Impact of Visualizing Teammates' Actions on Shared Situational Awareness in Virtual Reality

Teresa Gómez Mena

Supervisors: Marcus Specht, Nesse van der Meer

EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 24, 2023

Name of the student: Teresa Gómez Mena
Final project course: CSE3000 Research Project
Thesis committee: Marcus Specht, Nesse van der Meer, Ricardo Marroquim

Abstract

Background: Despite Virtual Reality being a relatively new field, it is steadily being introduced into numerous disciplines, such as education. Within these systems, particularly in collaborative environments, it is crucial to have a high Share Situational Awareness (SSA) in order to be aware of ones' surroundings, encourage efficient decision making and effective team coordination. This project aims to study the change in the SSA of a group of players by allowing them (or not) to visualize their teammates' actions.

Method: For this scope, an experiment was carried out: a VR game was played twice by two groups of three participants, at a distance of roughly two weeks, testing both visualizing and not visualizing their teammates' actions. During such event, their SSA was measured and analyzed through two techniques; namely Situation Awareness Rating Technique (SART) and Situational Awareness Linked Indicators Adapted to Novel Tasks (SALIANT).

Results: SART and SALIANT produced opposing viewpoints. According to the former, both groups showed higher degrees of SSA in their second sessions. According to the latter, both groups' first session had the highest level of awareness. Both of these results were unrelated to the addition of actions' visualizations. Possible explanations include the players' confidence rising in the second session (biasing SART), and the encoders not having an accurate inter-reliability analysis for SALIANT.

Conclusions: Overall, the experiment produced inconclusive results on the impact of visualizing teammates' actions on the groups' SSA. Nonetheless, some recommendations for future research may be made: to choose the SSA measurement tool in accordance to the layout of the system that will be examined, to extend the time between sessions, and to ensure the system is suited for the experiment.

1 Introduction

In complex setups that change quickly and offer a lot of information to keep up with, Situational Awareness (SA) is essential. In fact, people are less aware of their visual environment than they believe they are [1]. In order to guarantee that a collaborative environment promotes effective team coordination, it is crucial to create systems that allow for a high SA across team members.

More extensively, Situational Awareness (SA) is a phenomenon in which team members are more aware of each other's actions and the overall situation [2]. This can be applied to the group as a whole by considering their Shared Situational Awareness (SSA), which concerns "the degree to which every team member possesses the SA required for his or her responsibilities" [3]. Being aware makes it possible to finish tasks efficiently, that would otherwise require a team of experts to complete them. It also promotes informal social interactions and the growth of shared working cultures, both of which are crucial for preserving positive relationships in a team [1].

An example of a complex system where high SA is beneficial is Virtual Reality (VR). VR is defined as "the use of computer modeling and simulation that enables a person to interact with an artificial three-dimensional (3-D) visual or other sensory environment" [4]. Despite being a relatively new field, the positive impact of VR is being slowly included into multiple disciplines, such as medicine [5] or education [6]. VR has, in fact, been proven to augment the communication and collaboration between users [7] [8].

The purpose of this project is to determine what is the impact of visualizing teammates' actions during a VR game on their SSA. More specifically, this paper will answer the question "do visualizations of activities have an effect on the Shared Situational Awareness of group members inside Virtual Reality?". This query will be tackled by setting up an experiment in which a group of participants will play a VR game twice, testing both visualizing and not visualizing their teammates' actions, through vision cones and pinpointing. During such event, in which players will have to communicate to exit a maze by solving color puzzles, their SSA will be measured and analyzed. This paper will include a discussion of the related works, an in depth description of the experimental set up, an illustration of its ethical concerns and an interpretation of its results.

2 Related Works

This section will discuss existing research that relates to the main topics of the research question: visualizations, VR and SA.

Previous research has been made regarding the effect of visualizations in learning and team collaboration. As this occurs often in gaming environments, an experiment has been performed to test how useful different groupware is when playing a collaborative game. It was discovered that most players of Dota 2, a popular Multiplayer Online Battle Arena game, considered some groupware essential for winning, and used them frequently within planning or emoting, between other motivations. [9]. In addition, it has been proved that visualizations applied to resource searching, collaboration, reflection, and instructional design have the potential to help shape the learning process and encourage reflection on its progress and impact [10].

In regards to visualizations specifically in VR, players communicating with external users particularly enjoy, between multiple ways of communicating, pointing and vision cones [11]. In fact, also for other tasks, such as sharing an online remote space, adding augmented visual communication cues can improve the experience of collaborating together as well as the sense of being together. More specifically, pointing is the preferred additional cue, rather than voice or annotations [12].

Concerning SA, by performing an experiment on groups of people working together and analyzing their overall understanding, it has been discovered that their domain expertise and background experience has an influence on the individual and shared SA. The participant's qualifications should, therefore, be taken into account when analyzing their awareness [13].

These discoveries encourage the first hypothesis of this project to state that visualizing teammates' actions, while playing in VR, will lead to a higher SSA. Nonetheless, this paper will tie together all the above studies and discoveries by further studying how visualization improves learning and teamwork in a virtual setting.

3 Methodology

The following section will describe the steps taken to approach the research question by discussing the participants recruited, the materials used and the procedure followed.

3.1 Participants

The experiment included the need of two experimental groups, each composed of three human subjects. Due to the involvement of participants with different backgrounds, it is important to keep in mind any factor that could impact the group's SSA. Some examples include the following:

- Since VR is still a relatively new technology, the majority of people have never experienced it. For this reason, the first moments of the game may be confusing to the users, due to their possible lack of experience with the program. This might make users less confident, which could affect their actual awareness [14].
- As this experiment concerns the players' communication, it is important to understand that people communicate in different ways: some could be more talkative than others, or more confident, which could bias the results.
- Language barriers might be a challenge, as the experiment will be conducted in English with potentially no native English speakers.

Out of this sample of factors that could influence the final results, together with the competences needed to play the game, certain requirements for the participants were created:

- They must not be color blind, as most puzzles from the VR maze used for the experiment involve colors.
- Participants from the same group must not be friends, as that could affect their ways of communication.
- They must be able to communicate in English.
- They must not be prone to motion sickness or claustrophobia, as the VR maze is made of corridors.

These participants were recruited through personal connections of the research group and were informed of the experiment throughout an online Microsoft form. Their degree of experience was noted in order to address these additional issues in the analysis of the results.

3.2 Materials

This subsection will present the materials used in the project.

Software and Hardware

The system used for the experiment is a custom-made, multiplayer VR maze [15]. Within the game, the players have to collaborate to find the exit. The maze contains a series of obstacles that the users can overcome only with the help of their teammates, as it is shown in Figure 1. All obstacles concern color puzzles, in which each player can only perceive one color: red, blue or yellow.

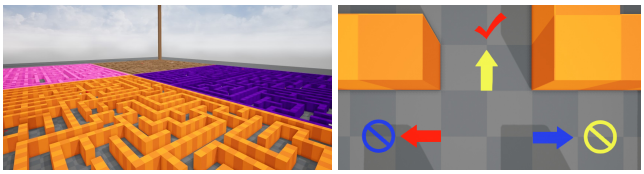

Figure 1: Top overview of the maze (left) and example of a puzzle (right), where each participant can only see one color. Participants are expected to understand which way to go (top) by communicating what they see.

This software allows for a collaborative game within three players, either in an experimental or in a controlled condition. In the former condition players have the ability to pinpoint interesting elements of their environment and are aided by vision cones that indicate the viewing area, whereas in the latter they do not, forcing them to communicate to point out elements to their teammates. The difference between the two conditions is illustrated in Figure 2.

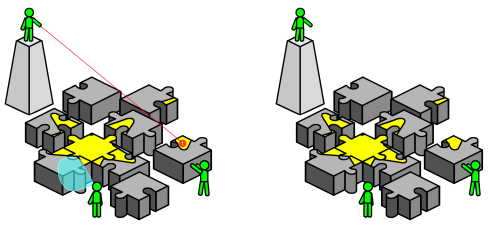

Figure 2: Example of experimental condition (left), containing vision cones and pinpointing, and controlled condition (right), without visualizations. Note that these images do not represent the VR game used for the main experiment of the project.

The hardware used for the experiment involved three VR Head-Mounted Displays (HMDs) from High Tech Computer (HTC) Vive, connected to the computers where the software was running, and two HTC Vive controllers, that allowed participants to move inside the virtual environment through pointing and clicking.

SSA measurement techniques

Multiple techniques exist to analyze SSA but, for the scope of this project, a low cost, non-intrusive and easy to use method for the analysis was required. These conditions reduced the ideal techniques to SART (Situation Awareness Rating Technique [16]) and SALIANT (Situational Awareness Linked Indicators Adapted to Novel Tasks [17]).

SART is a subjective rating approach which uses ten dimensions, all related to awareness (i.e. familiarity, concentration, spare mental capacity, etc.) to measure individual SA. Participants are asked to rate, half way through the experiment, each dimension on a seven-point scale (1 being the lowest and 7 being the highest) based on how well they are performing the studied task. The ratings are analyzed as follows:

$$Understanding = Q1 + Q2 + Q3$$
$$Demand = Q4 + Q5 + Q6 + Q7$$
$$Supply = Q8 + Q9 + Q10$$
$$SA = Understanding - (Demand - Supply)$$

The average of the individual ratings will give the SSA of the group. This method is non-intrusive, as it only stops the experiment once, allowing it to proceed smoothly without distractions that could affect the participant's awareness.

As for SALIANT, it is an objective methodology that connects observable behaviour to SSA. In addition to also being non-intrusive, SALIANT was picked because it is one of the few tools that focuses primarily on the team's SSA. In fact, this approach is based off the following steps:

1. Identify behaviours that demonstrate awareness.
2. Define the acceptable, or unacceptable, responses for the behaviours.
3. Develop general scenarios that correlate to the task.
4. Establish a table that includes, per behaviour and per scenario, the acceptable responses.
5. Use this table to rate the group's SSA.

A more in depth explanation of how this process was used for the experiment can be found in (subsection 3.3), and a small sample of a possible SALIANT grading table can be seen in Table 1. The final table used for this project was given by its supervisor, and will be used to analyze the SSA of the participants, together with the SART questionnaire.

Behaviour	Scenario: Being lost or stuck
Reporting problems	Participant acknowledges they are lost
Solving problems	Participant finds the correct way
Briefing status	Participant checks if group is together

Table 1: Small example of a grading SALIANT table, representing a scenario in which a group gets lost or stuck while playing a game. This table is not representative of the table used for the scope of this project, but is rather a generic example.

Questionnaires

The participants filled a total of three questionnaires. Two questionnaires are given to them at the start of the experiment: an informed consent form and a personal background survey, in which their domain knowledge and their previous experience with VR are noted, as well as their gender and age. In addition, as explained above, mid-way through the experiment the participants receive another questionnaire, containing the ten SART dimensions to rate.

3.3 Procedure

This subsection will present how the experiment was carried out and how the data was collected and analyzed.

Experiment

As stated in subsection 3.1, two groups of three completed the experiment. Each group did the experiment twice: one session with visualizations (the experimental condition) and one without (the controlled condition), at a distance of roughly 14 days, with slightly different mazes and puzzles. One trio started as experimental, and the other as controlled, as this allows for further comparison between the sessions.

The experiment took roughly 45 minutes to be completed. At the arrival of the participants, they were informed of the structure of the experiment and asked to complete the personal information and informed consent questionnaire. Once completed, the HMDs were collocated into the participants and the game started.

The game runs for 30 minutes, in which the participants need to find the exit of the maze by solving all the color puzzles. The experiment was interrupted once at the 15 minute mark. During the break, the moderators checked with the participants for any physical issues, and the participants were asked to fill in the SART survey.

Data collection

The data was collected by recording the point of view and the audio of each participant inside of the virtual environment, as well as through the questionnaires.

Data analysis

The SART questionnaire and the SALIANT technique were used to measure the SSA. Regarding SART, the formulas presented in subsection 3.2 resulted in individual SA scores, per teammate, that were averaged to get the SSA of the group. Regarding SALIANT, the following steps were performed on

the transcription of the dialogues together with the grading table, also explained in subsection 3.2 :

1. **Segmentation of the transcripts:** dialogues were segmented into what deem separate topics of conversation.
2. **Inter-reliability of the segmentation:** because the segmentation was divided amongst multiple encoders, an inter-reliability analysis was necessary to find a common ground and avoid extreme differences in the encoding.
3. **Scenario scanning:** each segment was scanned for instances of four possible scenarios:
   (a) Markings (symbols or text) on the floor (to guide participants)
   (b) Deciding which path to take when faced with multiple
   (c) Discussing the pass-code at gates
   (d) Participants are lost, they might be backtracking

   And every instance of a scenario is counted.
4. **Scoring:** for each instance of a scenario, every participant received a score based on the behaviour categories of the SALIANT table. This process gave a participant a 0, if there were only incorrect responses, or a 1, if there was at least one acceptable response within the instance. The general overview of the categories considered is the following:
   (a) Demonstration of Awareness of Surrounding Environment
   (b) Recognition of Problems
   (c) Anticipation of Need for Action
   (d) Demonstration of Knowledge of Tasks
   (e) Demonstration of Awareness of Information
5. **Averaging:** the sum of the scores of each participant was divided by the total number of instances in the four scenarios, giving the final SA score per participant. In order to get the SSA, the average was taken, by summing the results together and dividing by the number of participants.

The results given by these two techniques, SART and SALIANT, are discussed in section 5.

4 Responsible Research

This section will deal the with the ethics concerning the experiment, particularly the use of human participants, and its reproducibility.

4.1 Ethics

Because this project involved human subjects playing in VR and the collection of data, an ethics analysis of the possible risks and how they were mitigated is necessary. Nonetheless, a Human Research Ethics Committee (HREC) form was compiled by the supervisor of this project before the involvement of the research group to ensure a correct contribution.

Regarding the human subjects, the participants were informed of the experiment throughout an online Microsoft form, that also asked for their consent to share their name to the research group. At the experiment itself, a consent form was handed to them. Within this form, signed by each candidate, they were informed of the following:

- **All possible risks involving the experiment:** motion sickness, claustrophobia or other physical safety concerns.
- **Procedure for withdrawal of the study:** participants are allowed to quit their participation in this study at any time with no negative consequences.
- **Data collection and privacy:** although the experiment collects video and audio of the virtual environment during the game, as well as gender, age and personal experience, this data is anonymized, not usable for identification and analyzed only within the study group.

4.2 Reproducibility

The other important ethical concern regarding this paper's responsible contribution to science is the reproducibility of the main experiment. The experiment has been thoroughly described in this paper but, although all the necessary information is publicly available, the software of the maze is still in publication and the participants are anonymous. As a matter of fact, the experiment is reproducible only with a different system and different participants. It is to note that this will cause the results to vary; new participants may have different background and previous experience, and a different system may not have the same interface design, affecting the overall awareness. Nevertheless, this paper can be used as a guide to reproduce the experiment on other systems.

5 Results

This section will display the information of the participants, as well as the results obtained from the experiment's sessions.

5.1 Participants

All participants compiled a questionnaire giving their personal background (gender and age), as well as their degree of expertise in the VR and gaming field (weekly frequency of gaming and / or VR usage). Table 2 describes the personal data collected from the six participants:

ID	Age	Gender	Gaming	VR
1	25	Male	Often	Rarely
2	22	Female	Rarely	Rarely
3	21	Male	Rarely	Regularly
4	20	Female	Rarely	Rarely
5	21	Male	Rarely	Never
6	25	Male	Often	Sometimes

Table 2: Personal background of the participants

Group 1 is composed by candidates ID1, ID2 and ID3, while Group 2 is composed by ID4, ID5 and ID6. Both groups are balanced equally in regards to gender, and similarly in regards to gaming experience.

5.2 Individual scores

In session one, both the controlled and the experimental group completed roughly 3/4 of the maze. All participants compiled the SART questionnaire mid way through the experiment, and their dialogues were later on analyzed with the SALIANT method. These techniques provided the following information displayed in Table 3.

Condition	ID	SART	SALIANT
Controlled	1	4	0.095
	2	12	0.084
	3	-4	0.068
Experimental	4	-2	0.211
	5	1	0.263
	6	5	0.223

Table 3: Individual SA scores on session one

Compared to the first session, both groups were able to navigate the maze more in the second session. More specifically, Group 1 nearly finished the maze, whereas Group 2 advanced little more from the previous attempt, still only completing around 3/4 of the maze. Table 4 contains the specifics regarding the SA of the players.

Condition	ID	SART	SALIANT
Experimental	1	3	0.085
	2	9	0.083
	3	4	0.064
Controlled	4	5	0.074
	5	3	0.102
	6	23	0.094

Table 4: Individual SA scores on session two

5.3 Group results

The final SSA scores per session, derived from the average score of the participants, are described in Table 5. Overall, the average SSA value according to SART is 5.225 (maximum 10.3, minimum 1.3), while according to SALIANT is 0.1205 (maximum 0.263, minimum 0.077). More extensively, Figure 3 shows, per group, the mean and standard deviation of the individual scores of SART and SALIANT.

Group	Session	Condition	SART	SALIANT
1	1	Controlled	4	0.083
	2	Experimental	5.3	0.077
2	1	Experimental	1.3	0.232
	2	Controlled	10.3	0.090

Table 5: Group 1 and Group 2 SSA results from session 1 and 2.

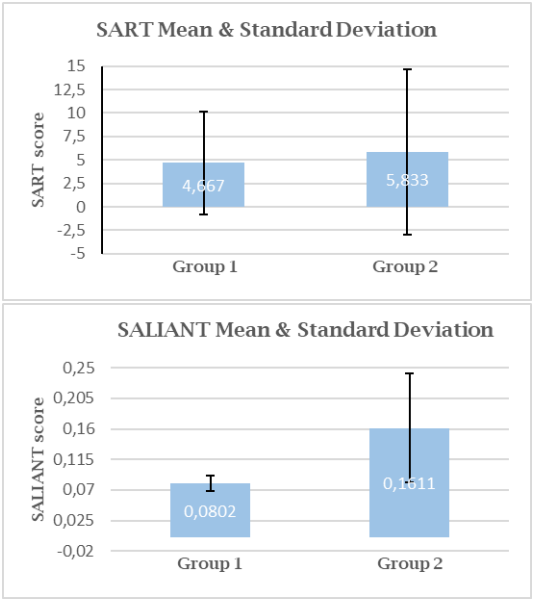

Figure 3: SART and SALIANT SA mean (blue bars) and standard deviation (lines) of Group 1 and Group 2, throughout both the experimental and the controlled condition.

6 Discussion

Within this section, results will be analyzed in regards to the two conditions, the two techniques used to analyzed SSA and other external factors that could have influenced them. It is important to note that the discussion of these results will mostly be descriptive and not statistical, due to the reduced size of the participants, which would not allow for a reliable analysis.

6.1 Experimental vs Controlled

According to SART, both of the groups participating in the experiment had overall higher levels of SSA in their second sessions. In fact, Group 1's awareness rose from 4 to 5.3, and Group 2 from 1.3 to 10.3. Contrary to the initial hypothesis stated in section 2, this rise was independent of either the experimental or the controlled condition; Group 1 had the greatest value in the experimental condition, while Group 2 had the highest value in the controlled condition. SART is a subjective measure, therefore, the increase on the second session may have been caused by a rise in the participants' confidence. This outcome is further confirmed by the fact that, regardless of the condition, both groups were able to navigate more of the maze during their second session.

SALIANT resulted in opposite results: the first session of both groups had the highest awareness, with Group 1 being more aware in the controlled experiment (0.0829 versus 0.0775), and Group 2 in the experimental one (0.232 versus 0.090). A possible explanation states that, during their first session, players tended to discuss more about the new things in their environment; while, in their second session, they were already familiar with the game and therefore communicated less and were more direct. SALIANT is, in fact, an objective approach that is solely based on the quality and quantity of the players' dialogue.

These results prompt to think that either the time passed between session one and session two was not enough, or that the environments were too similar; participants seemed more confident and familiar with the maze's system in session two, causing a skew in the results.

6.2 Limitations

Multiple external factors that could have affected the results should be considered. More precisely, the reliability of the techniques used to measure SSA and the varied backgrounds of the participants could have influenced the experiment.

Firstly, it may be that SART and SALIANT were not the ideal methods for this experiment. Despite being inexpensive and non-intrusive, both approaches had drawbacks. As a matter of fact, even though SALIANT underwent an inter-reliability examination, the encoding may have been still slightly uneven, resulting in the participants' grades not being assigned equally. In addition, because SART is a subjective measure, the confidence of the candidates plays a significant role. For instance, participant ID6, who had stated that he often plays video games, sometimes in VR, indicating that he is familiar with the concepts and tasks, gave himself the highest SART score during the second session (23 points in comparison to 3 and 5 of his teammates), despite not being the most aware according to SALIANT.

Secondly, as stated above, it is important to take into account the participants' backgrounds and personalities. Not all participants communicated in the same manner; players who didn't interact as much with their teammates had a detrimental impact on their SALIANT ratings, despite their actual awareness being higher. Additionally, in the experimental setting, some individuals used the pinpointing function less frequently than others, which made it more challenging for their colleagues to comprehend what they were trying to highlight. It could be that the maze was not designed to prompt players to communicate enough and, therefore, was not fully suitable for this experiment.

These ideas partly explain why the results do not correspond to the initial hypothesis, and encourage to believe that the techniques used to analyzed SSA were not ideal. The following section will track the final conclusions of the experiment.

7 Conclusions and future recommendations

Overall, this research paper aimed to analyze the impact of visualizing teammate's actions on their Shared Situational Awareness while playing in Virtual Reality. This study was tackled with one main experiment, in which groups of participants collaborated to exit a Virtual Reality maze, while their awareness was measured in two ways: first, through a subjective SART questionnaire and, second, through the SALIANT analysis of their dialogues.

The experiment did not demonstrate a difference in group awareness between visualizing teammates' activities and not, disproving the initial hypothesis. In fact, each measuring technique resulted in different outcomes; according to SART, the groups awareness was highest during the second session, while SALIANT concluded the opposite. Both techniques' results were independent from the condition of the session.

Concluding results were therefore not achieved, but possible causes have been discussed: wrong choice of SART and SALIANT, inexact inter-reliability analysis of SALIANT, design of the maze, not enough time passed between the sessions and background of the players. These factors can be transformed into future recommendations for the reproduction of this research. More specifically, it is recommended to choose the Situational Awareness measurement tool in accordance to the layout of the system that will be examined, as well as to extend the time between sessions, significantly alter the environment or, simply, repeat the experiment only once with an increased sample size.

References

[1] O. et al. Kulyk. Situational awareness in collaborative work environments. *Information Science Reference*, 2009.

[2] M. R. Endsley and D. G. Jones. Designing for situation awareness: An approach to user-centered design, second edition. 2016.

[3] M. R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 1995a.

[4] H. E. Lowood. Virtual reality, encyclopedia britannica. https://www.britannica.com/technology/virtual-reality, June 2023.

[5] R.M. Satava and S.B. Jones. Current and future applications of virtual reality for medicine. *Proceedings of the IEEE*, 86(3):484–489, 1998.

[6] Bambang Robiin Yovi Sepriansyah Achmad Soleh Dwi Sulisworo, Vera Yuli Erviana. The feasibility of enhancing environmental awareness using virtual reality 3d in the primary education. *Education Research International, vol. 2022*, 2022.

[7] G. C. et al. Dobre. More than buttons on controllers. *CM International Conference on Intelligent Virtual Agents (IVA '22), September, 2022, Faro, Portugal*, 2022.

[8] L. Zheng, T. Xie, and G. Liu. Affordances of virtual reality for collaborative learning. *Proceedings - International Joint Conference on Information, Media and Engineering*, 2019.

[9] Bateman S. Wuertz, J. and A. Tang. Why players use pings and annotations in dota 2. *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, page 1978–2018, 2017.

[10] Verbert K. Duval E. Klerkx, J. Enhancing learning with visualization techniques. *Handbook of Research on Educational Communications and Technology. Springer, New York, NY.*, 2014.

[11] Priyanka Pazhayedath, Pedro Belchior, Rafael Prates, Filipe Silveira, Daniel Simões Lopes, Robbe Cools, Augusto Esteves, and Adalberto L. Simeone. Exploring bidirectional pinpointing techniques for cross-reality collaboration. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 264–270, 2021.

[12] Seungwon Kim, Gun Lee, Nobuchika Sakata, and Mark Billinghurst. Improving co-presence with augmented visual communication cues for sharing experience through video conference. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 83–92, 2014.

[13] Bolstad C. A. Gonzalez C. Cuevas H. M. Saner, L. D. Measuring and predicting shared situation awareness in teams. *Journal of Cognitive Engineering and Decision Making*, 3:280–308, 2009.

[14] Olivier Compte and Andrew Postlewaite. Confidence-enhanced performance. *American Economic Review*, 94(5):1536–1557, December 2004.

[15] Meer, N. van der, Brinkman, W.-P., Specht, M. . How visualisations of actions affect social modes of co-construction inside virtual environments. Manuscript in preparation, 2023.

[16] Situation awareness rating technique (sart). https://ext.eurocontrol.int/ehp/?q=node/1608. Accessed: 2023-05-03.

[17] Situational awareness linked indicators adapted to novel tasks (saliant). https://ext.eurocontrol.int/ehp/?q=node/1606. Accessed: 2023-05-03.