Using Real-Time Gaze Based Awareness Methods to Enhance Collaboration

Christian Schlösser[1(✉)], Philipp Schlieker-Steens[1], Andrea Kienle[1], and Andreas Harrer[2]

[1] University of Applied Sciences and Arts Dortmund, Dortmund, Germany
{christian.schloesser,philipp.schlieker-steens, andrea.kienle}@fh-dortmund.de
[2] Clausthal University of Technology, Clausthal-Zellerfeld, Germany
andreas.harrer@tu-clausthal.de

Abstract. Using eye-tracking in applications can be used to identify which areas are looked at by their users. In collaborative software this information can be transmitted to partners in real-time to provide an additional information channel. This paper compares different types of real-time gaze data visualizations. For this purpose, a study with three groups is conducted, who have to solve a collaborative puzzle. In every group the gaze data from each participant is recorded and visualized in a different way depending on the specific group condition. The aim is to evaluate a new context-based visualization to be able to make use of the known advantages of coordinate-based gaze data visualization outside of the domain of What-You-See-Is-What-I-See (WYSIWIS) interfaces.

1 Introduction

Working in groups is increasingly mediated by computers. The spatial distance is bridged by distributed applications. Those collaborative applications serve as mediators and help to support the collaboration itself [1]. This support is a key factor and is done in different ways in the specific applications. Awareness tools help the user to give their respective partners the knowledge about their current context, so they can adjust their own actions accordingly [2]. In order to integrate such an awareness support in a collaborative application, eye-tracking can be used as an interactive method.

Efficient collaboration is highly dependent on the so-called joint attention, which denotes a common focus on an object by two or more persons [3, 4]. The concept is very general and represented in various fields of research. Shared Gaze is considered to be the weakest form of joint attention, which describes at least two individuals viewing at the same object. Pietinen et al. [5] assume that only the number and duration of shared gaze events may indicate the intensity of collaboration, even if it is explicitly stated that collaboration takes place outside of the sensor eye-tracker as well. In the personal interaction Shared Gaze can be achieved by keeping track of each other's gaze. Because this method is missing in computer-mediated interaction, an obvious solution is the transfer of the current gaze coordinates (gaze cursor) to each other. This method is known as Gaze Sharing [4, 6]. Prerequisite for gaze cursor sharing are What-You-See-Is-What-I-See (WYSIWIS)

N. Baloian et al. (Eds.): CRIWG 2015, LNCS 9334, pp. 19–27, 2015.
DOI: 10.1007/978-3-319-22747-4_2

interfaces [7], where all users see the same content. This is necessary because otherwise there is no correlation between the gaze data of the partner and the own screen content, which is why, for example, the aspect ratio of the monitor and its screen resolution must be taken into account. Since these identical environments are usually not present outside of laboratory conditions and the WYSIWIS principle is not applicable to all types of software, a more flexible approach to gaze sharing, without the constraint to use WYSIWIS interfaces is required. To achieve this, a transition from coordinate-based (Where?) gaze sharing to a context-based (What?) gaze sharing must be made. This is also supported by the results of [8] that information on the attentional processes can indeed be relevant, but are rarely needed in the detail of the exact gaze coordinates. She assumes, that the knowledge, that a partner is focusing a particular object is sufficient. The aim of this paper is to evaluate a new context-based gaze sharing method in comparison to the coordinate-based gaze cursor sharing. The question is, if the context-based method has similar positive effects and can thus be used outside of the WYSIWIS domain.

2 Experiment

To answer this question, the two types of gaze sharing visualization mentioned above were compared to a no gaze sharing control group. For this purpose, a between-subjects design with three groups was used. In the following, the term Gaze Awareness is used for the group with context-based gaze sharing, the term Gaze Cursor for the group with coordinate-based gaze sharing and the No Gaze for the group with no gaze support. A total of 60 college-level participants were acquired to record ten dyads per group. The average age was 23.2 with a standard deviation of 3.4. The gender ratio was 14 females to 46 males.

Material: The application used in this study was a collaborative puzzle, which was developed with reference to the turtle puzzle from Mühlenbrock [9]. In order to achieve an acceptable solving time, the pieces, in contrast to a real world puzzle, could not be rotated. As shown in Fig. 1, there were nine drop zones for solving the puzzle, as well as nine stack zones, which initially held the puzzle pieces.

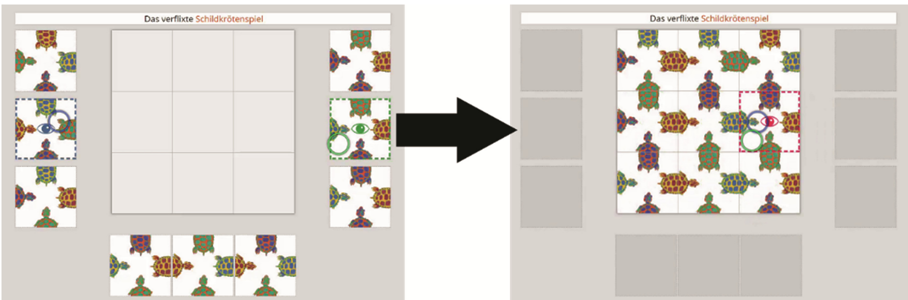

Fig. 1. Turtle puzzle from the initial piece distribution to the solved puzzle. Screenshots taken from the experimenters view with included gaze visualizations.

Every dyad had the same initial piece distribution. Pieces could not be stacked, doing so, led to the system swapping the involved pieces. It was implemented as a WYSIWIS real-time web application, which transmitted every drag and drop with a minimal delay of 1–5 ms to other clients. Conflicts like dragging the same piece or using the same drop zone was prevented by the system.

Participants were assigned to a color, which was used to highlight drop zones and gaze indicators. Figure 2 shows the support of gaze for each group mentioned above. The No Gaze group had no gaze support at all, the second group was supported by a mutual gaze cursor and the third group used a gaze enabled application, which highlighted the visible elements on the screen while fixating it.

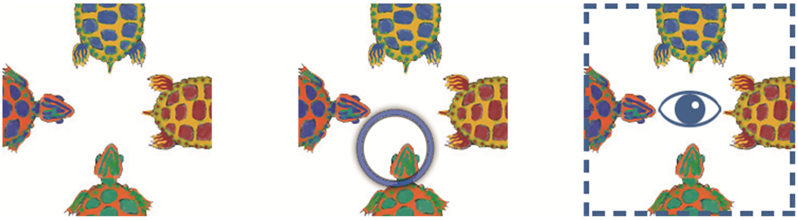

Fig. 2. The gaze visualizations used in each conditions: no gaze, gaze cursor and gaze awareness

Procedure: Each trial included two participants that were briefed about the procedure in general and told they had to collaboratively solve an online puzzle, each participant in a separate room, connected via audio chat. They were informed that the experimenter would join the audio chat for announcements. After that they were presented a description of the puzzle and its game mechanics, as well as condition specific features. The experimenter used a third client computer to observe the puzzle, each participant's eye-tracking status and the audio chat. Irrespective of the current condition, this experimenter's client displayed all gaze related features. The puzzle was started from the experimenters' computer after making sure that the task was understood and the participants were ready. The maximum duration was 20 min, without the participants actually knowing about this time-limit. If close to completion, up to two additional minutes were granted. The experimenter used the voice chat for start and stop signals, as well as required corrections of the participants seating position regarding eye-tracking data quality.

Eye-Tracking Setup: We used two desktop-based Tobii eye-trackers. One TX300 running at 300 Hz as well as one x 120 running at 120 Hz. The user specific eye-tracker calibration was performed using a five-point calibration at the beginning of the experiment. Due to 24 in displays and large areas of interest (puzzle parts and stack zones), no correction for gaze deviation was needed. An in-house server was used to synchronize and capture all gaze data.

Captured Data: During the experiment, various data from different channels were captured. Irrespective of the current condition, the same data was saved. On the one hand, raw interaction data that belongs to game mechanics like drags and drops were saved, as well as computed data like solution status after each drop. On the other hand, raw and computed data associated to gaze were captured, such as raw and denoised gaze data, fixations and pupil diameter. Fixations were complemented by information about underlying elements such as puzzle parts or stack zones. For an in depth retrospective analysis of each session, a video from the shared workspace complemented by the participant's voices and gaze data was recorded.

Hypotheses: To examine the research question three hypotheses were formulated. These hypotheses assume the effectiveness of context-based gaze sharing from different perspectives. The aim is to determine whether it is possible to apply the positive effects of coordinate-based to context-based gaze sharing mechanisms outside of a WYSIWIS interface.

1. The quality of a collaboration in the Gaze Awareness group is on the same level as the Gaze Cursor group and higher than in the No Gaze group.
2. The frequency of Shared Gaze events in the gaze sharing groups is on a comparable level but higher than in the No Gaze group.
3. While using both gaze sharing methods the cognitive load amount is lower than using No Gaze sharing.

3 Methods

For the evaluation of the three hypotheses mentioned above, the following methods were used:

Method 1 - Rating Scheme for Collaboration Quality: The quality of computer-supported collaboration has been quantified and measured by means of a multi-dimensional rating approach in the literature. In the original work [10] a rating scheme was presented for medical diagnosis tasks mediated by video-conferencing; the scheme consists of nine dimensions derived by a combination of bottom-up (empirically induced categories) and top-down (theoretically justified aspects) analyses. This rating scheme has been adapted to synchronous collaborative problem solving tasks and applied to shared workspace scenarios [11]. For our study two dimensions of the original rating scheme have been left out because they were not applicable in our scenario: Time management was not considered because there was no time limit known by the participants a priori, thus there was no need for the participants for timekeeping and scheduling (the used time limit of 20 min was because of practical purposes on the experimenter's side to schedule the dyads on proper intervals). Technical coordination was not relevant because additional tools were not used besides using a mouse and gaze information of the peer (if applicable for the condition). Because of the nature of our problem-solving task some of the dimensions have been adapted based on reformulation via the setting's specific constraints and expected/detected positive and negative behaviors for the adapted dimensions (following closely the approach in [11]). Dimension "Task division" was reformulated as "task coordination" because in the synchronously shared workspace an explicit division of labor and decomposition into sub-tasks is limited, but the coordination of accesses to cards and moves of these is relevant to coordinate the peers' efforts. The dimension of "individual task orientation" was replaced with a score that represents the balance of actions between the peers, because in contrast to the original scheme we didn't have any asymmetries in competence that require individual efforts on specific sub-tasks; in contrast, each participant engaged equally in problem-solving would result ideally in a balance between the contributions, thus we measured the substitute dimension "contribution balance".

A rating handbook was created by the two raters after watching three previously recorded videos with each rater taking notes for all dimensions. Afterwards the two raters discussed their individual perception of how a dimension should be evaluated. Due to the task of solving a puzzle and missing asymmetries in competence between the participants a concise handbook was sufficient. For each dimension requirements for a very good or very bad rating were formulated. If possible examples of communication and action were given. All videos were watched and rated in random order without the raters being aware about the current experimental condition of the dyad. In some cases participants mentioned the gaze visualization so that the rater could conclude the specific group.

Method 2 - Shared Gaze Occurrences: Within a collaboration a high frequency of Shared Gaze events can be an indicator for the efficiency of collaboration [5]. The closer the partners work together, the more often their gazes meet on the same elements which are in the current context of interest. The collision of the partners gaze on elements (puzzle parts and stacks) was automatically captured by the system, using a real-time distance-based hit detection [6]. Therefore no algorithm based on the spatial distance was needed. Accidental collisions were filtered afterwards by using the average fixation duration of all participants as a minimum length for a single Shared Gaze event. As there are no known numbers for high or low Shared Gaze frequencies, which are presumably strongly dependent on the scenario, a comparison with the No Gaze control group was made.

Method 3 – Pupillometry: The method for measuring the pupil diameter is referred to as pupillometry. Holmqvist [12] says that the pupil diameter can be used for interpretations of cognitive load and the recognition of emotions, because the pupil diameter varies depending on the mental workload. Linking that information to the current action of the participant, the mental workload could give information about the complexity of the current exercise. But he emphasizes, that there are many external factors which influence the pupil diameter. The most influencing factor are light conditions. During the experiment only artificial light was used. The puzzle did not cause any noticeable changes to the screen brightness and can be considered constant. To avoid measuring errors due to different eye characteristics of the participants, the lowest values were removed. Because we were interested in the overall cognitive load during a trial in each condition we divided each measure by the total number of data points for each participant.

4 Results

In total more than six hours of footage was recorded. The eye-tracking brings certain restrictions to the participant's movements, because otherwise false or No Gaze data can be recorded. For this reason, some of the participants had to be occasionally reminded to maintain a proper posture, whereby three of the thirty dyads had to be discarded, because no reliable gaze data was recorded. That led to a population for each group as follows: No Gaze n = 10; Gaze Cursor n = 9; Gaze Awareness n = 8. In evaluations where No Gaze data was needed, all 30 dyads were analyzed.

Hypothesis 1: To investigate the quality of collaboration the evaluation scheme described above was used. We evaluated the validity of the rating scheme for our research by checking for the robustness of inter-rater results, using the ICC (intra-class correlation coefficient) on the results of two independent raters on the whole dataset. The result with a mean value of 0.7 is an acceptable match (M = 0,701; SD = 0,035). The application of the adapted rating scheme with the original scale ranging between -2 (very bad) and +2 (very good) by the two independent raters brought the results in Table 1. The leftmost column refers to the numbering in the original rating scheme, the two reformulated dimensions are marked with an asterisk.

Table 1. Results of the adapted rating scheme for quality of collaboration

		No gaze		Gaze cursor		Gaze awareness				
#	Dimension	M	SD	M	SD	M	SD	F(2,27)	p	n^2
1	Sustaining mutual understanding	0,35	1,25	0,40	0,94	0,70	0,75	0,32	0,73	0,02
2	Dialogue management	0,95	1,01	1,10	0,83	0,75	0,51	0,42	0,66	0,03
3	Information pooling	-0,15	1,18	0,40	1,07	1,00	0,63	3,04	0,06	0,18
4	Reaching consensus	-0,65	0,74	0,35	1,07	0,75	0,72	6,33	0,01	0,32
5	Task coordination*	0,30	1,10	0,65	0,71	0,55	0,82	0,37	0,70	0,03
8	Reciprocal interaction	1,00	0,89	1,20	0,46	1,20	0,60	0,26	0,77	0,02
9	Contribution balance*	1,20	0,60	0,30	0,64	0,40	0,92	4,08	0,03	0,23
Sum of points		3,00		4,40		5,35				

In an overall comparison the groups with gaze sharing support achieve better results in terms of collaboration quality in comparison to the No Gaze group. The collaboration quality is therefore to be conditionally considered higher. Likewise, it can be seen that the Gaze Awareness group's collaboration quality is higher than the Gaze Cursor group.

The analysis of variance reflects the results of the dimensions in the comparison group. It can be seen that only the dimensions of Reaching consensus and Contribution balance vary significantly. The dimension of information pooling is almost significantly different with p = 0.06. Comparing the sums of the dimensions among the groups, it can be seen that the Gaze Awareness group performs best. However, this result is not to be interpreted as statistically significant ($F(2,18) = 0.90$; $p = 0.42$, $\eta2 = 0.09$), which is due to the small number of participants.

Hypothesis 2: To measure the frequency of Shared Gaze events in each group we used the computational data as described in Method 2. As Fig. 3 shows, the gaze sharing groups generated at least one more Shared Gaze event per minute compared to the No Gaze group. The average Shared Gaze event duration is nearly identical in all groups (No Gaze: 0,44 s; Gaze Cursor: 0,46 s; Gaze Awareness: 0,44 s).

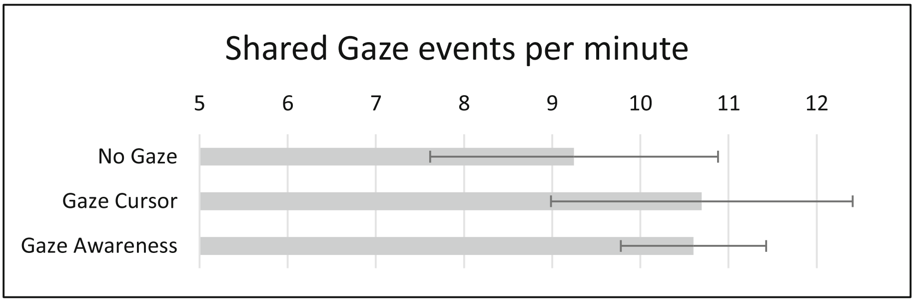

Fig. 3. Average shared gaze events per minute for each condition with standard deviation. The minimum length for each event was set to the mean fixation duration of all participants.

Hypothesis 3: The pupillometry described in Method 3 was used to estimate the cognitive load. Figure 4 shows the comparison of the mean pupil diameter of all participants grouped by the three conditions.

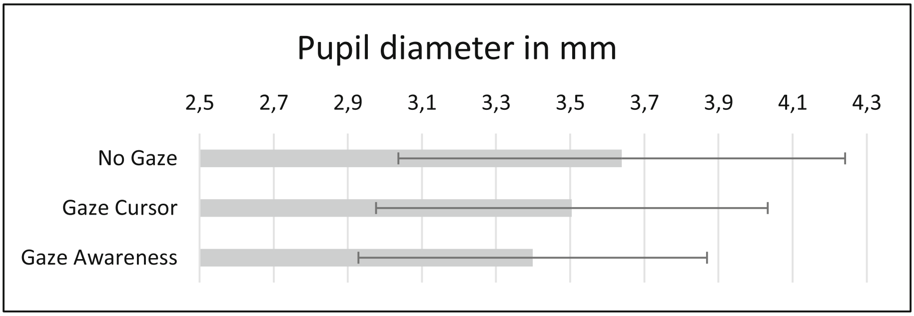

Fig. 4. Estimation of cognitive load by pupil diameter in mm with standard deviation.

The Gaze Awareness group shows a lower amount of cognitive load compared to the Gaze Cursor and No Gaze group ($F(2,27) = 26,53$; $p = 4,6E\text{-}12$; $\eta2 = 0,032$).

5 Discussion and Conclusion

Hypothesis 1: The quality of collaboration increases with the aid of mutual gaze data transmission. It is not important whether the participants are supported by the Gaze Cursor or the Gaze Awareness. Both visualizations achieve better results compared to the No Gaze group.

Hypothesis 2: The analysis of Shared Gaze events showed that the gaze sharing groups had a slightly higher Shared Gaze frequency as the No Gaze group which confirms the hypothesis. The gaze sharing groups generated more short Shared Gaze events in the beginning of the trials which were also sorted out by the minimum threshold of the mean fixation duration which was initially introduced to filter accidental gaze collisions. This is probably due to the fact, that all of the participants never used an eye-tracker prior to the study and had to get used to the gaze cursor and gaze awareness feature. We assume that those short Shared Gaze events are not part of the problem solving process and could be safely removed.

Hypothesis 3: Due to the constant movements of the Gaze Cursor, which reflexively draws attention to itself, an increased cognitive load in the Gaze Cursor group is measured in comparison to the Gaze Awareness group. The No Gaze group has, as expected, the highest amount of cognitive load because the coordination effort is probably the highest due to the lack of the additional information channel.

In summary, it should be noted that all three hypotheses have been confirmed. The context-based gaze sharing method achieved comparable positive effects on the collaboration as the coordinate-based gaze sharing, with the advantage of a slightly less cognitive load. Thus, a use of gaze sharing outside of the WYSIWIS domain is possible. Due to the limited scenario as well as the relatively small amount of participants no generalization should be made here. However, the results are promising and should be elaborated on in larger studies.

References

1. Reinmann-Rothmeier, G., Mandl, H.: Analyse und Förderung kooperativen Lernens in netzbasierten Umgebungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychol. **34**(1), 44–57 (2002)
2. Dillenbourg, P., Järvelä, S., Fischer, F.: The evolution of research on computer-supported collaborative learning. In: Balacheff, N., Ludvigsen, S., de Jong, T., Lazonder, A., Barnes, S. (eds.) Technology-Enhanced Learning, pp. 3–19. Springer, Dordrecht (2009)
3. Richardson, D.C., Dale, R.: Looking to understand: the coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. Cogn. Sci. **29**(6), 1045–1060 (2005)
4. Schneider, B., Pea, R.: Real-time mutual gaze perception enhances collaborative learning and collaboration quality. Int. J. Comput. Support. Collab. Learn. **8**(4), 375–397 (2013)

5. Pietinen, S., Bednarik, R., Tukiainen, M.: Shared visual attention in collaborative programming: a descriptive analysis. In: Proceedings of the 2010 ICSE Workshop on Cooperative and Human Aspects of Software Engineering, pp. 21–24. ACM, New York (2010)
6. Nüssli, M.-A.: Dual Eye-Tracking Methods for the Study of Remote Collaborative Problem Solving. École Polytechnique Fédérale de Lausanne (EPFL), Lausanne (2011)
7. Stefik, M., Bobrow, D.G., Foster, G., Lanning, S., Tatar, D.: WYSIWIS revised: early experiences with multiuser interfaces. ACM Trans. Inf. Syst. (TOIS) **5**(2), 147–167 (1987)
8. Müller, R.: Blickbewegungen in der computermediierten Kooperation. Chancen und Schwierigkeiten in der Verwendung einer weniger expliziten Kommunikationstechnologie. unveröffentlichte Dissertation, Technische Universität Dresden (2012)
9. Mühlenbrock, M., Hoppe, U.: Computer supported interaction analysis of group problem solving. In: Proceedings of the 1999 Conference on Computer Support for Collaborative Learning, p. 50 (1999)
10. Meier, A., Spada, H., Rummel, N.: A rating scheme for assessing the quality of computer-supported collaboration processes. Int. J. Comput. Support. Collab. Learn. **2**(1), 63–86 (2007)
11. Kahrimanis, G., Meier, A., Chounta, I.-A., Voyiatzaki, E., Spada, H., Rummel, N., Avouris, N.: Assessing collaboration quality in synchronous CSCL problem-solving activities: adaptation and empirical evaluation of a rating scheme. In: Cress, U., Dimitrova, V., Specht, M. (eds.) EC-TEL 2009. LNCS, vol. 5794, pp. 267–272. Springer, Heidelberg (2009)
12. Holmqvist, K.: Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press, Oxford (2011)

http://www.springer.com/978-3-319-22746-7

Collaboration and Technology
21st International Conference, CRIWG 2015, Yerevan, Armenia, September 22-25, 2015, Proceedings
Baloian, N.; Zorian, Y.; Taslakian, P.; Shoukourian, S. (Eds.)
2015, XII, 249 p. 74 illus., Softcover
ISBN: 978-3-319-22746-7