Effects of making unknown information explicit in computer-supported collaborative information retrieval

An empiric study in an urban search and rescue setting.

In partial fulfilment (30 ECTS) of the requirements for the degree of
Master of Science
in
Cognitive Artificial Intelligence

Author:
M.J.R. van den Heuvel
Studentnumber: 3020355
Cognitive Artificial Intelligence
Faculty of Humanities
Utrecht University

Supervisors:
prof. dr. J-J.Ch. Meyer
Utrecht University
Drs. N.J.J.M. Smets
Dr. J. van Diggelen
TNO Soesterberg
Additional reviewer:
Dr. S.F. Donker
Utrecht University

May 2013

There are known knowns; there are things we know we know.
We also know there are known unknowns;
that is to say, we know there are some things we do not know.
But there are also unknown unknowns – the ones we don't know we don't know.

–United States Secretary of Defense, Donald Rumsfeld

Acknowledgements

A lot of people have inspired and supported me during the entire process of this research. Without these people, I would not have been able to carry out the research to this extent.

Firstly I would like to thank my girlfriend Lisette, for her support and understanding during all the busy periods. I am convinced it was not always easy for her to listen to setbacks and my occasional less optimistic views on the whole process. I truly appreciate her patience and support during these times.

Secondly I would like to thank my supervisors at TNO, Jurriaan van Diggelen and Nanja Smets. Jurriaan supervised me during my internship, providing me with the initial topic of research and guidance on setting up the experiment. Jurriaan's input formed the basis for the entire research. Nanja supervised me during the actual experiment and afterwards during analysis, constantly pointing me in the right direction and anticipating any potential obstacles. Also, Nanja kept me motivated at any time, which for me was a necessity to get things done at some times. Her personal advice on topics other than the experiment and graduation internship is also greatly appreciated and was of great importance for my forthcoming career.

My gratitude also goes out to other people from TNO. Mark Neerincx was the first one to interview me when I applied for the internship at TNO. Without this positive interview with Mark, the internship at TNO might not have taken place at all. Moreover, Mark provided great insights in setting up the experiment and provided valuable feedback on analysis of the results.

Tjerk de Greef inspired and contributed in designing the task of the experiment, indicating points of interest and possible bottlenecks in the procedure.

Pjotr van Amerongen was essential for the process, as he developed a lot of code necessary to conduct the experiment. Even more appreciated are his efforts to help me out at any time, especially the stressful moments before commencement of the experiment itself.

Martin van Schaik provided me a lot of information on statistical analysis of the results. Since I was a novice in statistical analysis, Martin helped me in deciding which steps to take and which tests to use for analysis. This developed my knowledge in statistical analysis a lot, being a huge benefit in my future work. Also, his positive words when results turned out different than expected are appreciated.

Last but not least, I would like to thank all the other interns at TNO. Partially due to the fantastic sense of humor of these interns, I had a great time at TNO. Moreover, the interns gave me essential feedback on numerous occasions. Firstly, Corine Horsch helped me a lot during the first weeks at TNO, introducing me to the NIFTi project and the organization. Also, her research was of great help and inspiration for my own research. Being able to reuse some of her experiment's materials made it possible to conduct my experiment in the limited time I had. Tinka Giele and Christian Willemse gave me a lot of valuable feedback during the experiment and on analysis of results. Without their views on the small issues, I would never had kept a clear view on the process.

Abstract

When actors have to perform tasks in a team, whether the actors are humans or robots, communication and cooperation is of utter importance. Adequate information sharing and handling by team members can increase the performance of the whole team. A contemporary field in which this is important is urban search and rescue (USAR). Currently, in USAR situations the role of tele-operated robots increases in importance. Robot operators are able to share information in a common operational picture in order to create situation awareness. In this common operational picture, every team member can contribute instead of only one.

This thesis describes a between-subjects experiment to test a specific functionality, the *Explicit Unknowns functionality*, within such a common operational picture that enables the users to make relevant, unknown information explicit. In other words, users can indicate which information they need to know but cannot provide themselves. Team members with the right capabilities are made aware of this unknown information and could improve the shared situation awareness by answering these unknowns, improving team performance in turn.

Expected was that the use of this functionality would improve the efficiency of rescuing victims, cooperation, situation awareness and subjective performance. Measures were taken by individual questionnaires and registrations in the common operational display.

An experiment was conducted using a simulated environment in which participants were paired as a team. Each team member had different capabilities and teams were given the task to fully register as much found victims as possible in the common operational picture. Certain victims required specific information that could only be provided by one of the team members. These cases required cooperation between team members, which is why completeness was essential in the task. The explicit unknowns would enable team members to work more efficiently and hence perform better as a team.

Results, however, did not show significant effects between conditions. Nevertheless, tendencies were noticeable and seem most promising for cooperation, subjective performance and search strategy. While not significant, these tendencies did indicate slight improvements when the Explicit Unknowns functionality was available. Subjective feedback from participants also indicated a positive tendency towards the used functionality.

Therefore, the main conclusion in this research is that making unknown information explicit in collaboration tasks might yield effects for some topics, although it is essential to make changes in experiment setup and team formation.

Contents

1 Introduction

This thesis is written as part of a Master's study in Cognitive Artificial Intelligence. Cognitive Artificial Intelligence (CAI) is a multidisciplinary study, touching aspects of many different fields. Cognitive sciences, linguistics, computer sciences and philosophy are some of the fields of interest for CAI, resulting in a broad, general knowledge within these fields. This is an advantage in situations where separate existing fields need to be connected or have common properties.

Human-machine interaction can be seen as a unique field of cognitive and behavioral sciences intersected with the field of computer sciences as well as robotics. The cognitive aspect of CAI is addressed in this thesis by the focus on team performance between human actors using a computerized system as a tool. Human-machine interaction also applies to the concept of using computer-supported cooperative work systems to increase human performance. Central in the research in this thesis was the use of the *Explicit Unknowns* functionality, which is part of such a system and thus plays an important role in the interaction between human participants and computerized systems. This invigorates the role of CAI, connecting different fields of research resulting in practical applications.

1.1 USAR and NIFTi

Urban search and rescue (USAR) involves the first emergency response which deals with the rescuing of victims in an urban scenario [16]. Examples of such scenarios are collapsed buildings, tunnel accidents or natural disasters. Conditions in such situations are very poor; physical and emotional challenging circumstances impact a worker's competencies, motivation and cognitive abilities [4]. This, in turn, impacts teamwork, cooperation and team structure.

The use of computerized systems in high-risk professional domains, such as USAR, follows trends to increase efficiency and safety in cases where manpower is reduced or situational complexity is increased [3]. The use of such technologically advanced systems requires a seamless integration of human and machine capabilities in order to coordinate activities, which in turn prevent redundancy of work as well as ineffective deployment of resources.
Examples of such technologically advanced computerized systems are (shared) observability displays and purpose-built robots.

Nowadays the use of robots for such USAR missions is getting more ubiquitous because rescue workers are exposed to great dangers. Robots are replaceable and more agile in certain environments due to their small size. Human-robot cooperation is a fairly young field of research investigating the important link between human users and mechanical actors. In most automation techniques, the humans were simply taken ”out-of-the-loop”, resulting in loss of skill and situation awareness for operators [7]. This is not the same as actual *collaboration* between humans and robots; bringing the human back ”in-the-loop” might therefore improve performance and situation awareness in automated processes.
Urban search and rescue is a practical field of research where the human-robot collaboration can be explored. The NIFTi[1] project explores this field, researching human-robot cooperation in urban search and rescue situations. Making the robot a team player is one of the main aspects in this project. However, since robots in this field are not yet fully autonomous, they are mainly tele-operated by trained robot operators [15]. Human rescue teams use robots as a tool while an overview of the situation is maintained and distributed by a mission commander [14].

[1] see http://www.nifti.eu

1.2 Collaboration

For collaboration, whether it is between humans, machines or a combination of both, the concept of *joint activity* plays an important role in order to achieve a *joint goal* [3]. Joint activity is defined as an activity *"that is carried out by an ensemble of people acting in coordination with each other"* [2,3] and leads to measurable joint performance. Rescue workers operating in a team need to know relevant information; e.g. where they are, where other team members are, what these team members are doing and where potential hazards are. Each team member might have an individual task, role, different capabilities, priorities and resources. All these differences combined with a stressful situation and hazardous environment cause collaborative information retrieval to be subjected to errors and deficiencies or superfluities. Computerized systems are a successful aid to gather, process and distribute relevant information [3, 9]. This shifts the activity of keeping a situational overview from mission commander to every team member, decentralizing the concept of this situational overview. Loss of information by communication is prevented while situation awareness is distributed over the whole team. However, not every team member might be able to provide adequate and complete information, adding mere shreds of information to the shared display. Team members are not aware of this lack of information and can therefore not anticipate on this informational absence. Making the unknown information explicit to users could improve not only individual performance [18] but also collaboration, increasing team performance.

1.3 Explicit Unknowns

Users need to be made aware of required, but still unknown information. The quote from Donald Rumsfeld[2] implies there exists information of which it is unknown that it is unknown. When it is known which information is unknown, unknown unknowns will become known unknowns. An example can be found in the history of cartography and the discovery of land. In the fifteenth century, unexplored land was labeled on a map with the term *"Terra Incognita"*, meaning unknown land. This indicates it is only known that there exists land at that place, but nothing is known about this land. It is known this land is unknown, which is a substantial difference compared to mere unknown, uncharted land. The latter could be seen as an unknown unknown. The research described in this thesis is based on a functionality within a computerized system to make users aware of unknown information, the so-called *Explicit Unknowns functionality.* Making specific users aware of unknown information relevant to their interests (for example based on their role, function or location) and presenting these unknowns in an appropriate manner can help them direct their attention. This could result in a safer and more efficient accomplishment of tasks.

Making unknown information explicit by means of a dedicated *Explicit Unknowns* functionality as well as aspects of team performance and collaboration are central in the research described in this thesis. An important property of the functionality is machine readability. Due to the underlying framework, information is semantically structured and is, in a sense, interpretable by machines. Therefore, the used system is not limited to use by human actors, it may also be applied on other actors like autonomous robots. The Explicit Unknowns system abstracts content gathering, presentation and delivery in such a sense that it creates a framework that may be used by any type of actor, be it human, virtual or mechanical. Thus, besides a purely cognitive approach, technical sciences are also a substantive factor. In the use-case of a USAR scenario, an autonomous robot might automatically detect victims, add them to the common operational picture and use the Explicit Unknowns functionality to indicate unknown or uncertain information [5, 14]. The used functionality is by all means not limited to this specific domain

[2] See the epigraph.

and might be applied in several other domains where collaborative information retrieval or other collaborative work is central.

These, and other relevant topics will be elaborated on in section 2. The following section will expound on this thesis' research question and corresponding hypotheses. Section 3 will describe the method used to research the influence of the Explicit Unknowns functionality, while section 4 shows the results obtained by this method. Section 5 will explain the interpretation of these results and a conclusion is drawn in section 6.

1.4 Research question

The main research question of this thesis is: *"What is the influence of making unknowns explicit on human collaboration in a non-protocolled information retrieval task"*. To answer this question, several hypotheses were developed to be able to test different aspects of the main research question. The following hypotheses were designed based on the usage of a shared observability display, known as the *common operational picture*, **with** the functionality to make unknowns explicit compared to usage **without** this functionality.

Hypothesis 1: Expected is that the availability of the Explicit Unknowns functionality will make the rescuing of victims more efficient.
While the use of the functionality is not compulsory, operators will make use of the functionality. If an operator knows where his team mate already has been, redundancy in searching is reduced. Also, when an operator knows which victims need help in order to be rescued, the operator can direct his or her attention to these specific victims. This will mean limited resources like attention and time will be applied to victims in an efficient manner, increasing team performance.

Hypothesis 2: The availability of the Explicit Unknowns functionality will make cooperation more efficient.
The Explicit Unknowns functionality will provide more insight for team members in what their team mates will do in the near future. For participants, this will result in a more positive assessment of their degree of cooperation when the functionality is used. Also, it is expected that the forms used in the shared observability display will have less *incomplete fields*. Due to adequate cooperation, fields that are unanswerable by one team member will be answered by a team mate, leaving less fields unanswered for living victims.

Hypothesis 3: Situation awareness improves when the Explicit Unknowns functionality is available to operators.
Expected is that participants feel they have more knowledge about the environment and situation when they have used the Explicit Unknowns functionality. The use of the functionality makes a team member more aware of several aspects; what his or her team mate is doing, the information that is added to the common operational picture and where help is required. This all adds to an overall higher assessment of a team member's awareness of the situation. When team members do not have to direct their attention to low-level aspects of the environment, attention can be utilized for higher levels of situation awareness, improving overall situation awareness.

Hypothesis 4: Communication will be less verbal when the Explicit Unknowns functionality is used.
Since the Explicit Unknowns functionality attends team members on where help is requested, explicit verbal communication about this will be superfluous. Therefore, expectations are that

participants will assess their communication as less verbal when Explicit Unknowns are used. This means relatively more information is communicated asynchonously by means of the shared observability display instead of directly and verbally.

Hypothesis 5: Users feel they have performed better when they have used the Explicit Unknowns functionality.
Participants will assess their performance as a team higher since the functionality makes them aware of victims where help was required. Without the functionality there may exist uncertainty about this, decreasing the overall performance assessment.

Hypothesis 6: A different search strategy is applied by participants when using the Explicit Unknowns functionality.
Expected is that participants will search for victims in a more individual manner, keeping a larger distance to their team mate, since the Explicit Unknowns functionality shows them when help is needed. Without this functionality participants will stay close to each other to provide direct assistance when needed, with a possible trade-off in the number of found victims.

2 Related work

The following section will address literature relevant to the research of this thesis. To our best knowledge, specific research on making unknown information explicit has barely been done. As an actor in an USAR situation, being aware of specific unknown information can influence decisions, which in turn may influence situation awareness and performance. The use of computerized systems to share information is ubiquitous and can aid actors in sharing known knowledge as well as creating awareness of unknown knowledge. A good functioning team can perform beyond the sum of performance of the individuals in a team, because a specialist always is more effective and efficient than a generalist (Kenrick, Neuberg, & Cialdini, 2010, ch. 12 as cited by [13]).

Since the empirical research in this thesis makes use of concepts and techniques on these topics, the following sections will elaborate on these topics in order to provide information on their current state.

2.1 Situation awareness

Situation awareness is an important element in an USAR situation, especially when human-robot teams are involved [17]. Not only for individuals, but also for teams, situation awareness describes in what sense operators understand and anticipate on their environment in a certain situation.
Endsley [6] states *situation awareness* (SA) can be seen as an internalized mental model of the current state of the operator's environment, creating an integrated whole of the incoming data from all sources. This integrated picture forms the basis for all decision making and all actions. To perform and decide successfully in complex situations, an operator has to develop situation awareness and keep it up to date.
Situation awareness only exists in the mind of the human operators, hence it is useless to present vast amounts of data to such operators unless this data is successfully transmitted, absorbed and assimilated in a timely manner by the operator. Endsley's formal definition for situation awareness is *the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and projection of their status in the near future.* This general

definition is split up in three levels of situation awareness.

Level 1: Perception of the elements in the environment. An actor in an environment needs to perceive the status, attributes and dynamics of relevant elements in the environment in order to achieve situation awareness. This first step differs between actors as well as situations and is about getting aware of an actor its environment.

Level 2: Comprehension of the current situation. This includes assessing the significance of elements relevant to an actor's goals. An example is a warning sign which has to be interpreted for importance to determine subsequent actions.

Level 3: Projection of future status. The highest level of situation awareness allows operators to meet objectives by anticipating on the information resulting from previous levels of situation awareness. Endsley states that this ability to project is critical in many domains.

According to Endsley [6], all these levels are based on mechanism of short term sensory memory, perception, working memory and long term memory. Limited attention in perception has a negative influence on situation awareness, while (implicit) knowledge from the working memory and long term memory shape the perception of an operator. In most situations, working memory is heavily addressed to achieve higher levels of situation awareness; comprehending the situation and planning actions based on this information. The expertise of the operator influences the situation awareness, since experience develops long-term memory structures, these can be utilized in the higher levels of situation awareness, putting less load on the working memory. Also, an operator's goals determine what aspects are attended, and therefore influence all levels of situation awareness.

2.2 Collaborative information retrieval

In an USAR situation, a team of actors tries to create situation awareness by gathering information about the environment and situation. Therefore, team members have to collaborate in their information retrieval. Computer-supported cooperative work (CSCW) systems can be regarded as a useful tool for this task.

Foley and Smeaton [9] describe two important concepts for synchronous collaborative information retrieval (SCIR); *division of labor* (DoL) and *sharing of knowledge* (SoK). SCIR is literally described as *supporting two or more users who search together at the same time in order to satisfy a shared information need.* Therefore, searching for victims in an USAR mission can be viewed as collaborative information retrieval. Foley and Smeaton [9] state collaboration is common in synchronous information retrieval and occurs in any stage of the information retrieval process. CSCW systems, based on DoL and SoK enable users to execute coordinated collaborative information retrieval such that redundancy is reduced while group members can benefit from the discoveries of their collaborators.

DoL is used as a concept for dividing group tasks in order to share workload across collaborators, making the group function as a whole due to less redundancy and increased efficiency. In an USAR situation collaborators in a first-response team often are familiar with a structured division of tasks due to their background (e.g. firefighters) [16], hence DoL is essentially already present but must be acknowledged nevertheless when dealing with a domain different than USAR. According to Foley and Smeaton, if DoL is done effictively it enables the group to cover more of the information space. Translating this *information space* to the USAR domain, would mean more physical area could be searched.

SoK is the passing of information between collaborators during a group activity. Effectively sharing information is recognised as an important foundation of any group activity [19]. The personal experiences and expertises of a collaborator are introduced in the shared task, therefore the SoK should provide a seamless exchange of information across the group where users benefit from the expertises of their collaborators [9].

Observability display

Besides CSCW systems, observability displays play an important role in the current research. The concept of an observability display is a subset of CSCW systems and provides users with an interactive tool to share information and collaboratively build on the team's situation awareness.

De Greef [3] describes several experiments to test the use of electronic partners in dynamic task allocation and coordination. Electronic partners are computerized systems *partnering with* humans, instead of *automation extending* human capabilities. It is claimed that coordination becomes more difficult when actors are separated in space or time. Inter-predictability is the capability to plan actions based on predictions of what others will do. Observability is therefore defined as the perception of what others in the environment are doing in order to determine the impact on joint activities. An observability display is envisioned to increase awareness within a team and allow the facilitation of coordination of joint activities and increased resilience to unexpected events.
One of the experiments in de Greef's work describes an evaluation of an observability display in which actors had to collaborate in an USAR situation. A virtual environment was used to test whether performance and cooperation improved in situations where it was necessary to deviate from a predefined plan. Actors either had to find victims and report them on a map or assess information (triage) about these victims. Results showed an increase of observability when the display was used compared to when it was not used. When the observability display was used, participants found significantly more victims and assessed significantly more victims. A positive effect on coordination was observed, as well as a lower communication frequency, when the display was used.
Situation awareness did not improve when using the observability display, however it neither was significantly lower. Also, actors were more aware of, and able to cope better with, changes in the predefined plan when using the observability display. Concluded was that the use of an observability display leads to an increase in performance and satisfaction and ability to cope with changes in plan, while frequency of communication is decreased.

2.3 Explicit unknowns

Van Diggelen et al. [5] describe a framework intended for complex task environments that involves collaboration of networked computing devices, sensors, robots and human users. The $(RM)^3$ framework (Right Message at the Right Moment in the Right Modality) is designed to make optimal use of vast amounts of digitally stored information in a repository. An agent-based architecture is proposed to make decisions at runtime in order to deliver specific content to specific users in an apropriate way. Ontologies are a key part, describing classes and relationships of content, as well as users and devices. Figure 1 shows a high-level overview of the agent-based framework within the $(RM)^3$ architecture. The framework allows users to add information to a central repository by means of a common operational picture. The agent-based delivery of information to users allows selection of the content and mode of delivery in order to prevent unmanagable overload and interruptions.
This framework also supports a *Smart Questions* functionality, which allows users to mark fields explicitly as unknown; a *question* posed by a user that may be answered by a collaborator.

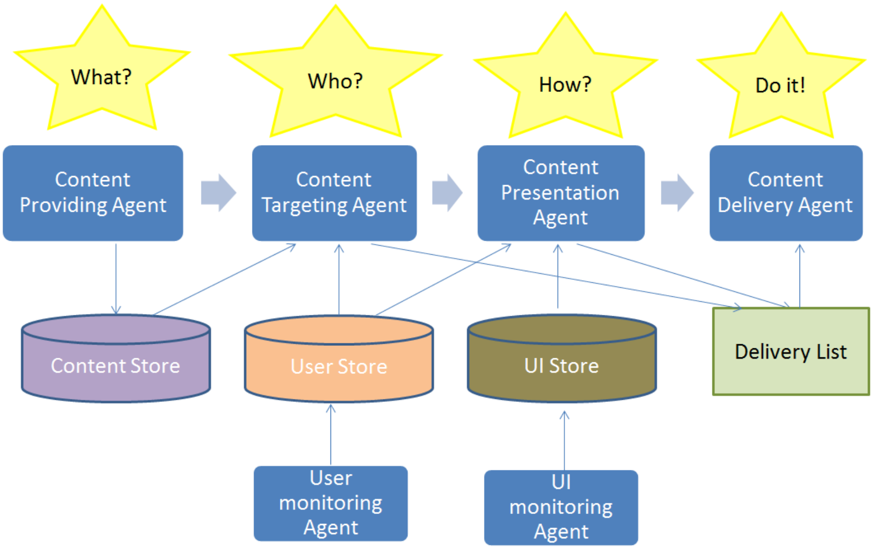


Figure 1: The agent-based framework of the $(\text{RM})^3$ architecture.

Based on the conceptual relations in the ontology and the decisions made by the agents, a message about such an unknown field can be delivered to a user capable of answering this field. This functionality translates to the *Explicit Unknowns* functionality used for the research in this thesis (see 3.2).

Wang et al. [18] describe an experiment in which unknowns were made explicit in order to research a person's trust in an automation. Participants had access to a system to classify targets in a military setting as friend or enemy. When the system was unable to classify a target as a friend, the target was deemed *unknown*. Differences between conditions were the degree of automation (manual versus automatic feedback) and the way in which the system provided feedback in case of unknowns ("no light" versus "red light"). Implicit feedback existed in the "no light" condition, where the automation did not send out any signal if it considered a target *unknown*. Explicit feedback was given in the "red light" condition, where the automation responded with a red light when a target was deemed *unknown*.

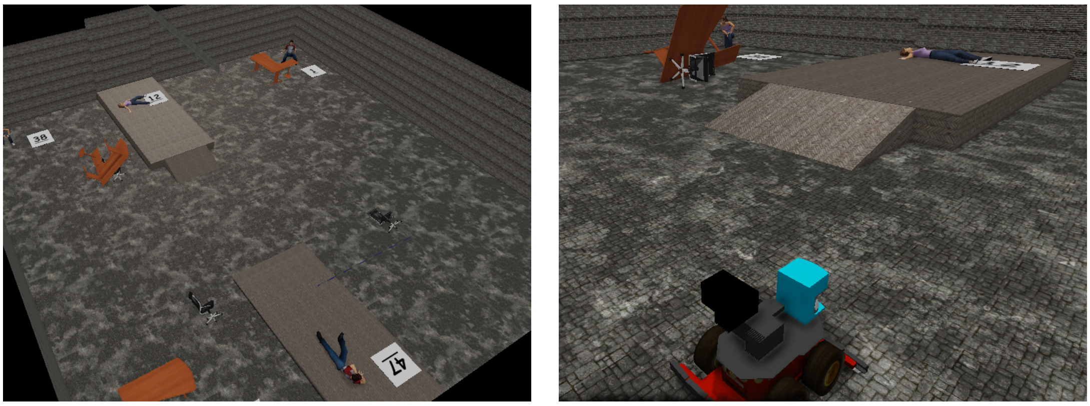
Results of this experiment showed that participants trusted the explicit "unknown" feedback more than the implicit feedback. They also responded quickest when explicit feedback in automatic mode was provided. While the research of Wang et al. aims at trust instead of performance, it indicates the importance of making unknown information explicit.

3 Method

The following section describes the method used for this research. It elaborates on the procedure and design of the experiment, the participants and the task they were given, used materials and the measures taken.

3.1 Procedure

Four participants conducted the experiment simultaneously. They were presented a consent form, providing substantive information about the experiment itself and the general procedure. Any questions or unclarities were answered and participants were given the first, demographic, questionnaire (see appendix B). Both teams were seated in separate rooms and given instructions about their task, robot and interface controls as well as a training procedure. Again, any unclarities were handled and the training was started. The training took place in a small virtual environment in which some obstacles and victims were placed, as can be seen in figure 2.



(a) Overview of the training environment.

(b) A UGV in the training environment.

Figure 2: The environment used for training.

During training, participants were taught the start-up procedures, basic controls and functions of the robot and the common operational picture[3]. Roughly 15 minutes were spent on training. When participants indicated that they finished the training procedure and understood the controls, the first session in the office environment was started. It was emphasized that they had 15 minutes to find as much victims as possible but more importantly, **register them completely** in the common operational picture. Registring victims completely was done by filling in all fields in the report form that was presented when a victim was added. Members of a team started their task simultaneously and were asked to start a timer in the common operational picture that indicated the remaining time.

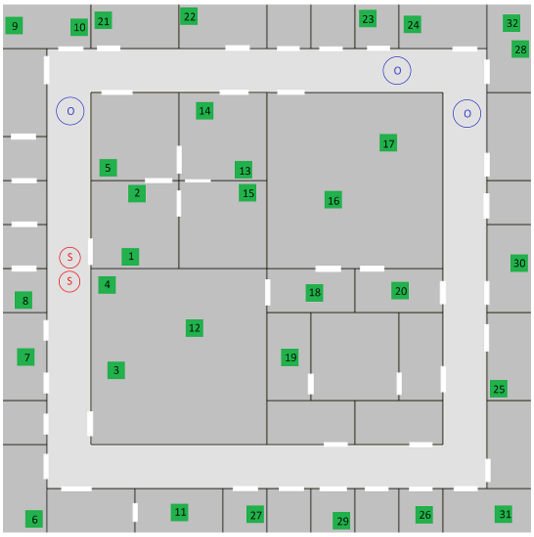
After 15 minutes participants were instructed to stop their activities and close all software. Participants filled in questionnaires regarding this session and all session data was saved on the server machines. After filling in the questionnaires, the four participants were recombined into two new teams.

The new teams were given the same instructions as for the previous session and started a 15 minute session in the same environment, which only differed in placement of victims and several obstacles. Again, questionnaires about this session were filled in afterwards. Next, participants were recombined into new teams for the last time and performed the experiment in yet another modified version of the same environment. Both the questionnaires about this session, as well as a questionnaire about all sessions in general were filled in by the participants. Finally, participants were thanked for their attendance and paid €30,- as compensation for their efforts.

[3]In the questionnaires, *tactical (awareness) display* was used instead of *common operational picture*.

3.2 Materials

Environment A virtual environment was used as a scenario for this experiment, in which simulated robots could be controlled. USARSim (Unified System for Automation and Robot Simulation) [1] is a high-fidelity simulation of robots and environments based on the Unreal game engine by Epic Games [10].
A collapsed office building was simulated in which 32 victims were placed, of which 12 victims were alive and 20 victims were deceased. Victims were placed arbitrarily in the environment. However, since living victims should force cooperation it was endeavored to distribute victims, especially the living ones, equally over the environment. Figure 3 provides an impression of the environment, whilst figure 3a shows an abstract overview of the complete environment, with one of three victim distributions indicated by green marks. Obviously, participants had only access to the map *without* the victim marks.



(a) Map of the environment.

(b) One of the rooms of the collapsed office.

(c) A living victim between debris.

(d) A deceased victim covered in debris.

Figure 3: The environment, including victims, used for the scenario.

Common Operational Picture Besides the simulation environment, participants also had access to the tactical awareness display, a *common operational picture*, in which they were able to share information. The underlying framework is a tagging-based real time exhibitor (TREx), maintaining the data for storage and delivery to users. This framework makes use of an ontological structure to model the relationships between concepts. Examples of these concepts are user properties (specific examples: capabilities, role, qualifications, location) and types of data. Based on these conceptual relationships and the (RM)[3] (see 2.3) architecture, data is stored in a repository and distributed to the correct users. Figure 4a shows several different types of datapoints, both UGVs and their travelled paths. This common operational picture shows users the current location and travelled path of their virtual UGV. The current location is an important reference point for the user when victims are found and need to be placed in the right area on the map. Also, knowledge of the location of another team member might influence operational decisions in a positive way. The travelled path is also important to see if the user itself or the other team member has already explored areas. This prevents redundant work since a user can assume an area has been successfully explored when the path indicates someone has been there, but did not report any information. This eliminates the need for a explicit indication to *clear* the area.

The display also enables users to add datapoints on a map, such as a *victim report*, *obstacle report*, *remark* or *snapshot*. Users could click the location on the map where they wanted to add a datapoint, select the type of data they wanted to add and were shown a form. Figure 4b shows a typical form for a victim report. The required information fields in this form asked for *gender*, *vivacity*, *injuries*, *heart rate* and *respiratory rate*[4]. When users submitted the form, the corresponding icon of the datapoint was displayed on the map in the main window (see figure 4a). Figure 6a shows the standard icon of a victim report. The icons of the datapoints could be clicked by any user to retrieve the entered information and modify or complement information fields.

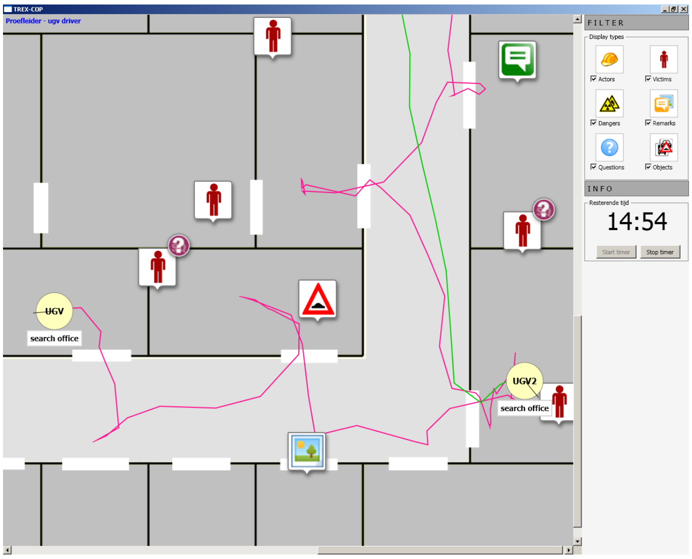
Additional functionalities within this common operational display are a timer to indicate the remaining search time and filter options to make a selection of what datapoints should be displayed.

Explicit Unknowns functionality The Explicit Unknowns system is an additional functionality within the common operational picture. The presence of this functionality was the difference between the conditions in the experiment. If a user adds a datapoint in the common operational picture, but is not able to answer all fields, the user can mark the field as being *unknown* by clicking a dedicated question-mark icon. Figure 5 shows typical information forms with the Explicit Unknowns functionality added to the common operational picture. Figure 5a shows the dedicated icons that users could click to indicate they were not able to answer the field. After clicking this icon, a border was displayed around the field (see figure 5b) to point out which field needs to be answered, and the icon in the main window was supplemented with the dedicated icon (see figure 6b).

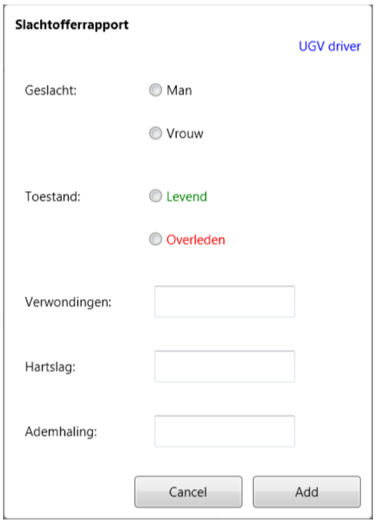
While this functionality might be fully automated, making users automatically aware of any *incomplete* form, it was decided to let users actively initiate an Explicit Unknown. Which information fields are in a form is decided beforehand and does not necessarily mean every field is essential in all cases. Therefore, in real-world situations, forms may be large and in most cases be incomplete. Automation would overflow the system with Explicit Unknowns messages, which are not all essential. Hence, the decision of which information is essential needs to be answered is left at the user.

In this specific experiment automation of the functionality would have been an option due

[4]Respectively *geslacht*, *toestand*, *verwondingen*, *hartslag* and *ademhaling* in Dutch on the form.



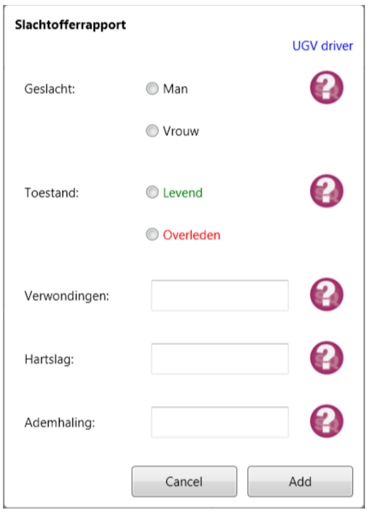
(a) Main window with map.



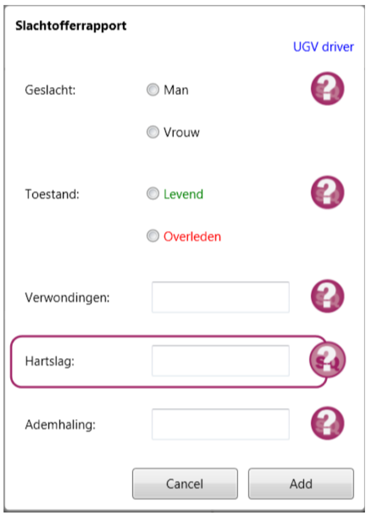
(b) Form of a victim report.

Figure 4: Screenshots of the common operational picture

to the small scale, small number of users and few fields in the information forms. However, automation would change the task substantially, since actors will not have to make unknowns explicit in an active way. Thus, while this specific setting would allow automation, it would make interpretation of results too specific for this case and task. The aim is to assess a general use of the Explicit Unknowns functionality and therefore automation of the functionality was deliberately omitted to maintain this generality.



(a) Blank victim form.



(b) Explicit Unknown initiated on unknown field.

Figure 5: Victim forms with the Explicit Unknowns functionality.

(a) Icon without unknown information. (b) Icon indicating unknown information.

Figure 6: Datapoint icons of a victim report

Hardware setup Two separated experiment setups were used at the same time, each one consisting of a client side and a server side. The client side of each setup consisted of two desktop computers (Dell Optiplex GX270), each with two 17" TFT monitors (Dell 1704FPV), keyboard, mouse and gaming controller (Logitech Rumblepad II). The server side consisted of a desktop computer as a UDK, UsarSim and TREx server. Machines were connected by UTP cables through a gigabit ethernet switch.
Participants were placed opposed to each other and were able and allowed to communicate directly with each other but were unable to see the monitors of their team mate. Figure 7 shows one of the two setups used in the experiment. It clearly shows the opposing workplaces of the participants. For every client-side setup, the left monitor displayed the USARSim view, whilst

Figure 7: One of the two similar setups used.

the right monitor displayed the common operational picture.

3.3 Design

The experiment featured a between subjects design in which individuals were paired as a team. The experiment consisted of two conditions: the first condition made use of the common operational picture without the *Explicit Unknowns* functionality, the second condition was done with the *Explicit Unknowns* functionality available in the common operational picture.

A between subjects design was chosen because a within subjects design needs participants to be randomly assigned to both conditions. Doing so would cause *carryover effects* in which results are influenced by participation in a previous condition. This could, for instance, result in more experience with the task or a more developed search strategy for participants that took part in the condition **without** the Explicit Unknowns functionality before taking part in the condition **with** the functionality. More importantly, participants first taking part in the condition **with** the functionality could be biased by their knowledge about the functionality and mimic it in the condition **without** the functionality. Therefore participants had to be naive and knew nothing about the workings of the functionality in the condition where it was not available.

3.4 Task

Participants were instructed to explore a virtual environment using a robot and *completely register* as much victims as possible in the common operational picture within 15 minutes of time. Registering victims completely meant participants had to add a victim report in the common operational picture and complete all information fields of this form. Hence, victim reports with one or more empty fields were incomplete. Figure 4b shows the used form with its required information fields. Participants were told that found, but not completely registered victims, were regarded as *not rescuable*. Hence emphasis was on completeness of information instead of purely the quantity of found victims.

3.5 Participants

In this experiment, 32 unexperienced persons participated of which 16 were male and 16 were female. Participants were initially untrained and naive regarding the domain. From these 32 participants, 48 unique teams of two persons were formed. This resulted in 14 male-male teams, 20 male-female teams and 14 female-female teams. Table 1 shows the distribution of these combinations over the conditions.

Team	Condition 1	Condition 2
male-male	10	4
male-female	7	13
female-female	7	7

Table 1: The distribution of gender combinations amongst teams.

Participants were aged 18 to 51 years old *(mean = 28,72, mode = 22).* All participants had at least higher vocational education or a scientific education[5], except for one participant who had an intermediate vocational education[6]. Of all participants, only two had met each other before. Of 32 participants, 25 owned a driver's licence (78%) which they owned for an average of 11 years and drove on average 12.000 kilometers per year.
15 of 32 (47%) indicated that they use a personal computer or game console to play games, for an average of 5,85 hours per week. Six participants indicated that they play either first-person shooter games or racing games. Participants were asked to indicate their experience with remote controlled machines or toys. On a scale from 1 (no experience) to 5 (a lot of experience), participants averaged to 1,78.
In general it can be concluded that participants had some experience in gaming, and little or no experience with robots or other remotely controlled objects.

[5]Respectively *HBO* and *WO* in Dutch
[6]*MBO* in Dutch

3.6 Measurements

Participants were asked to fill in a demographic questionnaire before the experiment started (see appendix B). This questionnaire asked about age, gender, possession of a driver's license as well as experience with gaming and remote controlled objects.

Measurements were taken in two different ways; firstly data was gathered by individual questionnaires for every session, secondly data was obtained by the manipulations by the participants in the common operational picture. The questionnaires were used to obtain subjective quantitative measures on different topics. These questionnaires were 5-scale Likert-style questionnaires of which participants were asked to indicate how much they agreed with the statements in the questionnaire. After execution of the experiment, factor analysis was performed on the quantitative questionnaires (questionnaires in appendix C.1 to C.4) to check for variability between questions. Factor analysis is a statistical method to describe variability amongst observed, correlated variables. In terms of questionnaires, this means a large number of factors would indicate little semantic correlation between questions. Since the questionnaires in this research are grouped per topic, ideally one factor with a large eigenvalue should result from factor analysis. More factors could indicate less semantic coherence between questions, meaning more than one topic is asked. Besides factor analysis, Cronbach's alpha was calculated to check the consistency between questions for each topic. After this, an aggregated value per participant per topic resulted by calculating the mean of all answer values or calculating a score based on this mean. The following sections elaborate on the different subjects and the way a score was calculated for each topic.

Cooperation

Cooperation was measured subjectively by questionnaire C.1 as well as an objective score based on the data from the repository of the common operational picture.
For the first, subjective score for cooperation, table 6 shows the results of factor analysis and Cronbach's alpha whilst table 2 shows the questions in the questionnaire. Two factors were present after factor analysis, a strong factor with a high eigenvalue and a second factor with a fairly low eigenvalue. The small second factor resulting from this factor analysis can be accredited to question 8 in table 2, which is on trust. Although question 8 is more about trust than cooperation itself, this question was not omitted because the eigenvalue of the second factor is very small and trust is regarded as an important part of cooperation in a team.
The mean value represents an individual score for cooperation.

Question
1. The whole time, I cooperated well with my team mate.
2. The whole time, my team mate perfectly understood me when I shared information.
3. My team mate fully did what I expected of him/her.
4. I could use the information that my team mate added in the tactical display[3] very well.
5. I used a clear strategy for myself for searching the environment.
6. My team mate and I used a common and discussed strategy.
7. Our team prevented redundant work.
8. I trusted my team mate completely.
9. I was a useful addition to my team.

Table 2: Questions in the cooperation questionnaire.

The second, objective score for cooperation was calculated by the inverse of the number of rescued victims divided by the number of found *living* victims; see equation 1. The higher this score, the better team members provided required information when a living victim was found.

$$P_{team\_cooperation} = 1 - \frac{n_{rescued\_victims}}{n_{found\_living\_victims}} \tag{1}$$

Situation awareness

Questionnaire C.2 was used to get an subjective measure of situation awareness. The questions in this questionnaire were mainly based on the short post-assessment of situation awareness (SPASA) items by Gatsoulis et al. [11]. Specific questions were added and afterwards factor analysis was performed. Factor analysis resulted in two factors, of which the second was relatively small. This small factor is mainly caused by question 6 in table 3, which is not part of the SPASA items. Since the eigenvalue of this factor is relatively small it was decided not to omit this question. The remaining questions are shown in table 3, the results of factor analysis and Cronbach's alpha are shown in table 6. The mean value represents an individual score for situation awareness.

Question
1. I had a good overview of the situation.
2. It was easy to know where I was and which direction I was facing.
3. It was easy to identify and avoid obstacles.
4. It was easy to keep track of time aspects.
5. It was easy to see which part of the enviroment was searched.
6. It was easy to keep track of which victims were found.
7. It was easy to maintain the task of the experiment.
8. I always knew where my team mate was located.
9. If I changed something in the tactical display[3], I verbally informed my team mate.
10. I had a good overview of the information my team mate added.

Table 3: Questions in the situation awareness questionnaire.

Communication

Communication was measured by questionnaire C.3. The mean of the answers on the questions in table 4 was used as a subjective measure for individual communication. Results of factor analysis and Cronbach's alpha is again shown in table 6. The mean outcome of this questionnaire represents a generalized score regarding the relevance, verbality and correctness. However, this questionnaire was also split up in two parts to measure verbality. Questions 1, 2, 3 and 6 are regarded as questions related to verbal communication, whilst questions 4 and 5 are regarded related to non-verbal communication. Therefore these two separate parts were analysed combined for a general communication score, as well as individually for two indications of verbality. Factor analysis and Cronbach's alpha is shown in table 6.

Subjective performance

A measure of how participants felt they had performed, both individually and as a team, was taken by the subjective performance questionnaire C.4. The specific questions are shown in

Question
1. What **I said to my team mate** always was related to the task.
2. What **my team mate said to me** always was related to the task.
3. The (verbal) information my team mate provided was always correct.
4. I feel my team mate added an sufficient amount of information to the tactical display[3].
5. I feel I have added a sufficient amount of information to the tactical display[3].
6. If I changed something in the tactical display[3], I verbally informed my team mate.

Table 4: Questions in the communication questionnaire.

table 5, the analytic results of this questionnaire are shown in table 6.
The mean value represents a score for subjective performance.

Question
1. I feel I have performed well in the last session.
2. I feel my team mate has performed well the last session.
3. I have cooperated well with my team mate.
4. I understood my team mate well.
5. In a next session, I would try to do the communication in the same way.
6. In a next session, I would try to cooperate in the same way.
7. In a next session, I would use the tactical display[3] in exactly the same way.
8. I have added a sufficient amount of information to the tactical display[3].

Table 5: Questions in the subjective performance questionnaire.

Manipulation checks

Two manipulation check questionnaires were presented to the participants. These questionnaires did not provide any quantitative measures but served as qualitative feedback in attempt to explain potential irregularities. One of these questionnaires was presented after each session (see C.5) asking session-specific feedback, i.e. irregularities and intentions for the next session. The second questionnaire was presented after the experiment had finished, asking for feedback on all sessions, i.e. which session went the best and general feedback about the experiment and the explicit unknowns functionality (only for condition 2).

Topic	No. of factors	Eigenvalues	Cronbach's alpha
Cooperation	2	*(4,735; 1,010)*	,886
Situation awareness	2	*(5,162; 1,064)*	,888
Communication (general)	1	*(3,667)*	,853
Communication (verbal)	1	*(1,705)*	,827
Communication (non-verbal)	1	*(2,369)*	,726
Subjective performance	1	*(5,387)*	,926

Table 6: The results of factor analysis and Cronbach's alpha on the questionnaires.

Team performance

Besides individual subjective measures, team performance was determined by taking objective quantitative measures from the repository of the common operational picture. This team performance is an indication for the efficiency and efficacy of the team. A relative measure of performance of a team was calculated by dividing the absolute number of rescued victims[7] by the maximum number of rescuable victims, which is 12 for every session, and multiply by 100 to yield a percentage (see equation 2).

$$P_{team} = \frac{n_{rescued\_victims}}{n_{max\_rescuable\_victims}} \times 100 \qquad (2)$$

Strategy indication

An indication for the used strategy was obtained by taking the average relative distance between UGVs in the virtual environment during a session. If participants split up to search for victims during the experiment, the average relative distance between UGVs would be larger than when they stayed together. This was done by calculating the Euclidean distance between UGVs every time step and divide by the total amount of time steps. Equation 3 shows how this average was calculated.

After obtaining the average relative distance, *k-means clustering* was used to separate all teams into two groups. The cluster containing teams with a relatively large distance (final center value = 20,91) were classified as using a *split-up* strategy. Teams in the second cluster (final center value = 8,05) were classified as using a *stay-together* strategy.

$$\overline{D}_{rel} = \frac{\sum_{i=0}^{n} \sqrt{(x_i^{ugv1} - x_i^{ugv2})^2 + (y_i^{ugv1} - y_i^{ugv2})^2}}{n} \qquad (3)$$

4 Results

This section will show the results of the analysis of the data obtained by the evaluation described in section 3. First, results for the measures *cooperation, situation awareness, communication* and *subjective performance* will be presented. In general, results were analysed as an average over three sessions to find main effects, however in order to find possible interaction effects most results were also analysed per session. Finally, team performance and strategy will be addressed.

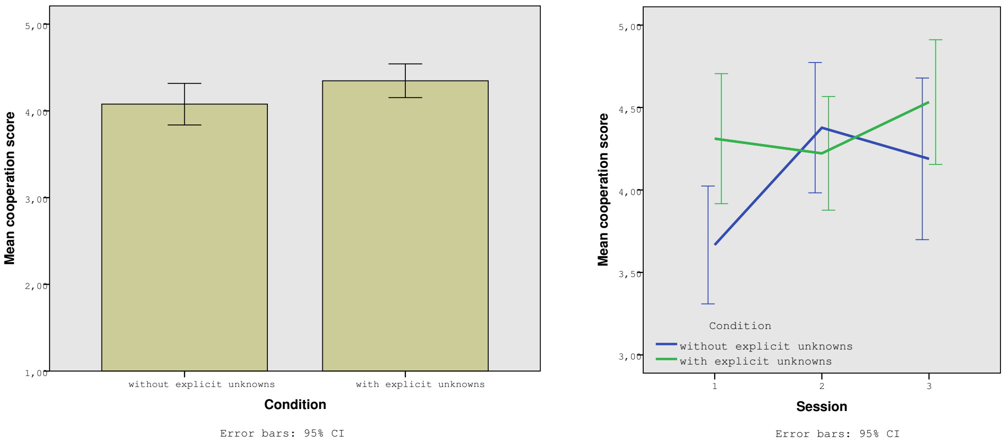
Due to technical problems, 17 of 48 participating teams were omitted from analysis, leaving a remainder of 31 teams. Of these, 15 teams participated in the condition without the Explicit Unknowns functionality, the other 16 teams participated in the condition with the functionality.

4.1 Cooperation

As stated in section 3, cooperation was determined in two seperate manners. The first is based on subjective questionnaire scores whilst the second is based on the data from the repository of the tactical awareness display.

[7] Rescued victims are completely registered living victims, see 3.4

Questionnaire data Figure 8a shows the mean scores of the cooperation questionnaire between the two conditions, over all sessions. A one-way ANOVA did not show a significant difference; *F(1, 60)=3,218, p=,078*. However, the graph and confidence clearly shows a tendency of increase in the score for the condition with the Explicit Unknowns functionality, which is as expected.



(a) Scores of combined sessions.

(b) Score course over sessions.

Figure 8: The mean scores of the cooperation questionnaires between conditions.

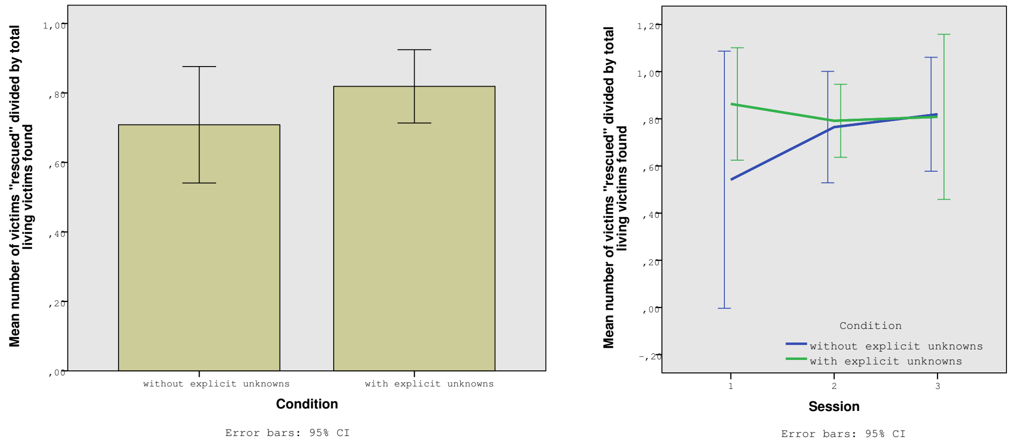
Besides analysis over all sessions, data of each separate session was analysed. Figure 10b shows the cooperation questionnaire score for both conditions per session.

A one-way ANOVA shows a significant increase in score for the condition with Explicit Unknowns in the first session, *F(1, 18)=7,501, p=,013*. The second and third session however show substantially smaller and non-significant differences, respectively *F(1, 20)=,441, p=,514* and *F(1, 18)=1,585, p=,224*.

Repository data The second measure to determine cooperation was a score based on the number of rescued victims divided by the number of found victims, as shown in equation 1. Since cooperation was essential when living victims were found, the result from this equation indicates the degree of cooperation for these living victims. If the number of victims rescued is less than the number of living victims found, apparently living victims remained without completed forms. This is attributed to a lack of cooperation and will result in a lower outcome of equation 1.

A one-way ANOVA did not show any significant results, *F(1, 29)=1,466, p=,236*, and is shown in figure 9a. Again, the tendency of the equation's outcome for the condition with Explicit Unknowns is consistent with the expectations.

Analysing the data per session yielded no significant results. Analysis is shown in figure 9b. The first session shows an increase for the condition with Explicit Unknowns, a one-way ANOVA for this first session resulted in *F(1, 8)=2,247, p=,172*. Note the large size of the error bars for this session, which are suspected to be caused by a combination of a small number of cases (4) which were also distributed diversely. One specific team did not rescue any victims, causing the wide distribution.



(a) Scores of combined sessions.

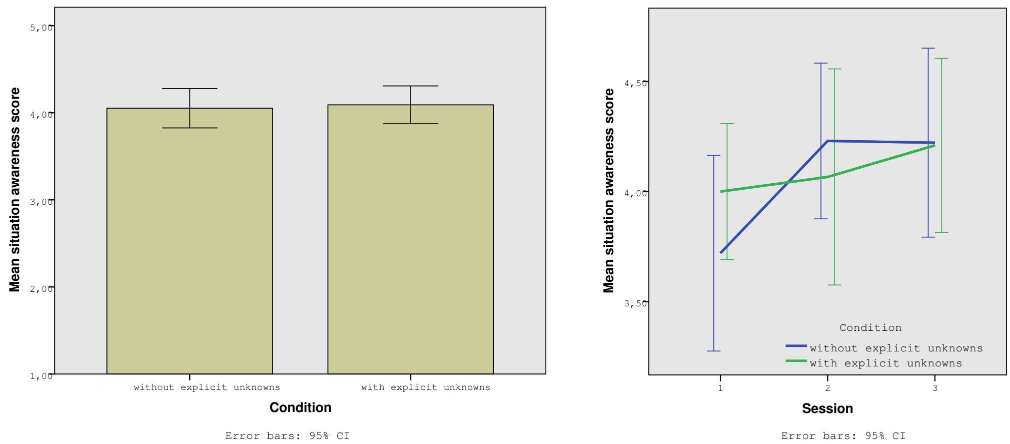
(b) Score course over sessions.

Figure 9: The mean scores of equation 1 per condition.

A one-way ANOVA on the second and third session, respectively *F(1, 9)=,069*, *p=,799* and *F(1, 8)=,005*, *p=,946* shows nearly any difference between conditions.

4.2 Situation awareness

Figure 10a shows the results of the situation awareness questionnaire scores. Almost no difference between conditions can be noted, which is also clear from the results of a one-way ANOVA, *F(1, 59)=,064*, *p=,801*.



(a) Scores of combined sessions.

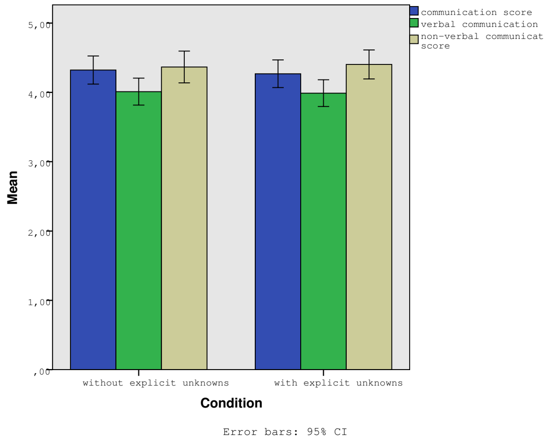
(b) Score course over sessions.

Figure 10: The mean scores of the situation awareness questionnaires between conditions.

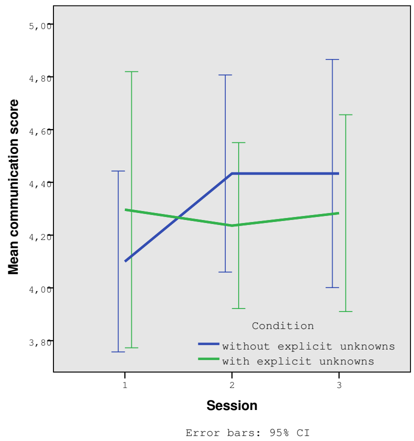
Analysing the data per session are shown in figure 10. Whilst one-way ANOVA results do not yield any significant results for the first, *F(1, 18)=1,369*, *p=,257*, second *F(1, 20)=,332*, *p=,571* and third session *F(1, 17)=,002*, *p=,962*. However, directions can be noted; a higher score for the condition with Explicit Unknowns in the first session, a large increase of scoring for the condition without Explicit Unknowns in the second session, exceeding the score of the condition with Explicit Unknowns and finally converging to almost the same score in the final session.

4.3 Communication

Communication was measured in a general sense but also specifically on verbality. For this, specific questions were analysed to obtain a seperate measure for *verbal* as well as *non-verbal* communication. Figure 11a shows all three of these measures. A one-way ANOVA results in no significant difference between conditions for overall communication (*F(1, 59)=,147*, *p=,703*), verbal communication (*F(1, 60)=,044*, *p=,834*) and non-verbal communication (*F(1, 60)=,135*, *p=,715*). These results do not provide any indication of the expected directions.



(a) Scores of combined sessions of communication.



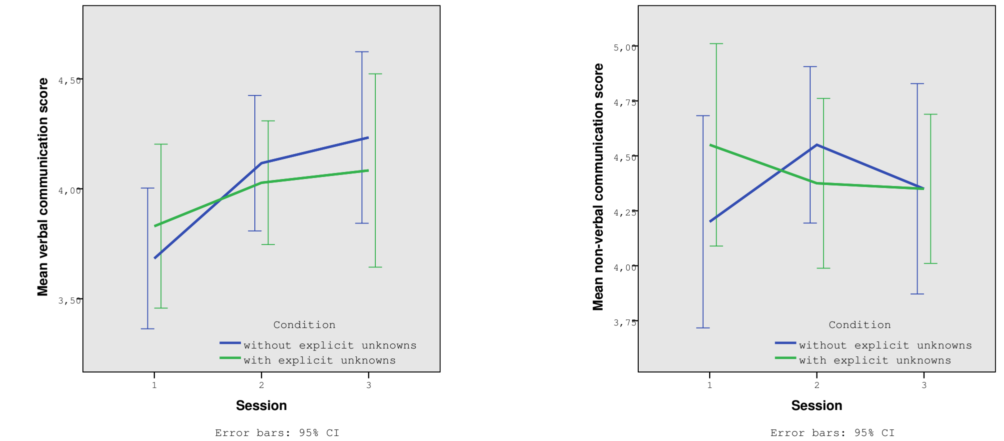
(b) Overall communication over sessions.

Figure 11: The mean scores of communication questionnaires.

Figure 11b shows the general communication scores over sessions. The one-way ANOVA for this measure did not result in any significant outcomes for any of the sessions. Results were *F(1, 17)=,537*, *p=,474* for the first session, *F(1, 20)=,824*, *p=,375* for the second session and for the last session, *F(1, 18)=,353*, *p=,560*. The graph does show a general direction; without Explicit Unknowns the first session scores low and passes the other condition to a seemingly stable score. The condition with Explicit Unknowns shows a fairly stable, almost linear course.

Figure 12a shows the scores of the verbal communication questions. Again, no significant results were observed. While the graph shows a slightly upward slope over sessions for both conditions, one-way ANOVA results show no difference between conditions (first session: *F(1, 18)=,456*, *p=,508*, second session: *F(1, 20)=,225*, *p=,640*, third session: *F(1, 18)=,333*, *p=,571*). However, while the course of the condition with Explicit Unknowns does not increase a lot over sessions, a larger increase for the condition without Explicit Unknowns can be observed after the first session.

Figure 12b shows the scores of the non-verbal communication questions, which yielded no significant results by a one-way ANOVA (first session: *F(1, 18)=1,409*, *p=,251*, second session: *F(1, 20)=,531*, *p=,474*, third session: *F(1, 18)=,000*, *p=1,000*). Remarkable is the convergence to *exactly* the same score in the last session for both conditions.

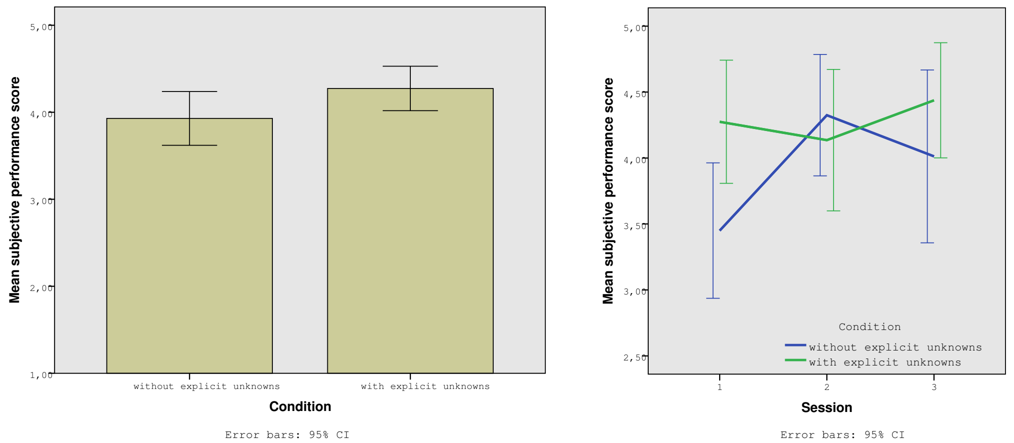


(a) Scores of verbal communication per session. (b) Scores of non-verbal communication per session.

Figure 12: The mean scores of verbal and non-verbal communication questionnaires.

4.4 Subjective performance

Figure 13a shows the results on the scores of the subjective performance questionnaires. The one-way ANOVA did not show a clear significant result, *F(1, 60)=3,115*, *p=,083*. Nevertheless, the increase in the Explicit Unknowns condition indicates the expected direction of the score between conditions.



(a) Scores of combined sessions. (b) Score course over sessions.

Figure 13: The mean scores of subjective performance questionnaires.

Results over sessions are shown in figure 13b.
The first session shows a significant increase in scoring for the condition with Explicit Unknowns, *F(1, 18)=7,233*, *p=,015*. The second session shows no significant difference, *F(1, 20)=,339*, *p=,567*, but does show a fairly large increase in score for the condition without Explicit Unknowns after the first session. The third session again does not yield a significant difference, *F(1, 18)=1,492*, *p=,238*, but just like the first session, indicates the direction in expected difference.

Combined questionnaire results To see how all questionnaire data changed over sessions it was visualized in figure 14. This figure shows the questionnaire scores of all topics, both conditions combined. It shows a clear increase of mean score as the sessions progress for all topics except non-verbal communication.

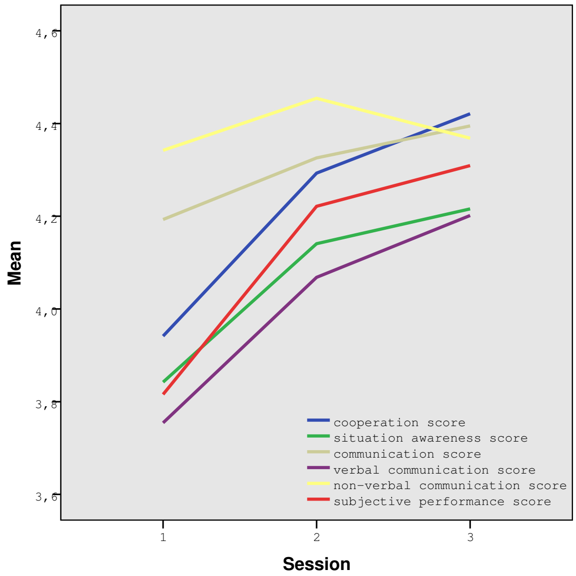


Figure 14: Average scores of all questionnaire topics per session.

4.5 Team performance

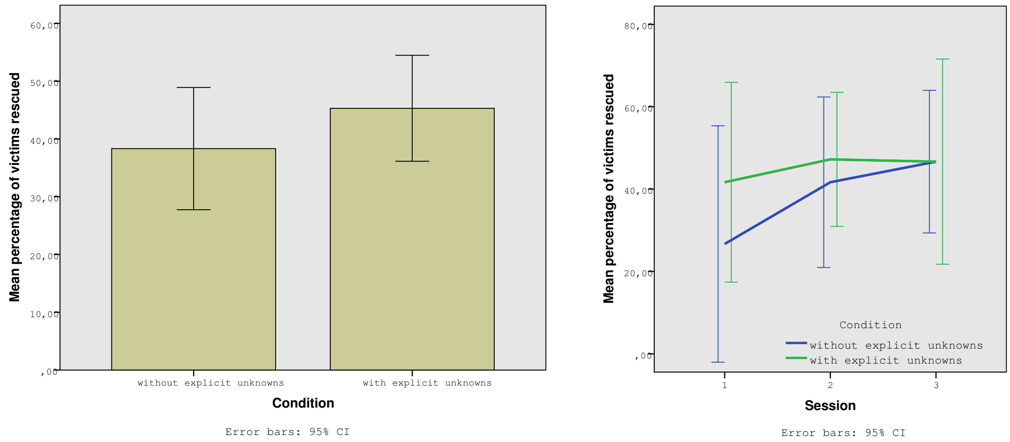
Team performance was, as stated in 3.6, measured by calculating the relative performance based on the number of rescued victims (see equation 2). Figure 15a shows the mean percentage of victims rescued. A one-way ANOVA shows no significant difference between conditions, *F(1, 29)=1,145*, *p=,293*. However, the direction of this team performance indicates a higher score for the condition with the Explicit Unknowns functionality, as expected.

Results per session are shown in figure 15b. One-way ANOVA results do not show a significant difference in any of the sessions. The first session, *F(1, 8)=1,227*, *p=,300*, does indicate the direction that was expected. This also is the case, be it in a lesser extent, in the second session *F(1, 9)=,327*, *p=,581*. The third session yields exactly the same score for both conditions, *F(1, 8)=,000*, *p=1,000*, and therefore cannot indicate anything on direction.

4.6 Strategy indication

An indication for the strategy used by participants was obtained by calculating the mean Euclidean distance between UGVs, as stated in equation 3. The results from this equation are shown in figure 16a. No significant results were obtained by a one-way ANOVA, *F(1, 29)=,073*, *p=,788*. Be it very small, the direction of the difference points in the way we expected.

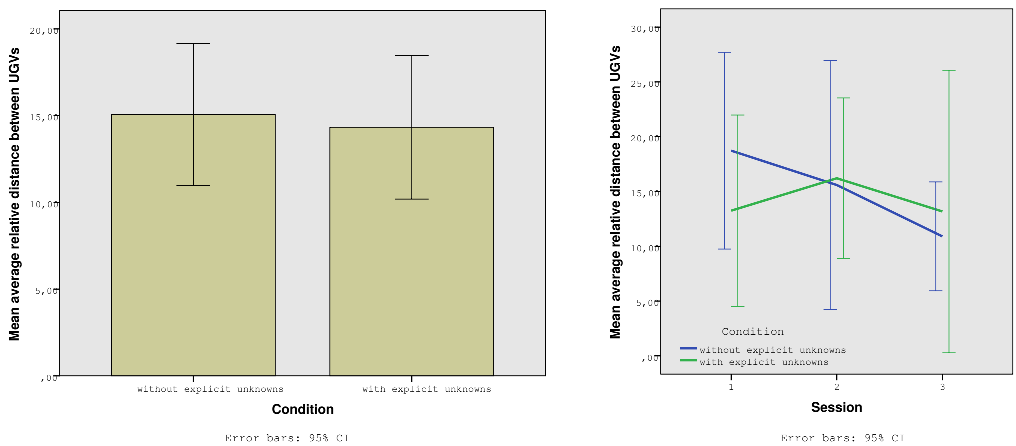
Since the search strategy could have been changed between sessions, it is important to analyse the relative distance per session. Figure 16b shows these results. Note the decline of relative distance over sessions for the condition without Explicit Unknowns, indicating a gradual shift



(a) Scores of combined sessions. (b) Score course over sessions.

Figure 15: The mean scores of team performance.

from a *split-up* strategy towards a *stay-together* strategy. One-way ANOVA results per session show no significance, but still the largest difference in the first session *F(1, 8)=1,476*, *p=,259*, where participants in the condition without Explicit Unknowns have a greater average relative distance. The second session shows nearly the same relative distance for both conditions, resulting in no significant difference after a one-way ANOVA, *F(1, 9)=,016*, *p=,902*. Average relative distance appears to drop for both conditions in the last session. Whilst the difference between conditions for this session is larger than in the second session, the one-way ANOVA yields no significant results, *F(1, 8)=,208*, *p=,660*. Overall, it may be observed that the average relative distance in the condition without Explicit Unknowns decreases after every session, whilst for the condition with Explicit Unknowns the distance is nearly the same between in the first and last session.



(a) Scores of combined sessions. (b) Score course over sessions.

Figure 16: The mean Euclidean distance between UGVs.

Analysis per strategy cluster (i.e. *stay-together* and *split-up*) shows remarkable results, shown in figure 17. It shows a significant difference between strategies used for the condition without Explicit Unknowns, while the condition with Explicit Unknowns remains nearly the same. The first shows a difference in team performance given the two strategies, for which a one-way ANOVA results in *F(1, 13)=5,102, p=,042*. The latter indicates no difference in team performance and the following one-way ANOVA results, *F(1, 14)=,014, p=,908*.

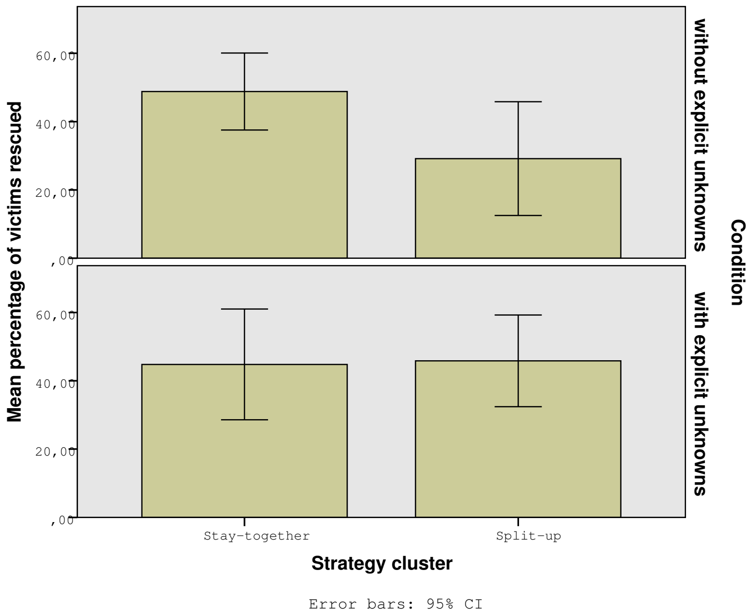


Figure 17: Team performance based on strategy, per condition.

5 Discussion

Cooperation The results of the questionnaires on cooperation show a tendency of a higher score in the condition where Explicit Unknowns were used. However, this cannot be regarded as an effect, since the two-tailed test does not yield proper significance (p=0.05 for a confidence level of 95%).

If we look at the cooperation scores per session, it can be observed that the condition without Explicit Unknowns initially has a very low score, then rising beyond the other condition in the second session and falling back in the last session. Only the first session yields a significant difference between conditions, while the other sessions do not differ much. When figure 10b is observed, participants in the condition without Explicit Unknowns seem to assess their cooperation initially low in the first session, but fairly high in the second session, as if there is compensation for their bad cooperation performance in the first session. This is also backed by the decrease in score in the last session for this condition, which is expected to be even higher than in the second session if it would be purely a learning effect.

Another explanation for the large initial difference between conditions might be that participants in the condition without Explicit Unknowns had to get used to sharing knowledge within the team. Since participants only received short training in using the tactical display and virtual

UGV without being trained in communicating information, this had to be empirically developed the first session. The Explicit Unknowns functionality facilitated in specific communication by notifying the other team member of required information, omitting the need for this specific communication. This can explain the initial difference and convergence to practically the same cooperation performance in subsequent sessions. It takes participants some time to develop adequate communication when they do not have access to the Explicit Unknowns functionality, giving rise to performance in the following session. Participants using the Explicit Unknowns functionality did not have a need of doing this resulting in a more *stable* overall (subjective) score.

The second measure for cooperation, based on the relative amount of rescued victims of total found victims, did not show any significant difference. Nevertheless, it shows a relatively large initial difference. Since this measure is objectively taken instead of subjectively, this cannot be attributed to subjective assessment of cooperation. Participants left more living victims without filling in the forms completely in the condition without Explicit Unknowns. However, the confidence level is very low in this session because of the large dispersion and little number of cases.

The convergence in the subsequent sessions shows no difference in cooperation between conditions. This could be explained in the same way as the initial higher score on the cooperation questionnaires; participants in the condition without Explicit Unknowns initially do not have an efficient way of communicating resulting in poorer cooperation performance. Again, in subsequent sessions they might have mastered this, resulting in nearly the same performance as participants in the condition with Explicit Unknowns.

It is expected that measuring cooperation by means of questionnaire will show a significant effect if the number of cases would be larger in both conditions. The relatively large initial dispersion in the second cooperation measure will also be reduced when a larger quantity of data is obtained. Furthermore, participants using the Explicit Unknowns functionality might experience more trust in themselves, their team mates and the software initially. Trust and confidence are therefore aspects to consider in future research. Also obtaining a measure of cooperation might be done in a different manner.

Situation awareness There is no main effect observed for the situation awareness measure. This may be explained by Endsley & Kiris [7], who state situation awareness even might decrease in situations where systems take over, or automate, a certain task. In this case, a system automates where attention should be directed to, taking this load off the user.

As Endsley states, several levels of SA exist. Automation might decrease SA on a lower level, raising the opportunity for increased SA on a higher level. The type of questions in a questionnaire might determine which level of SA is measured. Therefore, a different set of questions might yield different, more accurate results. Due to a limited time for background research on several topics, the choice for the SPASA questionnaire was made at a certain point. Instead of the SPASA questions, other measures exist. SAGAT is an objective measure that, although it is fairly intrusive during the simulated task, could have been used as a measure for SA [8]. Also, SART is a subjective way of measuring SA by means of getting a scaled response from an operator. Although this is fairly similar to SPASA, it might have been an alternative rendering different results. Subjective assessment however does not correlate with objective measurements of SA according to Endsley [8], however it still provides powerful information since a person's perceived quality of SA may be important in determining how a person will choose to act on that SA.

Communication No main effects can be observed for all three measures of communication. Overall communication was expected to decrease in the condition with Explicit Unknowns (see 1.4). Whilst results did not yield any significance, the direction over sessions shows an increase of communication score after the first session in the condition without Explicit Unknowns. The course of the condition with Explicit Unknowns barely changes. If these trends continue, it might indicate the expected direction. Nevertheless, more sessions are definitely necessary for this.

Communication was only measured by questionnaire. This is a very subjective way of measuring an activity that takes place very implicitly. Asking explicit questions about this by means of questionnaires might therefore cause unrealistic and unwanted results. Executing the experiment with team members isolated of each other enables the possibility to monitor communication in a different manner. In that case communication needs to be indirect and perhaps even asynchronous, for example by walkie-talkie or textual messaging. This in turn enables possibilities on analysis of frequency and content of communication.

Subjective performance While the results from the subjective performance questionnaires do not yield significance, they show a substantially higher score for the condition with Explicit Unknowns. The cause of this appears to be in the first session, where there is a significant difference between scores. In subsequent sessions, this significant difference disappears. This might either mean significant results might be achieved when more sessions are done with more teams to obtain more data, or participants will only initially feel they have performed better with Explicit Unknowns. The latter case would mean the influence of the first session decreases, resulting in less difference between subjective performance scores. Also, subjective performance score might be influenced by other scores, like cooperation, team performance or communication.

Interaction and learning effects Figure 14 shows the course of all questionnaire measures over sessions. It can be observed that all questionnaire scores, except non-verbal cooperation, increase in subsequent sessions. This might indicate either an interaction effect or a learning effect. Since participants only executed three sessions, they might have become more skilled in both robot operation, task execution, team communication and cooperation. This does not result in accurate and reliable data, therefore more sessions might converge data to a stable mean.

An interaction effect, meaning one independent variable is influenced by another independent variable, also might explain the increase of scores. When participants feel they cooperated well, this may influence their subjective performance assessment. Therefore, some independent variables might be influenced by other measures which in turn might have been influenced by experience, thus the learning effect.

Team performance Team performance was measured by calculating the relative amount of rescued victims. Whilst no significant main effect could be observed, results in figure 15a point towards the direction that was expected; a higher performance for the condition with Explicit Unknowns. If the performance over sessions is observed in figure 15b is observed, it can be noted that the largest difference is in the first session, decreasing in subsequent sessions converging to exactly the same performance in the last session. If this convergence would continue in subsequent sessions, team performance would eventually be the same for both conditions. More sessions should be performed to analyse this possible effect.

Strategy Participants used search strategies to cooperate and pursue optimal performance. Staying together or splitting up are both strategies that were used by participants. An indication of which strategy a team applied was obtained by calculating the average relative distance between UGVs and clustering this data into two clusters, one for the stay-together strategy and the other for the splitting-up strategy. Whilst no significant main effect was observed, over sessions it can be seen that the distance tends to decrease for the condition without explicit unknowns. This indicates a shift of strategy, from splitting up towards staying together. For the condition with Explicit Unknowns, the difference between average relative distance per session is smaller, indicating less change in strategy.

One important observation is the sudden emergence of the Explicit Unknowns functionality in the condition without Explicit Unknowns, which will be elaborated on in the next paragraph. The emergence of this functionality appeared to influence participants in choice of strategy, since the four teams with the largest average relative distance in the condition without Explicit Unknowns apparently used their own version of Explicit Unknowns. While these teams acted as if they were part of the condition with Explicit Unknowns, they were actually part of the condition without the functionality and therefore are regarded as such during analysis. This, however, might influence the results for strategy indication significantly.

For this reason it is important to observe figure 17, which shows team performance per strategy and condition. A significant difference was found for the condition with Explicit Unknowns between strategy clusters. This shows teams performed significantly better when they used the stay-together strategy instead of splitting up in the condition without Explicit Unknowns. No difference was observed in the condition with Explicit Unknowns, which might indicate strategy choice is of less importance in this condition.

Explicit Unknowns functionality During the experiment, the Explicit Unknowns functionality was only available in its basic form. No complex routing of information was available to users, all data was shown identically for all users. This is no issue for the abstracted setting used, with only two team members. However, the explicit unknowns were also shown to both users, which means the user making an unknown explicit saw its own explicit unknown. While this is positive in some sense (e.g. getting feedback from the system that the unknown is made explicit successfully), it can confuse users when a large amount of explicit unknowns exists. Simply distinguishing between one's own explicit unknowns and unknowns made explicit by other team members might solve this apparent problem. Moreover, explicit unknowns were displayed only by altering the datapoint icon. This can be overlooked by users, which is solved within the $(\mathrm{RM})^3$ framework by presenting unknowns in an appropriate manner and modality to a user.

Emergence of the Explicit Unknowns functionality Answers in the manipulation check questionnaires as well as data from the repositories showed the sudden emergence of an improvised version of the Explicit Unknowns functionality. Participants in the condition without Explicit Unknowns emergently developed the functionality by themselves, without having knowledge of the existence of this functionality in the other condition. These participants agreed upon a heuristic to add symbols used for remarks or obstacles at victims that needed cooperation. After cooperation, these symbols were removed to prevent redundancy. This shows the necessity of such a functionality in this task, however it does not make the functionality trivial. In this case, only two participants with fairly the same abilities formed a team. In practice, teams consist of a larger quantity of team members, each having their own expertise and abilities. This asks for a more complex system, like the $(\mathrm{RM})^3$ framework, to handle questions and messages. This in turn shows the need for the messages and unknowns to be machine readable.

The emergent functionality developed by some participants thus works for small scale situations, but lacks power and will be insufficient in larger, more realistic situations. Nevertheless, the emergence of the heuristic shows its necessity.

General Whilst great care was taken to influence participants as little as possible during instructions and training, participants might be biased in their subjective assessment of their peformance in the questionnaires. This bias might be caused by information in the instructions, causing participants in the condition without Explicit Unknowns to know that they were in the control group or participants in the condition with Explicit Unknowns that they had some sort of advantage. There is no real indication this was the case, but it should not be omitted as a probable cause.

Perhaps when more cases were available and more sessions done, more accurate results might have been obtained, yielding a clear effect on cooperation and subjective performance.

Technical issues In the experiment, two identical but independent setups were used simultaneously to be able to execute two independent sessions at one time. One of these setups was subjective to minor faults at arbitrary moments. While these faults were minor, they influenced the team's performance as well as subjective assessment by questionnaire beyond any doubt. For this reason it was decided to omit a substantial amount of data from analysis. This had its effect on the sample size during analysis, influencing both effect as well as significance.

Statistical analysis For parametric statistical analysis like ANOVA, a normal distribution of data is assumed. The data satisfied a normal distribution based on accepted guidelines, showing no excesses in skewness or kurtosis of the distribution. Since most of the hypotheses were directional (e.g. cooperation scores would be *higher* when using the Explicit Unknowns functionality), a one-tailed statistical test could have been conducted. Because there is a lot of scepticism about the use of one-tailed tests, the safe assumption of a normal distribution seems ciritical to justify a one-tailed test [12]. Because of the small sample size due to the technical problems, a Shapiro-Wilk test did not support a perfect normal distribution for most data. For this reason, the use of a two-tailed test provides a more liberal confidence interval, reducing the chance of falsely rejecting the null-hypothesis. Moreover, using a two-tailed test does not limit the analysis to one side of the distribution. If for any reason the opposite effect of the hypothesis was observed, it could be statistically supported with a two-tailed test, where a one-tailed test would not be able to do so.

Team formation Unique teams were created every session. However, the individuals in these teams were not unique and participated in multiple sessions. This renders analysis quite complex since teams were regarded as unique subjects, while the individuals in these teams participated multiple times in the same condition.

Another issue with this manner of team formation is the propagation of strategies through teams during sessions. One team might develop a certain strategy, apply and test this strategy, while another team does the same for a different strategy. In a subsequent session, participants were recombined to form a new team that decided on a (possibly earlier used) strategy. This way, not only the combination of participants varied, but also the way they acted as a team.

To avoid this, teams should be fixed and participants should not be recombined after each session. Fixed teams and more sessions will converge to stable results and search strategies per team, facilitating a more reliable analysis.

Structured information Because of the way the Explicit Unknowns functionality was implemented, structured information is always needed to be able to know which information is known or unknown. A frame of reference needs to be determined a priori to define necessary information. This can only be done by structuring the information, leaving little space for open remarks since these do not fit the bounded structure. The requirement of an information structure could make the functionality less deployable in practice, however methods might be developed to structure information based on their semantics, deducing necessary associated information.

Ceiling effect A ceiling effect might have occurred for the questionnaire data. The scores obtained in the first condition, without explicit unknowns, were fairly high on the scale (mean around 4 on the 5-point Likert scale). An overall more difficult task might cause participants to assess their performance lower in the condition without Explicit Unknowns, making the ceiling effect disappear.

Subjective feedback In order to identify any problems during the experiment, participants were asked for general feedback after every session as well as after the whole experiment. The majority (70%) of participants in the condition using Explicit Unknowns indicated they were positive about the functionality and thought they would have rescued less victims without the functionality. Some participants indicated they did not use the functionality in all sessions because they were able to communicate a lot or stayed together anyway during the task. The answers in these questionnaires indicated a positive effect of the Explicit Unknowns functionality on the user. Another indication is the need for such a functionality in a similar task when participants are not able to communicate directly. This may also mean the task needed more restrcictions to test the functionality itself properly. However, it might also indiciate the functionality is not nececessarily essential for any task, only when operators are restricted in some senses.

5.1 Future research

Future research might improve results in several ways. Firstly, it is advised to use a larger amount of teams, which should be fixed teams. Secondly, more sessions should be done with these fixed teams in order to prevent learning effects as much as possible. Both of the above will result in even better distributions of data, in turn making one-tailed testing possible for the directional hypotheses. Furthermore, an objective measure for cooperation should be used to determine the team's degree of cooperation. For situation awareness, other questions would render different results and for communication an objective manner of capturing an analysing communication between team members could be used. This also could involve seperation or isolation of team members to force specific ways of communication.

The task might be more restricted in terms of communication, which is also realistic when teams consist of more team members since the same resource (i.e. the maximum amount of communicable information on a channel) has to be divided over a larger amount of actors. Also, since some participants indicated they did not use the functionality all the time, several different tasks might be performed to see when such a functionality is most needed and desired by users.

Moreover, trust and confidence in own performance as well as in the system and other team members will be an important aspect in future research, as indicated by Wang et al [18]. When a system makes certain decisions for a user or takes over functions or tasks, it is important for users to trust the system in doing so. While the research in this thesis did not focus on trust and confidence in the system it is an aspect worth considering in future research.

Since automation of the Explicit Unknowns functionality could be possible for some tasks, effects of such automation could be researched. Automation by means of indicating incompleteness of forms could affect an operator's decisions and behavior. Operators might put less effort in trying to answer fields, since other team members are easily presented with the information request. This means automating the functionality might lower the social threshold to (anonymously) request support, which could either have a positive effect on collaboration or undesirably flood the system with explicit unknowns.

Finally, the Explicit Unknowns functionality has been improved in the mean time by making use of the $(RM)^3$ framework, which allows testing with larger teams due to complex routing of messages to users. This also means the users are not presented with their own *unknowns*, these unknowns are only made explicit to the appropriate user. The users also are notified in a more appropriate way rather than just being shown a modified icon. When the amount of data points and explicit unknowns increases, users will have a better overview of required information this way.

6 Conclusion

USAR situations are a practical case of complex situations in which people have to collaborate as a team in order to perform optimally. Situation awareness can be shared by retrieving and sharing specific information. Computerized systems are becoming ubiquitous for this collaborative information retrieval and the systems used for this should help users apply their valuable resources to the most important tasks. Being able to make explicit which information is required to retrieve by certain team members might influence performance of the team in several aspects.

In this experiment, participants formed USAR teams searching for victims by making use of robots. Participants in the first condition did not have access to the *Explicit Unknowns* functionality to make their team members aware of unknown information, while participants in the second condition did have access to this functionality. Six hypotheses were formed in order to test for effects between conditions.

The first hypothesis stated the rescuing of victims would be more efficient when the Explicit Unknowns functionality was available. No overall significant effects could be observed, hence there is no proof to support the hypothesis. However, there are indications significant effects could be observed when a larger sample size would be used. Still, convergence to indistinguishable results over sessions might occur in that case.

The second hypothesis expected cooperation to be more efficient when Explicit Unknowns were used. Only in the first of three sessions a significant improvement was observed for the condition with Explicit Unknowns. In the two sessions after this first session, the significant difference was no longer observable. Overall results do not confirm this hypothesis, however making unknowns explicit might provide an initial advantage in cooperation.

According to the third hypothesis, situation awareness should improve when using Explicit Unknowns. No significant results were observed to confirm this hypothesis. Scores for SA did not differ between conditions, but might be improved by altering the used measure in future work.

The fourth hypothesis stated communication would be less verbal when Explicit Unknowns were used. Communication scores did not differ significantly, therefore this hypothesis must be rejected. The manner of capturing and analysing communication between team members might influence communication itself as well as the results assessing this communication.

Teams would assess their performance as better when using Explicit Unknowns, according to the fifth hypothesis. Only in the first of three sessions a significant improvement of the condition with Explicit Unknowns over the condition without Explicit Unknowns could be observed. In the two subsequent sessions, no significant difference between conditions could be observed. Thus, the large difference between conditions in the first session decreases in subsequent sessions. Making unknowns explicit might make participants initially feel they have performed better, or subjective performance might be influenced by other independent variables. Overall results do not support the hypothesis.

The last hypothesis expected teams to use a different search strategy in the condition with Explicit Unknowns. No proof was found for this hypothesis, however a significant difference was found in team performance between strategies in the condition without Explicit Unknowns. Teams performed worse when they split up than when they stayed together. However, this does not confirm the hypothesis but shows strategy does influence performance when Explicit Unknowns are used.

Nevertheless, results show tendencies towards some of the expected effects. We have learned subjective measurements by questionnaire do not yield clear differences between conditions. Changing the specific task executed by the participants and using fixed teams could prevent ceiling effects. To assess cooperation, situation awareness and communication alternative, objective measurements could provide significant results. The benefit of a functionality of making unknowns explicit was also supported by subjective feedback by the participants, who indicated they would have performed less when they would not have had access to the functionality.

We thus conclude results of this experiment provide an indication of the improvement by making unknowns explicit on collaboration in information retrieval tasks. However, results suggest more research needs to be done in order to find significant effects. Cooperation, subjective performance and used strategy seem the most promising topics. Changes in experiment setup, using fixed teams and let these teams do more sessions probably all yield convergence for several scores and used strategies. Emergent improvisation by participants to unwittingly mimic the functionality and subjective feedback by participants shows the importance and usefulness of such a functionality.

References

[1] Usarsim project. `http://sourceforge.net/projects/usarsim/`.

[2] Herbert H Clark. *Using language*, volume 4. Cambridge University Press Cambridge, 1996.

[3] Tjerk de Greef. *ePartners for dynamic task allocation and coordination*. PhD thesis, TU Delft, 2012.

[4] Tjerk de Greef, A Oomes, and Mark Neerincx. Distilling support opportunities to improve urban search and rescue missions. *Human-Computer Interaction. Interacting in Various Application Domains*, 5613:703–712, 2009.

[5] J. van Diggelen, M. Grootjen, Ubink E.M., M. van Zomeren, and N.J.J.M. Smets. Content-based design and implementation of ambient intelligence applications, 2012.

[6] Mica R Endsley. Designing for situation awareness in complex systems. In *Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment*, 2001.

[7] Mica R Endsley and Esin O Kiris. The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(2):381–394, 1995.

[8] Mica R Endsley, Stephen J Selcon, Thomas D Hardiman, and Darryl G Croft. A comparative analysis of sagat and sart for evaluations of situation awareness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 42, pages 82–86. SAGE Publications, 1998.

[9] Colum Foley and Alan F Smeaton. Division of labour and sharing of knowledge for synchronous collaborative information retrieval. *Information processing & management*, 46(6):762–772, 2010.

[10] Epic Games. Unreal development kit. `http://www.unrealengine.com/udk/`.

[11] Yiannis Gatsoulis, Gurvinder S Virk, and Abbas A Dehghani-Sanij. On the measurement of situation awareness for effective human-robot interaction in teleoperated systems. *Journal of Cognitive Engineering and Decision Making*, 4(1):69–98, 2010.

[12] Dennis E Hinkle, William Wiersma, and Stephen G Jurs. *Applied statistics for the behavioral sciences*, volume 5. Houghton Mifflin Boston, 2003.

[13] Corine HG Horsch. Urban search and rescue robots, the influence of team membership of robots on team performance. Master's thesis, TU Eindhoven, 2012.

[14] Corine HG Horsch, Nanja JJM Smets, Mark A Neerincx, and Raymond H Cuijpers. Revealing unexpected effects of rescue robots' team-membership in a virtual environment. In *To be published in Proceedings of the 10th International ISCRAM Conference 2013*, 2013.

[15] Tina Mioch, Nanja JJM Smets, and Mark A Neerincx. Predicting performance and situation awareness of robot operators in complex situations by unit task tests. In *ACHI 2012, The Fifth International Conference on Advances in Computer-Human Interactions*, pages 241–246, 2012.

[16] Robin Roberson Murphy. Human-robot interaction in rescue robotics. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 34(2):138–153, 2004.

[17] Jennifer M Riley and Mica R Endsley. The hunt for situation awareness: Human-robot interaction in search and rescue. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 48, pages 693–697. SAGE Publications, 2004.

[18] Lu Wang, Greg A Jamieson, and Justin G Hollands. The effects of design features on users' trust in and reliance on a combat identification system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 55, pages 375–379. SAGE Publications, 2011.

[19] Ke-Thia Yao, Robert Neches, In-Young Ko, Ragy Eleish, and Sameer Abhinkar. Synchronous and asynchronous collaborative information space analysis tools. In *Parallel Processing, 1999. Proceedings. 1999 International Workshops on*, pages 74–79. IEEE, 1999.

A Questionnaires

The following pages contain the used questionnaires. The questionnaires used for the experiment were in Dutch. For the questionnaires that are used as measures (C.1 to C.5), in order to prevent any influence by translation, both the original questions are shown in Dutch as well as their English translation.

B Demographics

Participant number: ____

These questions are about personal characteristics that could influence performance during the experiment. The answers will only be used to analyse results.

Please fill in or check the answer that applies and provide a short elaboration, if asked.

1	Age:	______ years old.
2	Gender:	□Male □Female
3	Highest education program:	
4	Do you know one of the other participants?	□Yes, ______ □No
5	Do you own a driver's licence? (if not, proceed to question 8)	□Yes, ______ □No
6	For how many years do you own this licence?	______ years.
7	How many kilometers do you drive on average per year?	______ km.
8	Do you use a PC to play games?	□Yes □No
9	Do you use a game console (not PC) to play games? If yes, please specify.	□Yes, namely ______ □No
10	How often do you use the game console or PC to play games?	______ hours per week. □Does not apply.
11	What kind of games do you play the most? (e.g. FPS, race, strategy, ...)	□Does not apply.
12	How many experience do you have with the remote control of machines or toys? Mark between none (1) and very much experience (5).	(none) 1 2 3 4 5 (very much)

C Questionnaires after each session

We would like to ask you to fill in the following questionnaires after each session. These questionnaires are meant to retrieve information from every participant after each session about the experience of several aspects and their performance. The questionnaires propose several statements and an answering scale. You are asked to indicate how much you agree with the statements by circling the number that fits the best according to your feeling. The numbers in each questionnaire are scaled as follows:

Totally disagree	1
Disagree	2
Neutral	3
Agree	4
Totally agree	5

Because the statements are about personal experience and evaluation, there are no wrong answers. Important is that you indicate to what extent you feel the statements apply.

The questionnaires are about cooperation, knowledge of the environment, communication and performance.

Please answer honestly, there are no wrong answers.

C.1 Cooperation questionnaire

The following statements are about cooperation in the last session.
Please indicate how much you agree with the statements.

	Totally disagree	Disagree	Neutral	Agree	Totally agree
1. The whole time, I cooperated well with my team mate. *Ik heb de gehele tijd goed met mijn teamgenoot samengewerkt.*	1	2	3	4	5
2. The whole time, my team mate perfectly understood me when i shared information. *Mijn teamgenoot begreep mij de hele tijd perfect als ik informatie deelde.*	1	2	3	4	5
3. My team mate fully did what I expected of him/her. *Mijn teamgenoot deed volledig wat ik van hem/haar verwachtte.*	1	2	3	4	5
4. I could use the information that my team mate added in the tactical display very well. *Ik kon goed gebruik maken van de informatie die mijn teamgenoot toevoegde in de tactische display.*	1	2	3	4	5
5. I used a clear strategy for myself for searching the environment. *Ik gebruikte zelf een duidelijke strategie voor het doorzoeken van de ruimte.*	1	2	3	4	5
6. My team mate and I used a common and discussed strategy. *Mijn teamgenoot en ik hebben een strategie afgesproken en die gebruikt.*	1	2	3	4	5
7. Our team prevented redundant work. *Ons team voorkwam dubbel werk.*	1	2	3	4	5
8. I trusted my team mate completely. *Ik vertrouwde mijn teamgenoot volledig.*	1	2	3	4	5
9. I was a useful addition to my team. *Ik heb een nuttige bijdrage aan het team geleverd.*	1	2	3	4	5

C.2 Situation awareness questionnaire

The following statements are about the knowledge of the environment and situation awareness. Please indicate how much you agree with the statements.

	Totally disagree	Disagree	Neutral	Agree	Totally agree
1. I had a good overview of the situation. *Ik had een goed beeld van de complete situatie.*	1	2	3	4	5
2. It was easy to know where I was and which direction I was facing. *Het was eenvoudig om te weten waar ik was en welke kant ik op keek.*	1	2	3	4	5
3. It was easy to identify and avoid obstacles. *Het was eenvoudig om obstakels te identificeren en ontwijken.*	1	2	3	4	5
4. It was easy to keep track of time aspects. *Het was gemakkelijk om tijdsaspecten bij te houden.*	1	2	3	4	5
5. It was easy to see which part of the enviroment was searched. *Het was eenvoudig om te zien welk gedeelte van de ruimte al doorzocht was.*	1	2	3	4	5
6. It was easy to keep track of which victims were found. *Het was eenvoudig om bij te houden welke slachtoffers er gevonden waren.*	1	2	3	4	5
7. It was easy to maintain the task of the experiment. *Het was gemakkelijk om de taak van het experiment te volgen.*	1	2	3	4	5
8. I always knew where my team mate was located. *Ik wist altijd waar mijn teamgenoot zich bevond.*	1	2	3	4	5
9. If I changed something in the tactical display, I verbally informed my team mate. *Als ik wat veranderde in de tactische display zei ik dit verbaal tegen mijn teamgenoot.*	1	2	3	4	5
10. I had a good overview of the information my team mate added. *Ik had een goed overzicht van de informatie die mijn teamgenoot toevoegde.*	1	2	3	4	5

C.3 Communication questionnaire

The following questionnaire is about the communication, be it verbally or by using the tactical interface, that you had with your team mate.
Please indicate how much you agree with the statements.

	Totally disagree	Disagree	Neutral	Agree	Totally agree
1. I feel I have had a lot of conversation with my team mate, where I took the initiative. [a] *Ik heb voor mijn gevoel veel met mijn teamgenoot gepraat, waarbij ik het initiatief nam.*[a]	1	2	3	4	5
2. My team mate conversated with me where my team mate took the initiative. [a] *Mijn teamgenoot praatte veel met mij waarbij mijn teamgenoot het initiatief nam.*[a]	1	2	3	4	5
3. What I said to my team mate always was related to the task. *Wat ik tegen mijn teamgenoot zei had altijd te maken met de opdracht.*	1	2	3	4	5
4. What my team mate said to me always was related to the task. *Wat mijn teamgenoot tegen mij zei had altijd te maken met de opdracht.*	1	2	3	4	5
5. The (verbal) information my team mate provided was always correct. *Mijn teamgenoot gaf altijd (verbale) informatie die correct was.*	1	2	3	4	5
6. I feel my team mate added an sufficient amount of information to the tactical display. *Mijn teamgenoot voegde voor mijn gevoel voldoende informatie toe aan de tactische display.*	1	2	3	4	5
7. I feel I have added a sufficient amount of information to the tactical display. *Ik voegde voor mijn gevoel voldoende informatie toe aan de tactische display.*	1	2	3	4	5
8. If I changed something in the tactical display, I verbally informed my team mate. *Als ik wat veranderde in de tactische display zei ik dit verbaal tegen mijn teamgenoot.*	1	2	3	4	5

[a] Question removed for analysis to improve outcomes of factor analysis and Cronbach's alpha due to irrelevance to the intended subject.

C.4 Subjective performance questionnaire

The following questionnaire is about how you feel you have performed in the last session. Please indicate how much you agree with the statements.

	Totally disagree	Disagree	Neutral	Agree	Totally agree
1. I feel I have performed well in the last session. *Ik heb het gevoel dat ik deze sessie goed gepresteerd heb.*	1	2	3	4	5
2. I feel my team mate has performed well the last session. *Ik heb het gevoel dat mijn teamgenoot in deze sessie goed heeft gepresteerd.*	1	2	3	4	5
3. I feel I have found all the victims in the environment.[a] *Ik heb het gevoel dat ik alle slachtoffers in de ruimte heb gevonden.*[a]	1	2	3	4	5
4. I have cooperated well with my team mate. *Ik heb goed samengewerkt met mijn teamgenoot.*	1	2	3	4	5
5. I understood my team mate well. *Ik begreep mijn teamgenoot goed.*	1	2	3	4	5
6. In a next session, I would try to do the communication in the same way. *Ik zou de communicatie op dezelfde manier aanpakken in een eventuele volgende sessie.*	1	2	3	4	5
7. In a next session, I would try to cooperate in the same way. *Ik zou de samenwerking op dezelfdemanier proberen uit te voeren in een eventuele volgende sessie.*	1	2	3	4	5
8. In a next session, I would use the tactical display in exactly the same way. *Ik zou de tactische display op dezelfde manier gebruiken in een eventuele volgende sessie.*	1	2	3	4	5
9. I have added a sufficient amount of information to the tactical display. *Ik heb voldoende informatie toegevoegd in de tactische display.*	1	2	3	4	5

[a] Question removed for analysis to improve outcomes of factor analysis and Cronbach's alpha due to irrelevance to the intended subject.

C.5 Manipulation

The following questionnaire is about more general aspects of the last session.
Please try to answer short but clear.

1. How do you feel the last session went, and why? Did everything go well or do you feel certain things went wrong?	
2. In the last session, did you use a personal strategy for searching or cooperation of which your team mate was not aware?	If so, please explain this strategy.
3. In the last session, did you use a common strategy for searching or cooperation that you and your team mate agreed upon?	If so, please explain this strategy.
4. Do you plan to do anything different in case of a next session?	If so, please explain what you plan to do different.
5. Other remarks:	

D Questionnaire after all sessions

The following questions are general questions about *all* sessions including training. The questions are open questions, please try to to answer short but clearly.

1. Which session (1, 2 or 3) did you think went the best or the nicest? Please shortly try to explain why.	Session: ______ Because:
2. What did you think of the robot's controls? Was it easy or hard, and why?	I found it: Because:
3. If you have applied a strategy in one or multiple sessions, did the strategy change during sessions?	If yes, please describe the different strategies.
4. If you used a joint strategy in a session, for searching or cooperation, which strategy do you think was most successful?	If yes, I think the following strategy was the best:
5. What did you think of the tactical display? i.e. Did you miss certain things you needed? What could be better?	
6. What did you think of the *Explicit Unknowns* system?[1] Did it help you to see where help was needed? Do you think you would have performed less good if you did *not* have this functionality?	
6. Try to mention 3 things that you experienced as *negative* during the experiment. This can be anything general or specific for any session.	
7. Try to mention 3 things that you experienced as *positive* during the experiment. This can be anything general or specific for any session.	
8. Other remarks:	

[1]This question was only in the questionnaire for the second condition *with* the Explicit Unknowns functionality.