quincy Conley, Yvonne Earnshaw, and Grayley McWatters. 2020. Examining Course Layouts in Blackboard: Using Eye-Tracking to Evaluate Usability in a Learning Management System. *International Journal of Human–Computer Interaction* 36, 4 (2020), 373–385. https://doi.org/10.1080/10447318.2019.1644841 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2019.1644841.

Raimundo da Silva Soares, Jr, Katerina Lukasova, Maria Teresa Carthery-Goulart, and João Ricardo Sato. 2021. Student's Perspective and Teachers' Metacognition: Applications of Eye-Tracking in Education and Scientific Research in Schools. *Frontiers in psychology* 12 (2021), 673615–673615. https://doi.org/10.3389/fpsyg.2021.673615 pubmed:34366993 Publisher: Frontiers Media S.A.

Jr Fred D. Davis. 1985. *A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results*. dissertation. Sloan School of Management at the Massachusetts Institute of Technology. https://dspace.mit.edu/bitstream/handle/1721.1/15192/14927137-MIT.pdf

Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. 2017. EngageMeter: A System for Implicit Audience Engagement Sensing Using Electroencephalography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5114–5119. https://doi.org/10.1145/3025453.3025669

Poika Isokoski, Jari Kangas, and Päivi Majaranta. 2018. Useful Approaches to Exploratory Analysis of Gaze Data: Enhanced Heatmaps, Cluster Maps, and Transition Maps. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 68, 9 pages. https://doi.org/10.1145/3204493.3204591

Halszka Jarodzka, Tamara van Gog, Michael Dorr, Katharina Scheiter, and Peter Gerjets. 2013. Learning to see: Guiding students' attention via a Model's eye movements fosters learning. *Learning and Instruction* 25 (2013), 62–70. https://doi.org/10.1016/j.learninstruc.2012.11.004

Pragma Kar, Samiran Chattopadhyay, and Sandip Chakraborty. 2020. Gestatten: Estimation of User's Attention in Mobile MOOCs From Eye Gaze and Gaze Gesture Tracking. *Proceedings of the ACM on Human-Computer Interaction* 4, EICS (2020), 1–32. https://doi.org/10.1145/3394974

Fred A. J. Korthagen, Saskia Attema-Noordewier, and Rosanne C. Zwart. 2014. Teacher–student contact: Exploring a basic but complicated concept. *Teaching and Teacher Education* 40 (2014), 22–32. https://doi.org/10.1016/j.tate.2014.01.006

Prasanth Murali, Javier Hernandez, Daniel McDuff, Kael Rowan, Jina Suh, and Mary Czerwinski. 2021. AffectiveSpotlight: Facilitating the Communication of Affective Responses from Audience Members during Online Presentations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 247, 13 pages. https://doi.org/10.1145/3411764.3445235

Prasanth Murali, Lazlo Ring, Ha Trinh, Reza Asadi, and Timothy Bickmore. 2018. Speaker Hand-Offs in Collaborative Human-Agent Oral Presentations. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents* (Sydney, NSW, Australia) *(IVA '18)*. Association for Computing Machinery, New York, NY, USA, 153–158. https://doi.org/10.1145/3267851.3267904

Deepika Nambiar. 2020. The impact of online learning during COVID-19: students' and teachers' perspective. *The International Journal of Indian Psychology* 8 (2020), 783–793. https://doi.org/10.25215/0802.094

Sophia Nizielski, Suhair Hallum, Paulo N. Lopes, and Astrid Schütz. 2012. Attention to Student Needs Mediates the Relationship Between Teacher Emotional Intelligence and Student Misconduct in the Classroom. *Journal of Psychoeducational Assessment* 30, 4 (2012), 320–329. https://doi.org/10.1177/0734282912449439 arXiv:https://doi.org/10.1177/0734282912449439

Robert G. Packard. 1970. THE CONTROL OF "CLASSROOM ATTENTION": A GROUP CONTINGENCY FOR COMPLEX BEHAVIOR1. *Journal of Applied Behavior Analysis* 3, 1 (1970), 13–28. https://doi.org/10.1901/jaba.1970.3-13 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1901/jaba.1970.3-13

Dhaval Parmar and Timothy Bickmore. 2020. Making It Personal: Addressing Individual Audience Members in Oral Presentations Using Augmented Reality. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 55 (June 2020), 22 pages. https://doi.org/10.1145/3397336

Helena J. M. Pennings, Mieke Brekelmans, Pamela Sadler, Luce C. A. Claessens, Anna C. van der Want, and Jan van Tartwijk. 2018. Interpersonal adaptation in teacher-student interaction. *Learning and Instruction* 55 (2018), 41–57. https://doi.org/10.1016/j.learninstruc.2017.09.005

Thies Pfeiffer and Cem Memili. 2016. Model-Based Real-Time Visualization of Realistic Three-Dimensional Heat Maps for Mobile Eye Tracking and Eye Tracking in Virtual Reality. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (Charleston, South Carolina) *(ETRA '16)*. Association for Computing Machinery, New York, NY, USA, 95–102. https://doi.org/10.1145/2857491.2857541

Pernilla Qvarfordt and Matthew Lee. 2018. Gaze Patterns during Remote Presentations While Listening and Speaking. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 33, 9 pages. https://doi.org/10.1145/3204493.3204540

Michael Raschke, Tanja Blascheck, and Michael Burch. 2014. *Visual Analysis of Eye Tracking Data*. Springer New York, New York, NY, 391–409. https://doi.org/10.1007/978-1-4614-7485-2_15

V.P. Richmond. 2002. Use and Outcomes. In *Teacher nonverbal immediacy (Communication for teachers)*. West Virginia University, Virginia, 18 pages.

Marian Sauter, Teresa Hirzle, Susanne Hummel, Enrico Rukzio, and Anke Huckauf. 2022. A webcam-based eye tracking analysis on where students look during online lectures. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22)* (Seattle, WA, USA) *(EduEye '22)*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3517031.3529239

Kshitij Sharma, Michail Giannakos, and Pierre Dillenbourg. 2020. Eye-tracking and artificial intelligence to enhance motivation and learning. *Smart Learning Environments* 7, 1 (2020), 1–19. https://doi.org/10.1186/s40561-020-00122-x Number: 1 Publisher: SpringerOpen.

Wei Sun, Yunzhi Li, Feng Tian, Xiangmin Fan, and Hongan Wang. 2019. How Presenters Perceive and React to Audience Flow Prediction In-Situ: An Explorative Study of Live Online Lectures. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 162 (Nov. 2019), 19 pages. https://doi.org/10.1145/3359264

Zhongqiang Sun, Wenjun Yu, Jifan Zhou, and Mowei Shen. 2017. Perceiving crowd attention: Gaze following in human crowds with conflicting cues. *Attention, Perception & Psychophysics* 79, 4 (2017), 1039–1049. https://doi.org/10.3758/s13414-017-1303-z

Timothy D. Sweeny and David Whitney. 2014. Perceiving Crowd Attention: Ensemble Perception of a Crowd's Gaze. *Psychological Science* 25, 10 (2014), 1903–1913. https://doi.org/10.1177/0956797614544510 PMID: 25125428.

Niilo V. Valtakari, Ignace T. C. Hooge, Charlotte Viktorsson, Pär Nyström, Terje Falck-Ytter, and Roy S. Hessels. 2021. Eye tracking in human interaction: Possibilities and limitations. *Behavior Research Methods* 53, 4 (2021), 1592–1608. https://doi.org/10.3758/s13428-020-01517-x

Theo Wubbels and Mieke Brekelmans. 2012. *Teacher–Students Relationships in the Classroom*. Springer Netherlands, Dordrecht, 1241–1255. https://doi.org/10.1007/978-1-4020-9041-7_80

FRH Zijlstra and L Van Doorn. 1985. *The construction of a scale to measure perceived effort*. University of Technology, -.