

The Importance of Eye-contact for Collaboration in AR Systems

Erik Prytz*

Santa Anna IT Research Institute
Linköping, Sweden

Susanna Nilsson†

Swedish Defence Research Institute
Linköping, Sweden

Arne Jönsson‡

Santa Anna IT Research Institute
Linköping, Sweden

Abstract

Eye contact is believed to be an important factor in normal human communication and as a result of this a head mounted display (HMD) is often seen as something intrusive and limiting. This can be especially problematic when AR is used in a collaborative setting. The study presented in this paper aims to investigate the effects an HMD-based AR system can have on eye contact behaviour between participants in a collaborative task and thus, in extension, the effects of the HMD on collaboration itself. The focus of the study is on task-oriented collaboration between professionals. The participants worked through three different scenarios alternating between HMDs and regular paper maps with the purpose of managing the crisis response to a simulated major forest fire. Correlations between eye contact between participants and questionnaire items concerning team- and taskwork were analysed, indicating that, for the paper map condition, a high amount of eye contact is associated with low confidence and trust in the artefacts used (i.e. paper map and symbols). The amount of eye-contact in both conditions was very low. It was significantly higher for conditions without HMDs. However, the confidence and trust in the artefacts was generally rated significantly higher with HMDs than without. In conclusion, the decrease in eye contact with HMDs does not seem to have a direct effect on the collaboration in a professional, task-oriented context. This is contrary to popular assumptions and the results are relevant for future design choices for AR systems using HMDs.

Index Terms: H.5.1 [INFORMATION INTERFACES AND PRESENTATION]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 Introduction

Augmented Reality (AR) is an information medium that carries with it, just like any other medium, both strengths and weaknesses. One strength is that it can be used with Head Mounted Displays (HMDs) to provide a constant information overlay on the real world, and that this may enhance the users' understanding of a given situation, without restricting the use of arms and hands. A second strength is that it can be used to support collaboration in a team with mixed purposes and experiences, e.g. an emergency response team consisting of stakeholders from different agencies, as it allows for independence and individuality [3] meaning that each actor can independently have data tailored to her needs while working in a collaborative environment.

Ironically, these two strengths together form the basis of one of AR's weaknesses. That is, when using HMDs in a

team there may be negative consequences for team collaboration due to the impact on eye contact. A HMD obscures the eyes of the wearer and should thus lead to a lower level of eye contact between team members, which in turn will negatively affect team work. This weakness has so far largely been taken for granted in the research community, and this will be further discussed in Section 2.1.

The purpose of this paper is to investigate the effects of HMD-based AR systems on eye contact and how this may relate to and influence the workings of an emergency response team. Unlike many other studies on eye contact in collaborative settings, we are not interested in changes in communication patterns, e.g. turn taking, but rather in other concepts important for collaboration and teamwork. Concepts such as the users' perceptions of teamwork, trust, confidence and the like, and how they vary between AR and non-AR situations and where eye contact can be seen as an influencing factor.

This paper presents results that build upon further analyses done on a user study previously presented in [27, 25]. The user study included representatives from police, rescue service and military personnel, working jointly with the goal of coordinating work in a crisis situation. This means that the data in the study comes from professionals operating in an environment similar to one in which they would normally act. This provides a more realistic image of how a HMD-based AR system may influence eye contact, and in turn the collaboration, than using naive participants [22].

2 Related Work

Collaborative work has been studied extensively in many different research domains, from sociological and psychological perspectives as well as organisational perspectives. The use of technology as a tool to support collaboration has been studied in for example Computer Supported Collaborative Work (CSCW, see for example [33]), Distributed Cognition (see for example [14] or [30]), as well as Activity Theory (see for example [35] or [20]).

2.1 Head-Mounted Displays and Eye Contact

The effect of eye contact on communication has been studied in many different settings for a long time (see e.g. [18, 15]). Clearly, there are some well-documented effects of eye contact on communication in dyads or small groups, for example [18].

We do not wish to contend these findings, but rather to investigate whether or not any effects can be seen in task-oriented settings with professionals where HMDs are used.

The prevalent assumption of the AR-research community has so far been that it does. However, there is, to the best of our knowledge, no actual studies that have looked at this. On the other hand, research on solving the (possible) problem of lack of eye contact has been carried out. e.g. [4]. That study in particular motivated their research by saying that "the loss of gaze awareness is a critical problem when attempting to use the shared MR space for collaborative tasks" [4], but does not provide any references for the claim.

*e-mail:epryt001@odu.edu

†e-mail:susanna.nilsson@foi.se

‡e-mail:arnjo@ida.liu.se

Billinghurst and Kato [2] utilise a similar reasoning by saying that wearable computers "are ideal platforms" for collaborative AR-environments, while also claiming that "collaborators must be able to maintain eye contact and gaze awareness" in order to experience a seamless experience. However, it is also argued that HMD-based AR systems are beneficial to collaboration [2, 28]. One possible interpretation is that the positive effects of AR outweighs the potential negative effects of eye contact-impact.

2.2 Collaborative Command and Control

Another aspect to consider is the awareness of *team cognition* [10]. Gutwin and Greenberg [10] argue that team work, and thus collaborative work, depends heavily on real world interaction. In their paper, they argue that it is the situated nature of team work that enables people to successfully solve collaborative tasks, and that technological systems therefore also must provide workspace awareness. They define workspace awareness as "the up-to-the-moment understanding of another person's interaction with the shared workspace" [10, p. 5]. They divide the possible knowledge of a shared workspace into three dimensions: (1) conversation, gesture and intentional communication, (2) bodies and consequential communication, and (3) artefacts and feedthrough.

The first dimension is intentional on behalf of the sender, the second depends on the observer's ability to interpret the subtle signals sent out by the observed, the third ones are largely a consequence of the design of the artefacts in use. Gutwin and Greenberg [10] present a number of different techniques that can be used to provide feedthrough and transparency in distributed collaborative systems. Feedthrough is defined by Dix as "the mechanism of determining a person's interactions through the sights and sounds of artefacts" [10, p. 9], i.e. it is imperative that the participants can observe their own as well as the other participants gestures while using the technical artefacts, and also manipulate the same objects.

2.3 Collaborative Augmented Reality

AR research has illustrated many areas of use for single user applications, such as applications that provide the user with instructions, for assembling complex technical tools, or different game applications (for an overview see Azuma [1] or Haller *et al.* [11]). The AR system described in this paper was largely developed through experiences from user studies of a single user system in context [23, 24]. The results of these studies showed that AR has great potential as a way to give instructions on how to perform more or less complicated tasks in the health care domain. Other researchers have illustrated the use of AR in process industry and object assembly [34], training and education [7], mobile phones [13], mobile applications [29] etc.

The development of AR applications and solutions for several users is also an extensive field of research. Some of the earliest attempts of developing collaborative, multi-user AR applications were presented in the Studierstube projects [32]. Fuhrman *et al.* [8] presented an AR system for collaborative scientific visualisation with 3D interaction and customised views for several users. Billinghurst & Kato [3] presented a vision of shared space using AR technology, Henrysson *et al.* [12] developed a collaborative mobile phone application and since then several research groups have published papers that illustrate different ideas of merging AR with collaborative computing approaches. However, few of these attempts have studied collaborative AR in joint real

time operations such as the ones in emergency command and control work.

2.4 Using AR systems in Real Situations

Evaluating systems based on heuristics developed for computer-based applications may be common practice in the AR field, but there are few examples of studies on how users actually perceive the system in actual use situations [23, 24, 6].

As early as 1967, Drabek and Haas [5] argued for the importance of using what they referred to as "real groups" in experiments. "The first requisite for a realistic simulation is that a real group be utilised. Second, the type of task, activity, or demand placed on groups must be appraised. Third, the ecological setting in which a unit is located may significantly affect resulting interaction patterns" [5, pp. 342-343]. Similar arguments have been put forward by Samuracy and Rogalski [31] in their study of fire-fighting simulations (as in the case of the scenario used in this study), where Samuracy and Rogalski found important differences when comparing expert participants' (real fire fighters) behaviour with laymen in their study. Johansson *et al.* [16] has argued for the concept of evaluating novel technologies (such as the effect of HMDs on eye contact) by combining a representative task, such as a micro-world (like the C³Fire simulation used in this study) with "professional users", i.e. users with domain expertise and experience. Such evaluations are not as powerful as tests performed in a real work setting, but many times it is the only option, especially when studying crisis management systems.

3 Method

The data for this study was collected in conjunction with a previous user study of the usability of a HMD-based video see-through AR system. That user study was previously described in [26] where more information on method and experiment design also can be found. The participants in that study (and thus also in this) were professionals from three different branches typically involved in emergency management; the police, the fire department, and the air force. Representatives from each of these branches collaborated in a dynamic emergency management task during the user study. Three simulated scenarios of forest fires were used, where the participants used either a paper map or an AR system when coordinating the response. As the participants were videotaped throughout the user study there was a rich material of professional users acting in realistic scenarios on collaborative command and control tasks to analyse for this study.

3.1 The AR System

The AR system used consists of three identical high fidelity AR prototypes¹, one for each experiment participant. Each of the three AR system's headsets consisted of a Z800 3Dvisor from eMagin (<http://www.3dvisor.com/>) integrated with a firewire camera. The headsets completely obscured the eyes of the participants and all visual perception was mediated by the camera and displays save for small gaps in the sides and bottom of the headsets. The visors could also be flipped open upwards so that the participants' field of vision was unobstructed.

The system runs on a Dell XPS M1330, with a 2.10 GHz processor, 3 GB RAM and with a 128 MB NVIDIA GeForce 8400M GS graphics card. The ARToolkit marker tracking

¹We will sometimes use 'AR system' to refer to the set of three AR prototype systems.

technology was used for tracking and registration [17]. Each AR system was functionally independent in relation to the others. The three systems were linked via an internal Ethernet network in order to provide a shared work space. In this experiment it was a map that was used for the crisis response task.

There were one set of symbols specific for each organisation that included specific vehicles (e.g. police units, helicopters, or fire trucks) and personnel units (e.g. K9 units, military units or unidentified persons). These different libraries could be masked so that a participant could choose to see units from the other organisations or not. Each organisation could access and use all libraries. Some symbols were not organisation specific, e.g. the fire symbol. The participants also had access to functions such as zoom or ① which gave more information on a unit.

The users thus have access to a digital "playground" where they can add symbols, move them or remove them freely. The symbols were placed in relation to a marker attached on a joystick, meaning that there was no fixed menu in the user's field of view or related to the map. Instead the menu of symbols was related to the joystick interaction device. In order to place a symbol the user first moves the joystick-attached marker to the chosen position on the map and then selects and places the symbol.

The participants shared either a paper or digital map. They could independently place the symbols previously mentioned using the handheld interaction device and were free to discuss with the others how to place their own symbols and the common symbols, such as the fire symbol and break points.

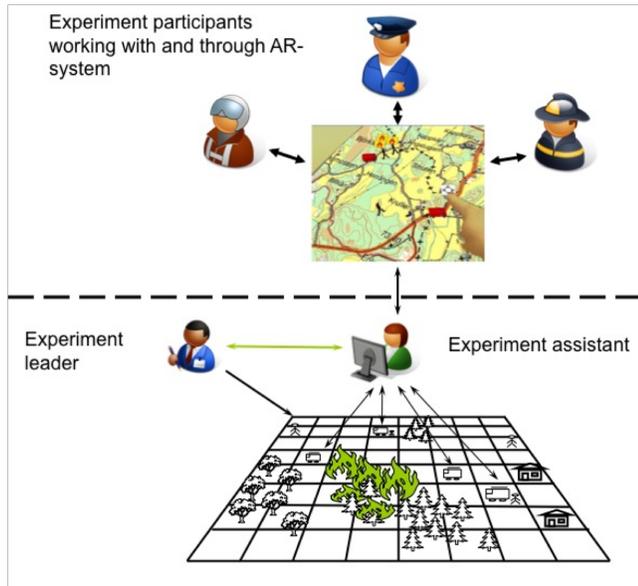


Figure 1: A schematic view of the C³ Fire gaming simulator used to create a dynamic and interactive scenario in the user study.

A computer-based micro world, C³Fire [9], was used to realistically simulate the fire and units involved. C³Fire is an advanced simulator that includes different types of terrain and structures, weather effects and can handle a large quantity of units and large fires. C³Fire was run by an experiment assistant on a laptop separate from the internal ethernet network used by the AR system. The experiment leader communicated changes, e.g. unit movements, weather

changes, or fire location to the participants and returned orders on unit movements to the research assistant. The experiment assistant also distributed paper slips with pre-defined events that occurred throughout the scenario to the experiment leader who brought the slips to the participants.



Figure 2: The gaming simulator that was controlling the input and responses to the participants was run by an assistant. The exercise leader worked as an information channel between the C³ Fire assistant and the participants.

The groups first trained with the AR system for 30 minutes before the three scenarios were played. While the scenarios themselves were balanced over time, the first always used the AR system, the second used a paper map and the third the AR system again. Each scenario took 20 minutes and was followed by an evaluation session where questionnaires were answered. The questionnaires used six-point Likert items as well as open-ended questions. The questionnaires included 15 closed response items and 6 open ended questions. A final AR system questionnaire was also used, distributed after the last AR-session, including 18 items and 10 open ended questions. The queries and data from the paper session and the second AR session are presented in Table 1². Analyses from the data in Table 1 are presented in [27, 25].

Each trial was filmed using five video cameras, three filming the AR screens and two filming the participants. Screen capturing software could not be used due to the high computational demands by the AR software. A schematic view of the experimental situation is seen in Figure 3. The two cameras used in the studies reported here are the ones on both sides of the AR systems' laptops table.

3.2 Participants

Eight small-team groups of three participants each participated in the study. All participants were professionals from the three different organisations (police, fire department, and the military). Each individual also had experience from within their organisation with team command and control situations as well as experience with working with at least one of the other two organisations. The individuals had never worked together before nor did they know each other before the teams were formed.

The original design included ten groups but two had to be excluded due to circumstances that comes with using professional actors from crisis response organizations. For instance, one experiment had to be excluded because one of

²The queries are translated to English by the authors.

Table 1: Cooperation questionnaire, average score and standard deviation from 6 point Likert scale items. As the statements in the questionnaire were both positively and negatively loaded (see for instance the first two items), the scores on the negatively loaded items were transformed in order to make the result easier to interpret. This means that in the table a high score is positive for the AR system/paper map and a low score is negative for the AR system/paper map.

| Request item | Paper | | AR | |
|-----------------------------------------------------------------------------|-------|----------|-------|----------|
| | μ | σ | μ | σ |
| 1. It took a long time to start to co-operate | 5.333 | 1.090 | 5.333 | 0,963 |
| 2. It was easy to co-operate | 5.087 | 1.474 | 4.696 | 1.550 |
| 3. I think that AR systems are good tools to use for co-operation | 4.250 | 1.391 | 4.417 | 1.136 |
| 4. The co-operation was fun | 5.208 | 0.833 | 5.292 | 0.690 |
| 5. I felt that the group controlled the situation | 5.000 | 0.885 | 4.792 | 1.141 |
| 6. It was easy to mediate information between the organisations | 5.000 | 0.834 | 4.833 | 0.917 |
| 7. The map made it easy to achieve a common situational picture | 4.167 | 1.404 | 5.041 | 0.999 |
| 8. The symbols made it easy to achieve a common situational picture | 3.417 | 1.472 | 4.833 | 1.050 |
| 9. The map became cluttered/messy | 2.375 | 1.469 | 4.000 | 1.474 |
| 10. I would have liked to have had more information than what was available | 3.167 | 1.659 | 3.292 | 1.459 |
| 11. I felt that I was certain that I could interpret what was on the map | 3.750 | 1.452 | 4.542 | 1.132 |
| 12. The map helped me trust the situational picture | 3.667 | 1.373 | 4.667 | 1.090 |
| 13. The symbols helped me trust the situational picture | 3.500 | 1.474 | 4.582 | 1.060 |
| 14. I though I had a good situational picture | 4.250 | 1.032 | 4.542 | 1.103 |
| 15. I thought the others had a good situational picture | 4.458 | 0.977 | 4.500 | 1.142 |

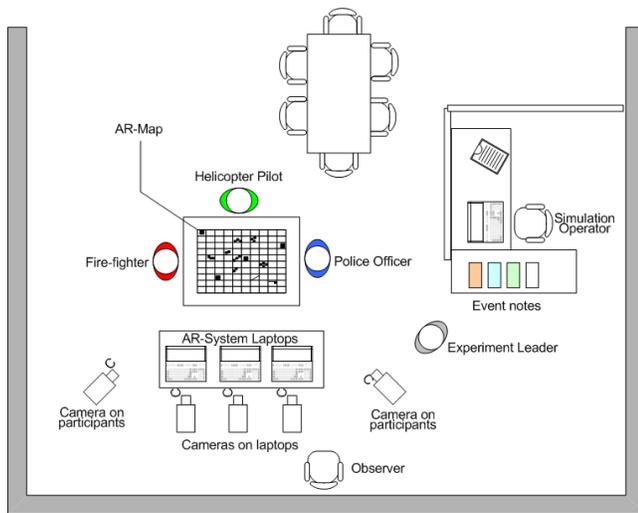


Figure 3: The experimental setting.

the participants was called away on active duty in the middle of the experiment.

There were 1 female and 23 male participants (24 total). The gender imbalance reflects the composition of the population from where the participants were recruited; all helicopter pilots in the area are male, the vast majority of the firefighters are male as well as the majority of the police. The average age was 39.1 years with a range from 25 to 57.

3.3 Video Analysis

16 videos of the 8 groups were analyzed. For each group the paper map condition and the second AR-condition were used, but for one group the second AR-videos were corrupted and the first AR-session was used instead³.

A total of 320 minutes and 20 seconds of video were analyzed. As the two video cameras did not capture all angles and the generally acknowledged difficulty for an observer

³For this group we still use the questionnaire items from the second AR session.

in determining exactly where a person is focusing his/her gaze [19], eye contact was defined as one individual looking in the vicinity of another individual's eyes e.g. the facial area. The individual looking was referred to as the subject and the person looked upon the object. The object did not have to respond and meet the gaze of the subject for it to count as eye contact, see Figures 5 and 6 where eye contact seeking behavior is considered established although only one person, the police officer, looks up. The reason for this classification is that HMDs are seen as restricting even one individual from seeking eye contact, not only when two individuals seek mutual eye contact.

While wearing the HMD, the participants could still have eye contact in some different ways. One alternative for the participant would be to open the HMD upwards which left the face and eyes unobstructed. This was however rarely used by the participants. Another method that the participants did utilize was to lean their head back and up and peering through the gap between the face and the bottom of the HMD. While not an intended design, this way of establishing eye contact was in fact used quite frequently by several participants. No distinction was made in the data concerning how the eye contact was established, e.g. through flipping the visor, removing the head mount completely or peering in gaps between the face the mount, in analyzing the videos and all types was included as long as they fulfilled the criterias outlined in the previous paragraph.

The videos were played at normal speed during the analysis. Enlargement tools were used to enhance the faces of the participants and aid in deciding where their gaze was focused. Whenever a potential eye contact was detected, the video was paused, backed up and then played frame by frame until the exact moment (in tenths and hundreds of a second) when the participant shifted his/her gaze toward another participant could be determined. The film was continued frame by frame until the exact moment of breaking the eye contact could be determined. The time period and number of eye contacts for each participant was recorded. An example using the enhancement tool to determine the start of the eye contact can be seen in the images in Figures 4 and 5.

There is only one frame between Figure 4 and 5, but the shift in the gaze of the police officer (to the right) can be



Figure 4: No Eye Contact.



Figure 7: End of Eye Contact.



Figure 5: Start of Eye Contact.

seen clearly. As the film was recorded at 25 FPS this means we can determine the start of the eye contact with an accuracy of 1/25th of a second. Likewise at the end of an eye contact, break could often be determined within one frame, see Figure 7.



Figure 6: During the Eye Contact.

4 Results

Table 2 presents data from the eye contact analyses. One participant was deemed to be an outlier and excluded due to abnormal data (outside of 1.5 inter-quartile range of all eye contact-data distributions), thus $N = 23$ for all tests performed.

Overall we can see that the amount of eye-contact is surprisingly low. On the average (Mean) we have 2% for the paper map, and 0.2% for the AR system. The most eye-contact is in the last paper map session with one participant having more than one minute of eye-contact and another participant having eye-contact for 49 seconds. Considering

that the whole session lasted for around twenty minutes, eye-contact at the most accounted for 5.6% of that time. For the AR-session the most eye-contact is 12 seconds, or 1%. Using the AR system we have six participants having no eye-contact at all.

Simple t-tests ($df = 22$ for all) were carried out to determine if there was any significant difference between the AR and paper map-conditions in terms of eye contact. The paper map condition was shown to have the most eye contact, both in terms of absolute eye contact moments, $t = 7.401, p < .001$ with 24.13 average eye contacts for paper map and 2.45 for the AR-condition, and in total percentage of time in the scenario spent in eye contact, $t = 5.895, p < .001$ with an average of 2.05% for the paper map condition and 0.21% for the AR-condition.

Correlations between eye contact and the questionnaire answers were then calculated using Spearman's R. This was done to see whether the amount of eye contact in any way was linked to the subjective experiences of e.g. confidence, teamwork, trust, etc. that was investigated by the questionnaire. "Amount of EC" refers to the number of eye contact occasions a participant had.

There is no correlation between the amount of eye contact in the AR-condition and any of the questionnaire items. There are however some correlations between eye contact in the paper map and the questionnaire items:

- There is a strong negative correlation between the amount of eye contact in the paper map-condition and question 8 on the questionnaire *The symbols made it easy to achieve a common situational picture*, $r = -0.684(21), p < .001$.
- There is a moderately negative correlation between the amount of eye contact in the paper map-condition and question 11 on the questionnaire *I felt that I was certain that I could interpret what was on the map*, $r = -0.467(21), p < .05$.
- There is a moderately negative correlation between the amount of eye contact in the paper map-condition and question 12 on the questionnaire *The map helped me trust the situational picture*, $r = -0.433(21), p < .05$.
- There is a moderately negative correlation between the amount of eye contact in the paper map-condition and question 13 on the questionnaire *The symbols helped me trust the situational picture*, $r = -0.466(21), p < .05$.

There is no correlation between the amount of eye contact in the AR-condition and the paper map-condition, $r = 0.184(21), p > .05$.

Table 2: Data from the eye contact analyses. The Paper Map session to the left and the second AR session to the right (the second session uses data from the first AR session, see text). Organisation (Org) corresponds to R for Rescue service, P for Police and M for Military. n(EC) is the amount of eye contact, i.e. the number of times that a person attempts eye contact, t(EC) is the total time in seconds for all EC's and t(nonEC) is the time with no eye contact. EC% is the proportion of eye contact for the whole session. Horizontal lines denote separate sessions. (The second session had one outlier, see text.)

| Org | Paper Map | | | | AR | | | |
|------|-----------|-------|----------|-----|-------|-------|----------|------|
| | n(EC) | t(EC) | t(nonEC) | EC% | n(EC) | t(EC) | t(nonEC) | EC% |
| R | 17 | 0.32 | 19.5 | 2.6 | 4 | 0.06 | 19.59 | 0.5 |
| M | 7 | 0.06 | 20.15 | 0.5 | 3 | 0.03 | 20.02 | 0.2 |
| P | 41 | 0.32 | 19.5 | 2.6 | 1 | 0.01 | 20.05 | 0.01 |
| M | 25 | 0.36 | 19.30 | 3.0 | 1 | 0.01 | 18.36 | 0.1 |
| P | 30 | 0.47 | 19.19 | 3.9 | 2 | 0.01 | 18.36 | 0.01 |
| R | 18 | 0.23 | 19.41 | 1.9 | 1 | 0.01 | 20.06 | 0.1 |
| M | 9 | 0.13 | 19.52 | 0.3 | 2 | 0.02 | 20.05 | 0.2 |
| P | 11 | 0.11 | 19.53 | 0.1 | 0 | 0.00 | 20.07 | 0 |
| R | 22 | 0.39 | 19.33 | 3.2 | 3 | 0.03 | 20.04 | 0.2 |
| M | 13 | 0.14 | 19.58 | 0.3 | 6 | 0.09 | 19.58 | 0.7 |
| P | 14 | 0.12 | 20.00 | 0.1 | 4 | 0.03 | 20.04 | 0.2 |
| R | 11 | 0.12 | 19.54 | 1.0 | 0 | 0.00 | 20.06 | 0 |
| M | 37 | 0.37 | 19.29 | 3.1 | 2 | 0.01 | 20.04 | 0.1 |
| P | 21 | 0.14 | 19.52 | 1.2 | 0 | 0.00 | 20.06 | 0 |
| R | 22 | 0.22 | 19.44 | 1.8 | 10 | 0.12 | 19.50 | 1.0 |
| M | 23 | 0.16 | 19.49 | 1.3 | 0 | 0.00 | 20.10 | 0 |
| P | 33 | 0.37 | 19.28 | 3.1 | 3 | 0.01 | 20.01 | 0.1 |
| R | 7 | 0.12 | 19.53 | 1.0 | 0 | 0.00 | 20.10 | 0 |
| M | 16 | 0.13 | 19.53 | 1.0 | 0 | 0.00 | 20.10 | 0 |
| P | 22 | 0.26 | 19.40 | 2.2 | 6 | 0.08 | 20.02 | 0.7 |
| R | 41 | 0.49 | 19.08 | 4.1 | 3 | 0.03 | 20.03 | 0.2 |
| M | 47 | 0.38 | 19.18 | 3.2 | 2 | 0.01 | 20.05 | 0.1 |
| P | 59 | 1.07 | 18.50 | 5.6 | 4 | 0.02 | 20.04 | 0.2 |
| Mean | 24.1 | 0.26 | 19.41 | 2.0 | 2.5 | 0.03 | 19.56 | 0.2 |

T-tests ($df = 22$) were carried out on these four questions and the corresponding question for the AR-condition to see whether if there was any difference.

The ratings were significantly higher for the AR-condition for questions 8 ($t = 3.246, p = .004$), 12 ($t = 2.185, p = .040$), and 13 ($t = 2.355, p = .028$). There was no significant difference for question 11 ($t = 1.727, p = .098$).

5 Discussion

The difference in eye contact between the paper map and the AR-condition are clearly significant. Wearing an AR system with a head mounted display does indeed decrease the amount of eye contact that individuals working together have, even if they have an external artefact to focus on in both cases (the map). The amount of eye contact in the AR-setting was roughly 10% of what it was in the paper map condition.

None of the questions in the questionnaire correlated with the amount of eye contact in the AR-condition. This is likely due to the very small amount of eye contact that could be observed. However, for all but one of the questions in the paper map-condition that correlated with eye contact, the corresponding AR-question was rated significantly higher.

It is worth pointing out that previous analyses of the questionnaire data used in this study revealed that on most issues the AR system is as good as or better than the paper map [27, 25]. For instance the participants did not consider it more difficult to achieve a common situational picture with the AR system than when using an ordinary paper map, nor did they regard the AR system to interfere with their communication to any large extent [27].

Interpreting the relationship between the questions and eye contact, we can see that the amount of eye contact in-

creases the more difficult the participants found it to use the symbols (on the map) to create a common operational picture, the less confident they felt interpreting the map, and the less trust they had in the symbols or the map.

While the statistical tests performed do not say anything about the nature of the causal relationship, if any, between the amount of eye contact and the questionnaire answers, it nonetheless seem likely that the lack of confidence and trust in the paper map and its symbols urged the participants to confirm information by increasing the communication with their coworkers. As a result of this, their amount of eye contact increased.

It then follows that in the AR-condition where the participants had less opportunity for casual eye contact, they should rate these questions even lower, as they cannot as easily confirm what the symbols and map mean for the common operational picture. However, this was not the case. In fact, the rating on 3 of the 4 questions increased (the fourth was also higher, but not significantly so).

Returning to the model by Gutwin and Greenberg [10] using three dimensions of team cognition awareness, it may be so that an increase in the capabilities of one dimension, in this case the artefacts and feedthrough, may compensate for losses in others, in this case the bodies and consequential communication. The team would thus be aware of the shared workspace through different means than eye contact.

While it appears that the commonly held assumption that using a HMD will decrease eye contact holds, the effects of this decrease are not as simple. The decrease does not appear to directly influence teamwork in a negative way (according to subject-ratings), nor does it appear to increase the uncertainty of the interpretation of the artefacts involved. Furthermore, the amount of eye-contact is very small, which

may well be explained by our experiment. It may be that professionals collaborating on a well known task do not have to communicate that much or have other means than eye contact to communicate. This leads to the question of the importance of expertise and context. Would a study using a novice, e.g. in a teacher-student dyad, provide the same result? What if the context was changed from a very task-oriented command and control collaboration to a less task-focused collaboration where social interaction is more important? Different cultures have different norms for the appropriateness of eye contact in different context, and this might also be an influencing factor. As may the size of the group and the design of the AR-equipment. For reasons outlined earlier in this paper, this study relied on experts collaborating in a task-driven context, but we acknowledge that there are many factors that may influence the effect of AR on eye contact and the effect of eye contact on the actual collaboration itself. The experimental setup used in this study was not designed to allow for independent measures of task performance, for a more elaborate discussion on this see [21, pp. 44–49]. Instead, user questionnaire items were used to benchmark performance. Changing this to include separate measures for task performance might yield further insights concerning the importance of eye contact in different collaboration settings. More user studies should be conducted to further explore these and other related questions. The results of this paper should be taken into consideration when designing AR systems for use in contexts similar to that in this user study.

Acknowledgment

This research is supported by FMV, Technology for Sweden's Security. We are indebted to XMReality for developing the system used in the studies and to Fredrik Köhler and Hanna Mårtensson who assisted in the studies.

References

- [1] R. Azuma. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.
- [2] M. Billinghurst and H. Kato. Out and about—real world teleconferencing. *BT Technology Journal*, Jan 2000.
- [3] M. Billinghurst and H. Kato. Collaborative augmented reality. *Communications of the ACM*, 45(7):64–70, July 2002.
- [4] M. Billinghurst, S. Weghorst, and T. A. Furness. Wearable computers for three dimensional CSCW. In *ISWC*, pages 39–46, 1997.
- [5] T. E. Drabek and J. E. Haas. Realism in laboratory simulation: Myth or method? *Social Forces*, 45(3):337–346, 1967.
- [6] A. Dünser, R. Grasset, and M. Billinghurst. A survey of evaluation techniques used in augmented reality studies. Technical Report Technical Report TR2008-02, Human Interface Technology Laboratory New Zealand, 2008.
- [7] A. Dünser, K. Steinbügl, H. Kaufmann, and J. Glück. Virtual and augmented reality as spatial ability training tools. In M. Billinghurst, editor, *Proceedings of the 7th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Design Centered HCI, 2006, Christchurch, New Zealand, July 6-7, 2006*, volume 158 of *ACM International Conference Proceeding Series*, pages 125–132. ACM, 2006.
- [8] A. Fuhrmann, H. Löffelmann, and D. Schmalstieg. Collaborative augmented reality: Exploring dynamical systems. In R. Yagel and H. Hagen, editors, *IEEE Visualization '97*, pages 459–462. IEEE, Nov. 1997.
- [9] R. Granlund. Web-based micro-world simulation for emergency management training. *Future Generation Computer Systems*, 2001.
- [10] C. Gutwin and S. Greenberg. The importance of awareness for team cognition in distributed collaboration. Technical Report Report 2001-696-19, Dept Computer Science, University of Calgary, Alberta, Canada, 2001.
- [11] M. Haller, M. Billinghurst, and B. H. Thomas. *Emerging Technologies of Augmented Reality*. IGI Publishing, Hershey, PA, USA, 2006.
- [12] A. Henrysson, M. Billinghurst, and M. Ollila. Face to face collaborative AR on mobile phones. In *ISMAR '05: Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 80–89, Washington, DC, USA, 2005. IEEE Computer Society.
- [13] A. Henrysson, M. Ollila, and M. Billinghurst. Mobile phone based AR scene assembly. In *MUM '05: Proceedings of the 4th international conference on Mobile and ubiquitous multimedia*, pages 95–102, New York, NY, USA, 2005. ACM.
- [14] E. Hutchins. *Cognition in the Wild*. MIT Press, 1995.
- [15] H. Ishii and M. Kobayashi. Clearboard: A seamless medium for shared drawing and conversation with eye contact. In *ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 525–532, Monterey, California, May 3-7 1992. ACM Press.
- [16] B. Johansson, M. Persson, R. Granlund, and P. Matsson. C3fire in command and control research. *Cognition, Technology & Work*, 5(3):191–196, 2001.
- [17] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99), San Francisco, USA, 1999*.
- [18] C. L. Kleinke. Gaze and eye contact: A research review. *Psychological Bulletin*, Jan 1986.
- [19] D. Knight, D. Langmeyer, and D. Lundgren. Eye-contact, distance, and affiliation: the role of observer bias. *Sociometry*, Jan 1973.
- [20] B. A. Nardi, editor. *Context and Consciousness: Activity Theory And Human-Computer Interaction*. The MIT Press, Cambridge, Massachusetts, 1996.
- [21] S. Nilsson. *Augmentation in the Wild: User Centered Development and Evaluation of Augmented Reality Applications*. PhD thesis, Linköping Studies in Science and Technology, Dissertation No. 1306, 2010.
- [22] S. Nilsson and B. Johansson. A cognitive systems engineering perspective on the design of mixed reality systems. In *ECCE '06: Proceedings of the 13th European conference on Cognitive ergonomics*, pages 154–161, New York, NY, USA, 2006. ACM.
- [23] S. Nilsson and B. Johansson. Fun and usable: augmented reality instructions in a hospital setting. In B. Thomas, editor, *Proceedings of the 2007 Australasian Computer-Human Interaction Conference, OZCHI 2007, Adelaide, Australia, November 28-30, 2007*, volume 251 of *ACM International Conference Proceeding Series*, pages 123–130. ACM, 2007.
- [24] S. Nilsson and B. Johansson. Acceptance of augmented reality instructions in a real work setting. In M. Czerwinski, A. M. Lund, and D. S. Tan, editors, *Extended Abstracts Proceedings of the 2008 Conference on Human Factors in Computing Systems, CHI 2008, Florence, Italy, April 5-10, 2008*, pages 2025–2032. ACM, 2008.
- [25] S. Nilsson, B. Johansson, and A. Jönsson. A co-located collaborative augmented reality application. In *Proceedings of VRCAI 2009, Yokohama, Japan, 2009*.
- [26] S. Nilsson, B. Johansson, and A. Jönsson. A holistic approach to design and evaluation of mixed reality systems. In E. Dubois, P. Gray, and L. Nigay, editors, *The Engineering of Mixed Reality Systems*. Springer, 2009.
- [27] S. Nilsson, B. Johansson, and A. Jönsson. Using AR to support cross-organisational collaboration in dynamic tasks. In *Proceedings of IEEE ISMAR-09, Orlando, FL, 2009*.
- [28] S. Prince, A. D. Cheok, F. Farbiz, T. Williamson, N. Johnson, M. Billinghurst, and H. Kato. 3-D live: real time inter-

- action for mixed reality. In *CSCW*, pages 364–371, 2002.
- [29] P. Renevier and L. Nigay. Mobile collaborative augmented reality: The augmented stroll. *Lecture Notes in Computer Science*, 2254, 2001.
- [30] G. Salomon. No distribution without individuals' cognition: A dynamic interactional view. In G. Salomon, editor, *Distributed Cognitions: Psychological and Educational Considerations*, pages 111–138. Cambridge University Press, New York, 1993.
- [31] R. Samuracy and J. Rogalski. Cooperative work and decision making in emergency management. *Le Travail Humain*, 56:53–77, 1993.
- [32] D. Schmalstieg, A. Fuhrmann, Z. Szalavri, and M. Gervautz. Studierstube: An environment for collaboration in augmented reality. In *Collaborative Virtual Environments (CVE'96)*, Nottingham, UK, 1996.
- [33] Schmidt, Kjeld. Editorial. *Computer Supported Cooperative Work*, 9(2):155, 2000.
- [34] A. Tang, C. Owen, F. Biocca, and W. Mou. Experimental evaluation of augmented reality in object assembly task. In *ISMAR '02: Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, page 265, Washington, DC, USA, 2002. IEEE Computer Society.
- [35] L. S. Vygotsky. *Mind In Society*. Harvard University Press, Cambridge, 1978.