

Linköping University Post Print

Using AR to support cross-organisational collaboration in dynamic tasks

Susanna Nilsson, Björn Johansson and Arne Jönsson

N.B.: When citing this work, cite the original article.

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Susanna Nilsson, Björn Johansson and Arne Jönsson, Using AR to support cross-organisational collaboration in dynamic tasks, 2009, IEEE ISMAR-09.

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-52588>

Using AR to support cross-organisational collaboration in dynamic tasks

Susanna Nilsson*

Department of Computer and Information Science
Linköping University, Sweden

Björn Johansson†

Swedish Defence Research Institute
Linköping, Sweden

Arne Jönsson‡

Santa Anna IT Research Institute
Linköping, Sweden

ABSTRACT

This paper presents a study where Augmented Reality (AR) technology has been used as a tool for supporting collaboration between the rescue services, the police and military personnel in a crisis management scenario. There are few studies on how AR systems should be designed to improve cooperation between actors from different organizations while at the same time support individual needs. In the present study an AR system was utilized for supporting joint planning tasks by providing organisation-specific views of a shared working. The study involved a simulated emergency event conducted in close to real settings with representatives from the organisations for which the system is developed. As a baseline, a series of trials without the AR system was carried out. Results show that the users were positive towards the AR system, and would like to use it in real work. They also experience some performance benefits of using the AR system compared to their traditional tools. Finally, the problem of designing for collaborative work as well as the benefits of using an iterative design processes is discussed.

1 INTRODUCTION

In complex collaborative situations, such as crisis management, actors from different domains and organisations must work together [4]. However, collaborative work across organisational borders is not simple and confusion emerging from differences in terminology, symbols or organisational structure is not rare. The information presented to the actors has to be simple enough to support cooperation between actors from different organisations but at the same time be rich enough for an actor from a specific organisation to facilitate her decision making.

The hypothesis in this paper is that Augmented Reality (AR) is especially suitable to support collaboration between actors from different organisations. AR allows for independence and individuality [2] meaning that each actor can independently have data tailored to her needs in various situations. AR also supports cooperation [2] as the actors can see each other and cooperate in a natural way.

This paper presents an evaluation of a multi-user AR application, where AR is used to aid cross-cultural collaboration. The system is intended to support collaborative work between representatives from police, rescue service and military personnel, working jointly with the goal of coordinating work in a crisis situation.

The purpose of this paper is threefold; It discusses the use of AR for collaborative command and control in crisis management operations (Section 2); It presents a design methodology for development of an AR system for this purpose (Sections 3 and 4); and it presents the results of a comprehensive user study conducted after two design iterations (Section 5). The paper ends with a discussion of the possible implications for design of AR systems for collaboration.

*e-mail:susni@ida.liu.se

†e-mail:bjorn.j.e.johansson@foi.se

‡e-mail:arnjo@ida.liu.se

2 RELATED WORK

Collaborative work has been studied extensively in many different research domains, from sociological and psychological perspectives as well as organisational perspectives. Technological tools which aid collaboration have also been developed within the broad range of research on computer supported collaborative work, such as decision support systems combined with teleconferencing systems. Virtual environments have been used as tools for training and simulating collaborative work (for instance the CAVE system and the Virtual Workbench [8]), but few, if any, systems have actually been aimed for use in crisis management situations.

2.1 Collaborative command and control

When personnel from different organisations work together under stress, as in many crisis situations, there is always a risk that misunderstandings emerge due to differences in terminology or symbol use. In linguistics, it is common knowledge that time has to be spent on establishing a 'common ground', or a basis for communication, founded on personal expectations and assumptions between the people communicating with each other [3, 20]. The maintaining of common ground is an ongoing process, which demands both attention and coordination between the participants. Exercising command and control is an attempt to establish common intent to achieve coordinated action [21]. Successful communication is obviously necessary to achieve this. The AR system can facilitate the establishment of a common ground, e.g. by having shared objects that can be viewed and manipulated from different perspectives.

In addition to this, there are situation-specific problems that emerge in collaborative command and control tasks. Such tasks often circle around a shared representation of the current activities, as in the case of a situational map. Most organisations involved in such tasks, like the military or rescue services, have developed a library of symbols that can be utilised for representing units and events. A problem arises when representatives from different organisations are to work together, since they are used to working with their own, organisation-specific, symbols and conventions. This means that time has to be spent explaining and negotiating meaning when jointly creating and manipulating a shared representation, a tedious task to undertake when there is little time, as for example in the case of forest fire-fighting in, or close to, urban areas. The idea behind this study is to improve initial common ground by providing commanders from different organisations with personalised views of a situational map, as can be seen in Figure 1.

Another aspect to consider is the awareness of team cognition [11]. Gutwin and Greenberg [11] 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" [11, 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) artifacts and feedthrough.

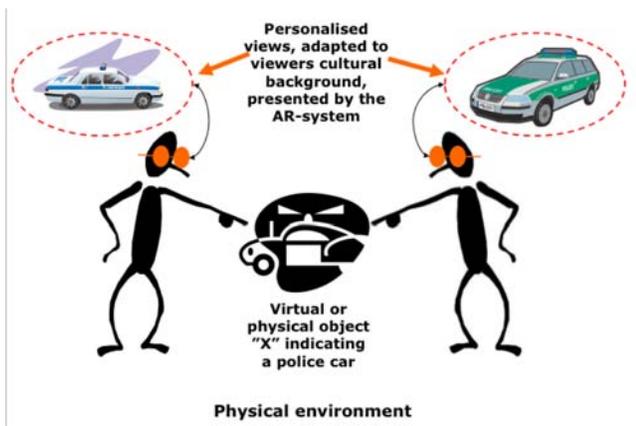


Figure 1: Using AR to improve initial understanding in a collaborative task

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 artifacts in use. Gutwin and Greenberg [11] 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 artifacts” [11, 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 artifact, and also manipulate the same objects. Gutwin and Greenberg [11] do not address AR systems, but an AR system like the one suggested in this study may provide an excellent example of feedthrough.

2.2 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.* [12]). 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 [28], training and education [6], mobile phones [14], mobile applications [25] 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 [27]. Fuhrman *et.al.* [8] presented an AR system for collaborative scientific visualisation with 3D interaction and customised views for several users. Billingham & Kato [2] presented a vision of shared space using AR technology, Henrysson *et.al.* [13] developed a collaborative mobile phone application and since then several papers have illustrated 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 as described above. Even less research exists regarding the use of AR as a support for improving the pre-conditions for communication between personnel from different organisations.

Improving shared understanding between commanders has the potential to speed up coordination work, something that may prove

to be an important enabler of success in many real-world situations.

2.3 Design and evaluation issues of AR systems

Even though AR systems are designed differently with different applications and tasks in focus, the usability methods used to evaluate them are similar and mainly based on usability methods used for more traditional graphical user interfaces, sometimes in combination with usability for VR applications [29, 22, 6]. Designing systems based on heuristics developed for computer based applications may be common practise in the AR field, but there are few examples of studies on how users actually perceive the system in actual use situations [23, 24]. In contrast to traditional methods, which analyse the user and system as separate parts, the Cognitive Systems Engineering approach [16, 30] emphasises a systemic view in which the system, including the user, is studied as a whole rather than as one technical device that the user interacts with. The natural environment of the user is also an important part of the analysis. In this way the analysis focuses on function rather than structure, which is more useful for analyses of novel systems such as the AR system presented in this paper [23, 24].

Usability methods such as cognitive task design [15] where the design approach is based on observations of how a user completes a task in which the system or artifact is involved, also have to deal with the so called ‘envisioned world problem’ [17, 31] The ‘envisioned world problem’ states that even if a good understanding of a task exists, the new design, or tool, will change the task, rendering the first analysis invalid. Acknowledging the ‘envisioned world problem’, we have adapted an iterative design approach where realistic exercises are combined with focus groups in an effort to catch both user behaviour and opinions.

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 [26] 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.* [18] has argued for the concept of evaluating novel technologies by combining a representative task, such as a micro-world (like the C³Fire simulation used in this study) with ‘professional’ users. 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 THE AR SYSTEM USED IN THE STUDY

The AR system used for our study comprise three identical high fidelity AR prototypes¹, one for each experiment participant. The AR system provides the capability to work in a shared space, in this case a map, which is the basis for the task. The starting point for any cross-organisational operation involving the police, the fire and rescue services and the military helicopter platoons is a relatively serious crisis situation. For instance a wide spread forest fire which is not under control and forces the fire department to request back-up from the police and the military in order to limit the damages of the fire. The police assist with the evacuations, traffic control, finding missing people, etc., while the military assist the fire department both on the ground and in the air with water bombing. Usually a forest fire that requires this involvement has been going on for a couple of days, and the weather conditions are not favourable for the fire

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

fighters. This means a scenario where the events have forced the on-scene commander from the fire department to request backup from military at which stage the field commanders from the three organisations will meet to evaluate and assess the current situation, the events that has lead up to the situation and finally to agree on a course of future action. It is in this stage that there is a need for a common situational picture and overview of all available resources and tools for planning the operation. This is the stage for the study presented below.

3.1 The first iteration of the AR system

The AR system was iteratively designed in three steps of which the third evaluation is the user study presented in section 4. In the pre-design phase field experts took part of a brainstorming session to establish the parameters of the AR system. This brainstorming session was used to define the components of the software interface, such as what type of symbols to use, and what type of information is important and relevant in the task for creating common ground between the three participating organisations.

After the brainstorming session a first design was implemented and evaluated in a study conducted with the purpose of evaluating the system design as a tool for collaboration between organisations. To promote realistic results, the participants were representatives from the three organisations in focus; the fire department, the police and the military helicopter platoon. The setting was at a military helicopter base.



Figure 2: The study was conducted in a simulated natural setting at a helicopter base.

The AR system was designed to support cooperation as advocated by Billingham & Kato [2] and Gutwin & Greenberg [11] and thus emphasised the need for actors to see each other. Therefore, it was equipped with hand-held displays that are easier to remove from the eyes than head mounted displays, see Figure 2. We used a digital map where participants had personal, individual views, allowing them to see an organisation specific map and the symbols they normally use. In this way each actor has her own information mapping to the AR markers on the map to facilitate independence and individuality. A feature allowed each participant to send their view of the situation (i.e. their individual map) to the other participants when necessary. Hand pointing on the map was not possible as the hand was occluded by the digital image of the map in the display.

This design was then evaluated using a scenario in which participants, one from each of the three organisations, had to interact and work together to complete tasks in a dynamic scenario. The exercise was observed and the participants also answered questionnaires pertaining to the AR system design, and finally a focused

group discussion was held.

The evaluation revealed a number of issues regarding the design of the system as well as the scenario being used. In general, the participants were positive to the AR system. What they appreciated most was the easy overview of what was going on. Being able to see all resources placed on the map facilitates the joint task.

Several suggestions were given for redesign including a map with more details, more events in the scenario played and changing the physical interaction devices. Especially the design of the AR displays as a handheld device did not receive a positive response and the observations clearly illustrated this problem.

The participants also commented on more positive aspects of the system, such as the possibility of spatially distributed collaboration. Other findings in the first evaluation questionnaires were that despite the relatively clumsy design of the prototype, all participants thought it was easy to use and that it was quick to learn. Despite flaws in the system, all participants could also see themselves using the AR system in their professional life as well as in other situations.

3.2 The second iteration of the AR system

As a result of the design evaluation the system was redesigned. The handheld display was replaced with a head mounted display allowing freedom of movement. The interaction device was also considerably redesigned and in the new AR system the user can easily manipulate objects using only one hand as opposed to using both in the previous prototype, see Figure 3.



Figure 3: The redesigned interaction device, which allows the user to choose a virtual object and place it on the digital map.

Another improvement made was a simplified interaction in which the user can use their hand to point at things in the digital map. In the previous design this pointing manoeuvre could not be seen as the digital map was superimposed over the pointing hand giving the impression that the user was pointing 'under' the map rather than on the map. The first prototype therefore had a pointing function in the interaction device. The new, improved technical design has eliminated the need for this pointing device as the system now allows the users hand to be superimposed over the digital map image using blue-screen technique, see Figure 4. This allows the users to use deictic gestures like pointing since their hands are visible above the digital representation. The system thus presents several properties of a normal paper map with the added functionality of adding, moving and removing digital objects that carry in-

formation and can be manipulated by any user working with the system.



Figure 4: The users display showing the digital map with symbols and pointing used in the collaborative AR application

The redesigned AR system was evaluated in a focus group discussion where the participants first were asked to reflect on their experience in the previous study. Then the redesigned system was presented and the participants were observed using it to complete simple tasks from the scenario in the pre-study. After this the focus group discussion continued with reflections on the new design.

The results from the discussions were positive. The problems that they reported on previously had been addressed. The head mounted display was a big improvement and allowed them to move around and interact more freely. The new joystick interaction device was also appreciated and the participants found it very easy to use. The added possibility to see hand gestures such as pointing, on the digital map has simplified the interaction considerably and also resulted in a more natural interaction and better communication between the participants. In the redesigned application the participants had exactly the same view allowing them to alter their personal image but still seeing the same map, and not as previously the organisation specific map. As noted by one of the participants during the group discussion:

“A common picture, everything is better than me telling someone what it looks like...you need to see the picture and not to hear my words.” (participant from the second evaluation)

3.3 The AR system used in the third iteration

As stated above, the AR system used in the study comprise three identical AR systems, developed by XMReality. 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 system ran on a Dell XPS M1330, with a 2.10GHz processor, 3 GB RAM and with a 128 MB NVIDIA GeForce 8400M GS graphics card and the marker tracking was based on ARToolkit [19]. Each AR system was independent in relation to the others, i.e. the systems were not dependent on each other in order to function properly.

In order for the users to share the same view the AR systems must be interlinked and responsive to what each system user does. In the three system setup the individual AR systems communicate through an internal Ethernet network. Each system listens for

changes in the internal representation in the other AR systems and updates its own internal representation to reflect the changes made to the representation in the other two AR systems.

3.4 Functionality of the AR system

The users have access to a personal, organisation-specific symbol library which they can use to create a situational picture. Examples of symbols are police vehicles, fire trucks, helicopters, and personnel. Other types of symbols are the function symbols, for instance the ① symbol which when used allows the user to see additional information about the already placed symbols, such as information about how many hours personnel has been on duty, or how much water is left in the tank of a tank truck. Other functions include zooming in or out and saving or retrieving an image (i.e. a screenshot of the current layout). The symbols are simplified to some degree in order to be understandable by users from other organisations. There is organisation-specific information connected to the symbols that can be displayed on demand. It is also possible to personalise the system by filtering out symbols belonging to one or more organisation, showing for instance only symbols from the own organisation on the map.

If necessary, the users can manipulate each others symbols, e.g. a fire-fighter can place, delete and move a police vehicle. There are also a set of symbols that are common to all users of the AR system, such as fires and smoke (this is particularly important in this case as the participants in the study are confronted with a forest-fire fighting task). 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 in the menu by using the buttons on the joystick. The same procedure is used to remove a symbol, see additional information about a symbol, or zoom in the map.

4 THE THIRD ITERATION - THE FINAL USER STUDY

As noted, the cognitive systems engineering approach to studying human computer interaction advocates a natural setting and a realistic task. The aim of the study was not to measure performance in terms of metrics such as task completion time [10], as these types of measures require a repeatable setting and identical trials for all participants in order to give meaningful comparable results. In a natural setting, unforeseen consequences are inevitable and also desirable, which means that no trials will be identical. The performance measures of interest in this study are instead the users experience of the AR system and how well the system achieves the intended goals. Unfortunately current AR systems are not developed enough for use in critical real life situations, especially not if used in situations where enormous values are on stake, such as large forest fires. Consequently, we use simulations in this study, cf. [31].

4.1 Participants

The AR application was evaluated in a study where ten groups, with three participants in each group, used the system in a simulated scenario of a forest fire. The theoretical starting point was that in order to find real world applicable results we need real world end users. To meet this demand participants from three different organisations involved in crisis management were recruited. In total 30 participants took part in the study during ten sessions distributed over ten days, with three people in each session. The participants were all at the level in their organisation where they in real life are assigned to team-coordinating situations. This means that they all either have experience from working in teams with partners from

at least one of the other organisations, or have a position in their organisation which require that they have a minimal education and training in these types of command and control assignments. The groups formed here had never worked together before and they did not know each other prior to this study.

Of the ten trials, two were spoiled due to unforeseeable events (in one case one participant was called to active duty due to an emergency and in the other case external technical problems forced the trial to end prematurely). This resulted in a total of eight complete trials with 24 participants, of which 23 were male, one female and the ages ranged from 25 to 57 (median: 36, average: 39,1). There is a clear gender imbalance which is mainly due to the composition of the user groups, the vast majority of the firemen in this area are male, all helicopter pilots are male, and a majority of the police are male, thus the selection of participants is representative for the user group populations.

4.2 Procedure

The setting was at a military helicopter base in which the environment was designed to simulate a rough in-the-field command and control environment (meaning that the users only had a table and basic equipment such as pens and paper available, see Figure 5).

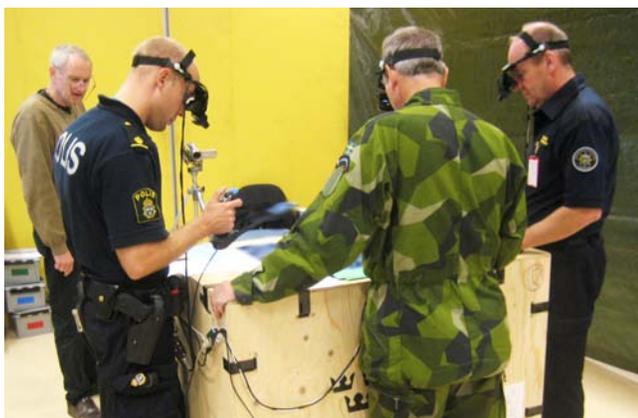


Figure 5: The simulated natural setting (a helicopter base).

In order to create a dynamic scenario and realistic responses and reactions to the participants' decisions in the three sessions, we used a gaming simulator, C³Fire [9]. C³Fire generates a task environment where a simulated forest fire evolves over time. The simulation includes houses, different kinds of vegetation, computer simulated agents, vehicles etc. that can be controlled by an experiment assistant. The simulator was run in the background by the research team, see Figures 6 and 7, where one member, the experiment assistant, inserted information into the gaming simulator, for instance, that several police cars have been reallocated to attend to a traffic incident. The experiment leader acted as a feedback channel to the participants in order for them to carry out their work. In other words, the experiment leader took the role of a communication system between the commanders and the field personnel. For instance, when the reallocated police cars had reached their new destination the experiment leader returned with information to the participants. Other examples of information from the gaming simulator are weather reports, status of personnel and vehicles, the spread of the fire etc.

The application was designed around a scenario in which the participants, one from each organisation, had to interact and work together to complete tasks in a dynamic scenario. Three different scenarios were used, each describing a forest fire that has been going on for a couple of days. The description was rather detailed and

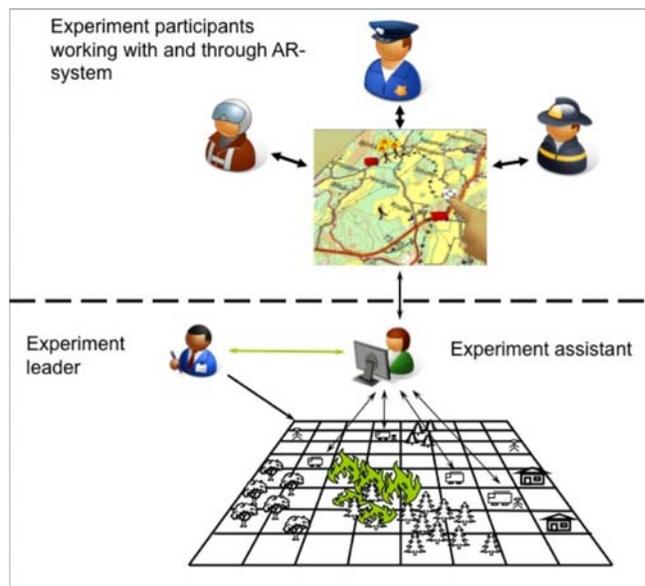


Figure 6: A schematic view of the C³ Fire gaming simulator used to create a dynamic and interactive scenario in the user study.

included information on when the fire has started, where people had been seen, weather conditions etc. Each organisation had a number of units that they had to place on the map as they would have done in a real situation². The participants all have the same digital map in their view. They can independently place symbols using the handheld interaction device and they can also discuss with the others how to place their own symbols and also common symbols, such as the fire symbol and break points.



Figure 7: 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.

After a 30 minute training session, each group of three participants performed three simulations, each lasting 20 minutes. The first simulation session was conducted using the AR system, the second was conducted using a traditional paper map and the third session was again conducted using the AR system. The paper map

²All participants are used to various similar training exercises from their own organisations, so this never posed a problem.

session was included to be able to compare the use of an AR system to a 'system' that they normally use, i.e. a paper map, marker pens and transparencies. We used three different simulation scenarios permuted between sessions. All three scenarios are identical in number and type of events, but the events are distributed differently to avoid learning effects on the specific tasks.

After each 20 minute session the participants filled in a questionnaire on cooperation using the AR system or the paper map and after the final session they also filled in a questionnaire on the AR system. The questionnaires used six-point Likert items and also had open ended questions, such as *Did you experience anything as troublesome, and if so, what?, How did you experience the system? Can you compare it to anything else?*, see Section 5 for more examples of open ended questions. The questionnaires filled out between sessions included 15 closed response items and 6 open ended questions. The final AR system questionnaire included 18 items and 10 open ended questions. Finally the participants could more freely express their views in a semi-controlled group discussion on different topics related to the AR system design, the scenario, aspects of collaboration and communication.

To summarise the experiment:

Activity	Duration
Introduction to AR and the experiment	30 minutes
AR practise using a predefined set of tasks	≈ 15 minutes
Paper map exercise using a predefined set of tasks	≈ 15 minutes
AR session 1	20 minutes
Co-operation questionnaire	≈ 10 minutes
Paper map session	20 minutes
Co-operation questionnaire	≈ 10 minutes
AR session 2	20 minutes
Co-operation questionnaire	≈ 10 minutes
AR questionnaire	≈ 15 minutes
Focus group discussion	≈ 20 minutes

5 RESULTS AND DISCUSSION

In this section we present results from using the AR system the second time (AR session 2, above) compared to using a traditional paper map for collaboration. We also present results specifically addressing the use of the AR system.

5.1 Comparing the AR system to the paper map

As stated above, we had a questionnaire after each session where the participants responded to items on how the AR system supported collaboration, see Figure 8³. Using one way ANOVA with Bonferroni post hoc tests we found interesting differences between the paper session and the AR session on Items 7, 8, and 9, as well as a tendency on Item 13.

In general the results on the questionnaire were positive for the AR system. The average scores for all items, but one, were above 4 out of 6 which can be considered good for a new system.

On Item 1, *It took a long time to start to cooperate*, the participants do not experience that the AR system influences the time it takes to start to cooperate, nor does the paper map, both have high scores, 5.33 for both the AR system and the paper map. That the paper map does not influence the time it takes to start to cooperate is not surprising as this is what they are used to.

When asked if it was easy to collaborate, Item 2, the results were positive in both sessions - the mean score was 4.70 (AR) and 5.09 (paper map) on a 6 grade scale, no significant difference.

When asked about the AR system as a tool for collaboration, Item 3, the average scores were above 4 in both sessions. There was no significant difference between the paper map and the AR system sessions.

³The queries have been translated from Swedish to English by the authors.

#	Question/Item	AR System		Paper map	
		μ	σ	μ	σ
1	It took a long time to start to cooperate.	5.333	0.963	5.333	1.090
2	It was easy to cooperate.	4.696	1.550	5.087	1.474
3	I think that the AR system is a good tool for cooperation.	4.417	1.136	4.250	1.391
4	The collaboration was fun.	5.292	0.690	5.208	0.833
5	I felt that the group was in control of the situation.	4.792	1.141	5.000	0.885
6	It was easy to mediate information between the organisations.	4.833	0.917	5.000	0.834
7	The map made it easy to achieve a common situational picture.	5.041	0.999	4.167	1.404
8	The symbols made it easy to achieve a common situational picture.	4.833	1.050	3.417	1.472
9	The map became messy/cluttered.	4.000	1.474	2.375	1.469
10	I would have liked to have more information than what was available.	3.292	1.459	3.167	1.659
11	I felt certain that I could interpret what was on the map.	4.542	1.132	3.750	1.452
12	The map helped me trust the situational picture.	4.667	1.090	3.667	1.373
13	The symbols helped me trust the situational picture.	4.583	1.060	3.500	1.474
14	I thought I had a good situational picture.	4.542	1.103	4.250	1.032
15	I thought the others had a good situational picture.	4.500	1.142	4.458	0.977

Figure 8: Questionnaires, average score and standard deviation (using the 6 point Likert scale). 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

Concerning whether or not the participants enjoyed the collaboration, Item 4 *The co-operation was fun*, the scores were 5.29 for the AR system and 5.21 for the paper map. There was no significant difference between the sessions.

On Item 5, *I felt that the group controlled the situation* we note a lack of significant difference between the paper based session and the AR session. The use of the AR system does not seem to result in an experienced loss of control in comparison to the paper map, despite that it is more difficult to get eye contact with the head-mounted display. In the paper map session the average score was 5.0 in comparison to an average of 4.79 in the AR session.

An important aspect of collaboration is information sharing which was targeted in Item 6, *It was easy to mediate information between the organisations*. The overall average score on the item was 4.83 using AR and 5.0 in the paper session. However, there was no significant difference between the AR session and the paper map session which means that sharing information was experienced as easy to do while working on the paper map as with the AR system.

A group of questionnaire items specifically addressed the map and the symbols on the map; Item 7, Item 8, and Item 9. Here the scores for the AR system are higher than for the paper map, Figure 9, suggesting that the use of the AR system made it easier to achieve a common situational picture. Regarding the map, Item 7 *The map made it easy to achieve a common situational picture*, we only see a tendency to difference between the AR systems mean score of 5.04, and the paper maps mean score of 4.18 ($F(2,42)=6.1, p\approx 0.052$). Regarding the symbols, Item 8 *The symbols made it easy to achieve a common situational picture*, there is a significant differ-

ence between the mean scores of 4.83 and 3.42. The symbols in the AR system made it easier to achieve a common situational picture compared to the paper map ($F(2,42)=15.3, p<0.05$). The map is also regarded as less messy when using the AR system, Item 9 *The map became cluttered*, with a significant difference ($F(2,42)=12.7, p<0.05$) between the mean score of 4.0 for the AR system and 2.38 for the paper map.

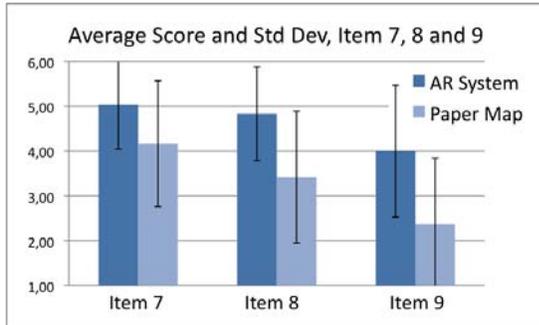


Figure 9: Results from Items 7, *The map/AR system made it easy to achieve a common situational picture*, 8, *The symbols made it easy to achieve a common situational picture* and 9, *The map was cluttered/messy*. The AR system scored significantly higher than the paper map. See text for further details.

We also note that the users wanted even more information than we had on the map, Item 10 *I would have liked to have had more information than what was available*, scoring rather low (mean score of 3.29 for the AR system and 3.17 for the paper map) on this item. The participants had, however, no problems to interpret the symbols and the map, giving Item 11, *I felt that I was certain that I could interpret what was on the map*, a mean score of 4.54 when using the AR system and 3.75 with the paper map.

When asked if the map and symbols helped the participants trust the situational picture, Item 12 and Item 13 there are differences. Concerning whether the map helped the users trust the situational picture, Item 12, we have a tendency to a difference between the paper map (mean score 3.67) and the AR system (mean score 4.67), $F(2,42)=4.6, p\approx 0.051$. The symbols, Item 13 *The symbols helped me trust the situational picture*, helped the users more using the AR system (mean score 4.58) than the symbols on the paper map (mean score 3.50), $F(2,42)=5.1, p<0.05$, see Figure 10.

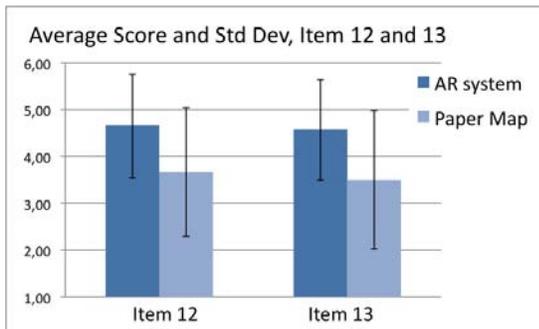


Figure 10: Results from Items 12, *The map/AR system helped me trust the situational picture*, and 13 *The symbols on the map helped me trust the situational picture*. There are tendencies to difference between the sessions, see text for further details.

Finally we had two items, Item 14, *I thought I had a good situational picture* and Item 15, *I thought the others had a good situational picture*, where users had to provide a more subjective view of

the situational picture. Our participants scored high on these items in all three situations, all above 4, but there were no significant differences between the sessions.

Overall the comparison between the AR system and the paper based map shows that the AR system is as good as or better than the paper map in many respects.

5.2 Evaluating the AR system

The questionnaire used to evaluate the AR system contained items specifically addressing the use of the AR system and did not include other aspects such as collaboration, see Figure 11. The queries were used in a previous study investigating AR systems for single users [23, 22], and here modified to reflect the task carried out in this study.

#	Question/Item	μ	95% conf. interval	
			Lower Bound	Upper Bound
1	It was easy to use the AR system.	4.21	3.796	4.621
2	The map was difficult to read due to the colours.	4.63	4.085	5.165
3	The map was difficult to read due to image resolution.	3.08	2.528	3.639
4	The map was difficult to read due to the symbols.	3.38	2.835	3.915
5	It took a long time to learn to use the AR system.	4.96	4.423	5.493
6	There were too many symbols in the AR system.	4.58	4.028	5.139
7	There were too few symbols in the AR system.	3.86	3.328	4.422
8	I thought that the interaction tool was easy to use.	3.86	3.274	4.476
9	I would like to use the AR system in my work.	4.46	3.930	4.986
10	I would like to use the AR system in other situations than work.	3.79	3.086	4.496
11	I felt confident that the AR system gave me correct information.	3.86	3.261	4.489
12	The AR system was clumsy to use.	3.92	3.361	4.472
13	The AR system had no major flaws.	3.04	2.521	3.562
14	I felt sick during the trial.	5.75	5.314	6.186
15	I felt dizziness during the trial.	5.50	5.002	5.998
16	I experienced other discomfort during the trial.	5.04	4.343	5.740
17	The AR system was fun to use.	5.71	5.512	5.904
18	It took a long time to learn how to cooperate using the AR system.	4.67	4.143	5.190

Figure 11: AR system questionnaire, average score and lower/upper confidence bound on the 6 point Likert scale. 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 and a low score is negative for the AR system

The participants found the system easy to use and learn, as seen in Item 1, *It was easy to use the AR system* and Item 5, *It took a long time to learn to use the system*, with the mean scores of 4.21 and 4.96⁴ respectively. They had only used the AR system that day but had no difficulty using it. Learning to cooperate using the AR system was not a problem either, Item 18 *It took a long time to learn how to cooperate using the AR system*, scored 4.67.

The participants liked to use the system. They gave high scores, 4.46, on Item 9, *I would like to use the AR system in my work*. They were slightly less interested to use the system in other situations, Item 10, scored 3.79.

On the general open ended question on what they thought of using AR technology in these types of situations in their everyday professional life, *What was the best about the paper MR system for collaboration (one or several things)?*, several users pointed out that they do think that this technology can be a reliable help:

⁴Note that we have transformed the scores for consistency in the table meaning that a high score on this item is positive for the AR system, i.e. positive, indicating it did not take a long time

”Good! It gives a credible situational picture and when you are secure in using the system my hope is that you can focus more on your task and less on verbal communication” (RS0924, question 3:6)⁵

The quote also highlights the key issue of training and practise in order to fully take advantage and feel secure in the technology they use:

”Fun, but takes practise so people are comfortable using it” (HP0925, question 3:6)

The participants trusted the system, Item 10 *I felt confident that the AR system gave me correct information* scored 3.86.

The questionnaire also addressed the AR system display, in Item 2, Item 3 and Item 4. The users had no problems reading the map due to the colours, Item 2 *The map was hard to read due to the colours* (mean score of 4.6), but on Item 3, *The map was hard to read due to picture resolution*, we see that the users are less positive (mean score of 3.1). We believe that this is due to the instability of the image, i.e. when they focus on the outer parts of the map the system sometimes loses the marker and hence the image projected to the user, as explained in the open ended question *What were the worst aspects of the AR system for collaboration?:*

”That the map image disappeared when I looked at the edges of the map. I felt somewhat isolated from the surrounding.” (P0925, question 4:2)

The users did not use the ability to switch between showing all symbols and only their own symbols as frequently as we had anticipated, and consequently they found the map a bit cluttered, Item 4 *The map was hard to read due to the symbols* scored a mean of 3.38. However, the AR system seems to be experienced as less cluttered than the paper based system, see Item 9, *The map became cluttered*, in the questionnaire on collaboration, Figure 9. We believe the main reason for not using the functionality allowing them to declutter the map by choosing to see only selected symbols, is that it was a bit too cumbersome to use the interaction device. Users had to manually select each organisation that they did not want to see in a menu.

When asked about specific features in the AR system in the open ended section of the questionnaire, 14 of the 27 participants said that they did use the feature allowing them to see additional information about the objects on the map *How often did you use the possibility to see additional information?.* Of the remaining participants several would have used it if they had had more time and/or training.

”My ambition initially was to to work with ①. However, I chose not to focus on this in order to save time. It’s a necessary feature for the future in my opinion” (RS0924, question 4:6)

The interaction device was evaluated in Item 8, *I thought that the interaction device was easy to use* (mean score of 3.9), and to a certain extent in Item 12, *The AR system was clumsy to use* (mean score of 3.9), and Q13, *The AR system had no major flaws* (mean score of 3.0). The rather low scores on these items can to some extent be explained by the result from the responses to the open ended question *What are your general impressions of the AR system?:*

”It needs further development in order for it to be more user-friendly” (RS0929, question 4:3)

⁵In the following text quotes of the participants are coded as follows: the first letter/s indicate organisation (P-Police, RS- Rescue services, HP-Helicopter Pilot), the following four numbers are the team number, and the final number combination indicates which questionnaire and open ended question the quote is related to.

The main issues of concern for improved user-friendliness are related to the symbol menu and the lack of shortcut buttons on the joystick interaction device:

”Flipping through units and icons. There is no need for RS [rescue service] vehicles in the police joystick for example” (P0923 question 4:3)

The interaction device and the symbol management in the design is also the main concerns for the participants when asked about what they would change in the system *What would you change in order to improve the collaboration over the paper map/MR system?:*

”Make the symbol management simpler” (HP1001, question 3:5)

”Move the symbols by pointing and dragging the finger” (P0925 question 3:5)

It is evident that the open approach derived from the design phase, where the decision was made to make all symbols visible to all participants, was misguided. This illustrates one of the difficulties in all design processes - a feature may be desirable in one cycle of development but perhaps not in the next (the envisioned world problem). Including more iterations in the process will likely reduce this problem.

Two items addressed the number of symbols, Item 6 *There were too many symbols in the AR system* and Item 7 *There were too few symbols in the AR system* scored 4.58 and 3.86 respectively. Again positive results, maybe adding some symbols, which poses no problems.

Addressing the more ergonomic or physical aspects of the system were Items 14, 15 and 16. As can be seen in Table 11 the users did not experience feeling sick, Item 14, dizziness, Item 15, nor did they feel discomfort, Item 16, due to using the AR system. The discomfort they felt appears to be mainly due to the head-mounted system. It became too heavy after a while as illustrated by this quote in response to a question about the negative aspects of using the system *Did you experience anything as troublesome, and if so, what?:*

”Standing for a long time with a front-heavy headset.” (HP0925, question 4:4)

One important aspect of interacting with systems is whether or not the users enjoy working with the system. The result indicate that they did, which is evident by viewing the score on Item 17, *The AR system was fun to use*. The result has an average score of 5.6 on a 6 point scale. In the open ended question regarding any positive aspects of the AR system, several issues were highlighted *Did you experience anything as positive, and if so, what?:*

”Easy to get an overview, good to get information about people, hours of every unit. Also the possibility to see other units information. Easy to move units” (HP0930, question 4:5)

”Good and clear situational picture that is shared and updated” (RS0930, question 4:5)

”You learn quickly. Good situational picture” (P1003, question 4:5)

These responses are important in relation to the purpose of the application. The aim of the study was to implement and develop a system supporting the establishment of a common situational picture among participant from different organisations. The quotes above illustrate that in many aspects the AR system succeeded to

aid the users in achieving an overall situational picture. However, as the results indicate there are several important issues to address in terms of interaction design, symbol management and the general physical design of the system.

The open-ended part of the questionnaire also included a question regarding whether the participants could see any other potential use for AR technology than this application and several users did *Is there any other situation where a system like this would be useful?:*

”Education, team work which requires a separation of tasks” (HP0929, question 4:9)

”Practise navigation on your own ship but still in dock, i.e. you fine tune the things you’re supposed to do in reality. Real time supervision of divers in the water, i.e. you have a sea chart with 3D (available in military environments)” (P1003, question 4:9)

A common theme for most suggestions of potential use of AR include training, simulation and strategy testing before an operation. The last quote above does however indicate another direction of real time use of AR in an ongoing operation or event as does this quote:

”An extremely interesting and exciting way to build an operation. I can already now see the advantages with placing different groups and units in connection to a larger football [soccer] command” (P0930, question 4:3)

The final part of the user study was the focused group discussion during which the participants discussed the experiences of the study, the AR system and their collaboration. One interesting issue raised during several of these focus group discussions were how the AR system could be used to share and distribute the command team’s situational picture to other people outside the control room. For instance the potential to work collaboratively but distributed as they do now but instead of having only telephone contact with commanding strategic officers at ‘home’ they could all share the same view of what is going on in the field, in real time.

6 IMPLICATIONS FOR DEVELOPMENT OF COLLABORATIVE AR SYSTEMS

Designing a single user AR application with a sequential interaction scheme is in many respects relatively easy compared to designing applications for multiple users performing dynamic interactive tasks. Simply converting a single user application into multiple users may be technically relatively easy but the complexities of the task demand careful consideration before adapting a single user system into a multiple user one (which is often the case in many computer supported collaborative applications [7]).

The results of the study clearly indicate that the AR system was experienced as a possible future system not only for the task used in the scenario but also for other tasks within the three organisations. AR technology is not limited to creating and maintaining a common situational picture, it also has potential for training and information sharing throughout the chain of command and control. The most important lessons learned during this study of implementing this AR application for use in a complex, dynamic scenario for emergency management can be summarised as follows:

- Involve real end users in the design of the system/application development process
- Involve real end users in the design of the user study task
- In order for real end users to feel involved in the task, make sure the task, or scenario, is realistic and not artificial
- Do several design iterations
- Involve real end users in the evaluation of the application

6.1 The iterative design process

Working iteratively with re-design and evaluation, involving real users is invaluable for several reasons. Firstly, it allows us to cope with the envisioned world problem – by moving from making major changes to the system to minor changes, the discrepancy between the envisioned usage and the actual usage slowly erodes. It is however important to remember that the process is two-sided; the original vision also changes as the designer begins to understand possibilities and limitations with the system.

The iterative design of this application is by no means finished with this end user study. As noted in the results, the interaction device needs to be carefully designed to facilitate a situation where users not only access and modify their own objects and information but also can access and modify objects and information from other users. Not only is the interaction more complex than in a single user application, as there are many more symbols to manipulate, users also manipulate their own symbols more frequently than the others’ symbols. Regarding specific design features the next iteration of this application will make sure to simplify the interaction modality even further, taking note to the participants comments. The click-and-drag feature requested by one participant was considered during the development phase but was unfortunately not implemented due to time/resource constraints at that time, but has been implemented since. The menu has also been restructured to allow faster navigation.

6.2 Using AR to establish a common ground

One of the most interesting findings in this study was the fact that the participants in most issues gave the AR system an equal or better score than the regular paper map. The quote by one of the fire fighters above (RS0924, question 3:6) gives a clear indication that the AR application designed has in one important aspect reached its purpose. One of the most problematic issues during command and control work in the field is the messy and cluttered map over ongoing events. As several different actors give their input, and their view of what is going on and what needs to be done in what order, the notes and sketches tends to pile up, leaving a very difficult to read map (or white board or paper) to interpret. The AR system allows all individuals to work on the same map, or in the same interaction space, both individually as well as collaboratively. But the added benefit of the AR system, compared to a paper map, is that it is possible to quickly switch perspectives and follow one organisation at a time, as well as see the overall view of all available resources and their status and distribution. The experience of the AR system as less messy and cluttered than the paper map (Item 9) in the first questionnaires illustrates this issue. Even though the participants felt that they had a good situational picture in both settings, the clutteredness of the paper map compared to the AR map significantly affected their rating of the paper map versus the AR system. The participants see a great potential in the AR system to present them ‘with a credible situational picture’ allowing them to focus more on their actual task at hand rather than spend time on verbal communication, talking about what resources are where and when.

Paper maps are something that these participants are very used to working with, whereas this study was the first time they ever encountered AR technology. The high scores given to the AR system indicate that they actually can perform their tasks to the same level of satisfaction as they normally perform, i.e. with a paper map. 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.

The results of the study are also successful in relation to the demands made by Gutwin and Greenberg [11] regarding team cognition, since the users did not seem to be hampered in their joint ef-

fort of creating a shared situational picture. The system thus seems to provide enough feedthrough for joint work, possibly because it allows gesturing and joint manipulation of symbols. The more specific aspects of work provided by the participants, like the wish for an extended symbol library, is probably a result of the effort of using real participants, as proposed by [5] and [26]. Although an artificial task and non-professional users probably could have provided basic usability input such as the experience of motion sickness, image quality etc, these task-specific findings are only likely to emerge in more work-like settings with real users. The participants in this study, the fire fighters, police officers and helicopter pilots have the specific knowledge and experience to assess the AR application's potential in their professional work in a way no novice user could.

7 CONCLUSIONS

This paper has described an AR application developed in close cooperation with real end users, and illustrated how an iterative design and evaluation method can be used in this field. The results of this study has illustrated that although collaborative command and control is a rather complex field, with ever changing needs and tasks, AR as a technology can, if carefully designed, be successfully used for collaboration in dynamic tasks.

ACKNOWLEDGEMENTS

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, Hanna Mårtensson and Erik Prytz 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. Collaborative augmented reality. *Communications of the ACM*, 45(7):64–70, July 2002.
- [3] H. H. Clark. *Using Language*. Cambridge University Press, Cambridge, 1996.
- [4] M. Cross and C. Bopping. Collaborative planning processes in command and control. In *Fourth International in Command and Control Research and Technology*. DoD CCRP, 1998.
- [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, 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.
- [7] M. Fjeld, K. Lauche, M. Bichsel, F. Voorhorst, H. Krueger, and M. Rauterberg. Physical and virtual tools: Activity theory applied to the design of groupware. *Computer Supported Cooperative Work*, 11(1/2):153–180, 2002.
- [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] R. Grasset, P. Lamb, and M. Billinghurst. Evaluation of mixed-space collaboration. In *ISMAR '05: Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 90–99, Washington, DC, USA, 2005. IEEE Computer Society.
- [11] 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.
- [12] M. Haller, M. Billinghurst, and B. H. Thomas. *Emerging Technologies of Augmented Reality*. IGI Publishing, Hershey, PA, USA, 2006.
- [13] 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.
- [14] 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.
- [15] E. Hollnagel. Is affective computing an oxymoron? *Int. J. Hum.-Comput. Stud.*, 59(1-2):65–70, 2003.
- [16] E. Hollnagel and D. D. Woods. Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies*, 18(6):583–600, 1983.
- [17] E. Hollnagel and D. D. Woods. *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. CRC Press, Boca Raton, FL, 2005.
- [18] B. Johansson, M. Persson, R. Granlund, and P. Matsson. C3fire in command and control research. *Cognition, Technology & Work*, 5(3):191–196, 2001.
- [19] 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*.
- [20] G. Klein, P. J. Feltovich, J. M. Bradshaw, and D. D. Woods. *Common Ground and Coordination in Joint Activity*, volume Organizational Simulation, chapter 6, pages 139–184. John Wiley & Sons, 2005. <http://dx.doi.org/10.1002/0471739448.ch6>.
- [21] C. McCann and R. Pigeau. The human in command. In M. C and P. R, editors, *The Human in Command; Exploring the Modern Military Experience*. Kluwer Academic/Plenum Publishers; New York, 2000.
- [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] P. Renevier and L. Nigay. Mobile collaborative augmented reality: The augmented stroll. *Lecture Notes in Computer Science*, 2254, 2001.
- [26] R. Samuracy and J. Rogalski. Cooperative work and decision making in emergency management. *Le Travail Humain*, 56:53–77, 1993.
- [27] 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*.
- [28] 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.
- [29] M. Träskbäck. Toward a usable mixed reality authoring tool. In *VL/HCC*, pages 160–162. IEEE Computer Society, 2004.
- [30] D. D. Woods and E. Hollnagel. Mapping cognitive demands in complex problem-solving worlds. *International Journal of Man-Machine Studies*, 26(2):257–275, 1987.
- [31] D. D. Woods and E. M. Roth. Cognitive engineering: human problem solving with tools. *Hum. Factors*, 30(4):415–430, 1988.