Distributed Collaborative Situation-Map Making for Disaster Response

Lucy Gunawan[a,*], Hani Alers[a], Willem-Paul Brinkman[a], Mark A. Neerincx[a,b]

[a] *Man-Machine Interaction Group, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*

[b] *TNO Human Factors, Dutch Institute of Applied Science, Soesterberg, The Netherlands*

ARTICLE INFO

ABSTRACT

A situation map that shows the overview of a disaster situation serves as a valuable tool for disaster response teams. It helps them to orientate their location and to make disaster response decisions. It is, however, a complicated task to rapidly generate a complete and comprehensive situation map of a disaster area, particularly due to the centralized organization of disaster management and the limited emergency services. In this study, we propose to let the affected population be utilised as an additional resource that can actively help to make such a situation map.

The aim of this study was to investigate the possibility of constructing a shared situation map using a collaborative distributed mechanism. By examining earlier research, a detailed list of potential problems is identified in the collaborative map-making process. These problems were then addressed in an experiment which evaluated a number of proposed solutions. The results showed that more collaboration channels led to a situation map of better quality, and that including confidence information for objects and events in the map helped the discussion process during the map-making.

1. Introduction

After a disaster, such as the 2011 tsunami in Japan and the 2010 earthquake in Haiti, the local infrastructure can be devastated. In such an event, the extent of the damage needs to be continuously analyzed in order to understand the conditions on the ground. This is necessary as a key component of naturalistic decision-making, where decisions and actions in a dynamic environment are based on the available information. The process of familiarizing, analyzing, understanding and keeping track of what is going on, is known as sensemaking, and the outcome of this action is situation awareness (Klein, Moon, & Hoffman, 2006).

A situation map is one form of situation awareness as the product of the sensemaking process. This map serves as a tool to see the overview of the disaster situation based on geo-spatial information at a particular time. It may contain information indicating whether an area is safe, a road network is passable, a place is destroyed, etc. It can also indicate the locations of evacuation shelters and emergency facilities. This information is needed so that emergency services and supplies can be mobilized.

Unfortunately, due to the lack of professional resources available to collect and assemble complete situation information in combination with a centralized model of disaster management, it is usually not trivial to construct this kind of map. Learning from the mass-casualty disasters in the past, the local emergency services were often overwhelmed due to the limited capacity and

resources of emergency facilities, which was the biggest problem faced immediately after these disasters struck.

In order to overcome this problem, one of the possibilities is to find potential resources to support emergency services in gathering situation information collectively and continuously. Research conducted by disaster sociologists suggests that the population affected in a disaster can be considered as a potential resource instead of helpless victims, since they are still capable human beings during disaster situations (Drabek & McEntire, 2003; Dynes, 1994; Quarantelli, 1999). Examining the numbers of natural disasters over the last thirty years, it can be seen that the affected people who are neither killed nor injured in the disaster are the majority of the group, accounting for about 90% of the affected population (Guha-Sapir, Hargitt, & Hoyois, 2004). Therefore, it can be argued that by supporting the affected population with simple technology, a collective effort may improve and expedite the process of constructing a high quality situation map compared to currently used mapping methods.

Despite the fact that the affected population can serve as potential collaborators during the mapping process, there is still little technological infrastructure developed to support this. Recently, there were several attempts to use micro blogging services such as Twitter in reporting events in disaster situations (Vieweg, Hughes, Starbird, & Palen, 2010). In a similar manner, an open source map project called OpenStreetMap was used extensively during the Haiti earthquake in 2010 (n.d., 2010). Nonetheless, little research has focused on how to provide the affected population with support to help emergency services, and

specifically to allow them to rapidly construct a collaborative situation map.

Therefore, the aim of this study is to investigate the potential of a collaborative distributed mechanism for making a shared situation map by non expert users (laymen) representing the affected population. It tries to explore important factors in collaborative map-making. This work focuses on including different modalities in the communication and collaboration process between the different actors in the disaster scenario, and it evaluates how they improve the results.

The paper starts with a brief coverage of the literature, taken the position that affected population in a disaster are capable individuals instead of helpless victims. This is followed by a description of how these new potential resources can be used actively during disaster response. Recent disaster events provide some showcases where new technologies such as microblogging and opensource-mapping are effectively used in assisting rescue efforts. The paper then briefly talks about related work (Gunawan, Oomes, Neerincx, et al., 2009) to give a background about the suggested hypotheses and the work methodology. This is followed by the study handled in this paper, which involved a controlled and detailed experiment carried out to evaluate the ideas gathered from the earlier research. The paper concludes with presenting its main findings stating that providing more additional collaboration channels led to a situation map of a better quality, and that the availability of confidence information for objects and events in the map can help the discussion process during map-making. Finally, lessons learned about collaborative map-making are listed.

2. Related Works

2.1. Public as a resources

The most common model of disaster management, derived from the military centralized command-and-control model (Neal & Phillips, 1995), assumes and treats the affected population as helpless victims without ability to help themselves (Dynes, 1994; Neal & Phillips, 1995), let alone help other human beings. However, disaster sociologists have shown that the opposite is actually true regarding the affected population. Studies over 50 years of human response to disaster situation discredits disaster myths, such as panic and anti-social behaviour (Quarantelli, 1986), and instead shows the cohesive and unified emergent phenomena, such as a calm and helpful behaviour, of individuals or groups during situations of collective stress (Drabek & McEntire, 2003; Quarantelli, 1999; Wenger, Quarantelli, & Dynes, 1986). As a result, Drabek & McEntire (2003) and Dynes (1994) suggest the possibility of expanding the command and control model with a participative effort, allowing for a decentralized and flexible structure that accommodates collaboration between professional actors and the public. This is the base of our study, the distributed mechanism, where the affected population can help the emergency services

during disaster response, especially in the process of situation-map making.

2.2. Collective effort

Through extensive ethnographic fieldwork with a number of different fire and rescue services in Sweden, Landgren (2007) suggested that the collective effort of sensemaking is fundamental for successful response work. This collective effort should focus on information-sharing in social interaction among the response-actors involved and the transparency of their actions in the ongoing work. He suggested that, by making use of digital traces of team member actions, it is possible to provide collaborative visibility of the work, so that it improves the capacity for sensemaking in time-critical ambiguous events. These factors should be carefully taken into consideration when designing a collaboration support system so that they are built into the way such a system is designed and used.

2.3. Collaboration technology in disaster response

There are already some efforts aiming at using new technology in disaster response. These kinds of efforts are continuously gaining more popularity and recognition. For example, a recent study by Vieweg, et al. (2010) investigated microblogging (Twitter) as a medium to harvest information during the Oklahoma Grassfire of 2009, and the Red River Floods of 2009 for the purpose of improving situation awareness. The study identified geo-spatial information and situational updates as two important features generated during emergencies that accounted for improvement in situation awareness. However in microblogging, the users are limited to describing the geo-information either by mentioning geographical information such as city and road or location-referencing. Thereby they use a prominent landmark as the reference base, since they are not capable of pinpointing their exact location. Hence, an additional intermediate step is needed to convert this information into data that can be pinpointed on the situation map.

During the recent Haiti 2010 earthquake (n.d., 2010), there were substantial efforts using OpenStreetMap (Goodchild, 2007), where there was a massive rapid mapping progress in a very short period of time. Haiti did not have a digital map before the disaster. However, within 48 hours of the earthquake, a complete map of Port-Au-Prince and Carrefour were completed. This was achieved by the collaboration of hundreds of mappers around the world using post-quake aerial imagery. The resulting digital map was used extensively for the disaster response in Haiti, by emergency services and humanitarian organizations for damage report, search and rescue missions, and transportation purposes. In this example, the collective effort shows a successful collaboration during crisis events. It might have been more useful if the affected people in the disaster area contributed to the collaboration effort of making the map.

Additionally, for supporting collaboration among the professional rescuers in the field, some research has been done for aiding Urban Search and Rescue (USAR)

missions, specifically to make the actions of distributed teams observable (de Greef, Oomes, & Neerincx, 2009). Moreover, some research efforts attempted to implement a visual communication system providing drawing tools and predefined sets of icons that support a free and natural way to sketch and describe crisis situations. The aim was to allow for fast interaction, as pictorial signs can be recognized more quickly than written words (Fitrianie, Yang, & Rothkrantz, 2008; Yang & Rothkrantz, 2007).

2.4. Field observations and preliminary experiment

Field observations collected at a disaster response exercise at the Rotterdam Rijnmond Safety Region in the Netherlands (Gunawan, Oomes, Neerincx, et al., 2009), which focused on the process of creating disaster situation map, identified three important issues. First was the use of improper modalities for specific tasks, in this case the use of verbal communication only to relay geo-spatial information across distributed team members. Since it is difficult to pinpoint an exact location using a verbal description, this often resulted in an inaccurate exchange of location information. Secondly, since the situation map was not shared across the distributed team members, errors committed as a result of the above-mentioned problem were not quickly detected. The third issue was that as the rescuers having to process many different information chains, occasionally they neglected to forward important information to the map plotter. This rendered the situation map not only inaccurate but also out of date.

The above mentioned observations showed that it may be useful to test the effectiveness of other forms of communication modalities (i.e., visually through a map) to improve the process of creating a situation map. In addition, sharing the map across collaborators may help improve the overall situation awareness. Therefore, one option is to alter the way the maps are currently created (by a single map plotter) into a collaborative system where multiple actors at different places can continuously check the accuracy of the map and make corrections when needed.

Earlier research (Gunawan, Oomes, Neerincx, et al., 2009) has already examined face-to-face collaboration and map making, highlighting potential benefits and pitfalls. It concluded that a better joint map can be created when collaborators cooperate and help each other by recheck the story, facts, and the certainty level of the events. On the other hand, joint maps can have a lower quality than the individually created maps when there is an unbalanced power or dominance relationship between the participants, e.g. if one actor is more dominant in the discussion, or if one of the actors has a more senior position. Other conclusions from the same work also mentioned that during the discussion, while making a collaborative map, participants tend to repeatedly express their confidence about objects and events they remembered from the scenario. However, there is a potential for occasional information-loss where collaborators, who may be quite confident about a particular event, give up their stance when the uncertainty of another collaborator dominates the discussion. This usually occurred when there was an unbalanced relationship between collaborators.

While examining the problem of dominance among participants in focus groups, earlier work (Carey, 1995) has already concluded that such a problem can be bypassed by explicitly putting all relevant information on the table before starting the collaboration. Hence, the implementation of a collaborative map making system should have a mechanism that explicitly states the confidence levels of the presented information in order to overcome the dominance factor in the collaborative process. In regard to information sharing modalities, the system should have a combination of different modalities which can be used to continuously refine the generated map. One possible approach is to allow each user to construct an individual map as a first stage of the map making process. The system then can share the created map with other collaborators so that it allows them to compare the presented information and come up with a better shared map. An additional stage of collaboration can then introduce the modality of voice discussion (currently the only used channel of information by professional rescuers on the field) to allow collaborators to resolve ambiguities in the information on the shared map.

Based on the above discussed work (Gunawan, Oomes, Neerincx, et al., 2009) three hypotheses were formulated:

(H1) Additional stages of increased collaboration channels improve the quality of a share situational map.

(H2) Explicitly indicating confidence information of objects and events shown on a map improves the quality of a share situational map.

(H3) Explicitly indicating confidence information of objects and events shown on a map supports the communication process.

3. Remote Collaboration Experiment

3.1. Experimental Methodology

With the three hypotheses defined, a detailed experiment was designed to examine their validity. In order to test the first hypothesis, the experiment involved different stages of collaboration adding different types of communication modalities. The tested additional stages in the experiment were: (1) no collaboration, i.e. individual map making, (2) updating maps after individual situational maps were exchanged, and (3) adding voice communication between collaborators to discuss and alter their maps. Testing the second and third hypotheses required two experimental conditions in which collaborators were able or unable to show explicitly their level of confidence on the map itself. The idea was that by making the confidence level information explicit, it would be possible to see whether this helped the participants during the map-making and discussion process by allowing them to focus on discussing objects that were more important or they were less certain about.

3.2. Preparations

3.2.1. Scenarios

In order for participants to be able to make a collaborative map with overlapping information on the same incident situation, two different scenarios were created. These two scenarios were an explosion in a gas station and a collapsed bridge due to collision. The scenarios were verified for their plausibility by a fire-fighter commandant. Each of the scenarios was divided into two parts: the accident unfolding and the rescue response. So at the end, four sets of stories were created.

3.2.2. The miniature world for incident setting

After creating the scenario, a disaster setting in which the scenario took place was constructed. Making incident scenes in the real world setting proved to be impractical, therefore, it was decided to make a miniature world to simulate the incidents using Playmobil toy sets, an approach already shown to be appropriate for such experiments (Gunawan, Oomes, Neerincx, et al., 2009). These toys were chosen due to their simplicity and flexibility, which made them practical for rapidly simulating real world environments.

The scenes were constructed in such a way that they could be viewed from two different angles representing the different vantage points of the two observers. Photos were taken from two specific locations, while the Playmobil world was adjusted as the accident storyline developed. The goal was to stimulate the exchange of information, requiring the participants to collaborate in order to figure out the complete scenario.

In the study, four sets of photo slideshows for each scenario were created, giving eight slideshows in total. For each scenario, the first two slideshows presented the story of the unfolding accidents (from two different angles) while the other two showed the rescue effort (again, from two different angles). Some of the images were later manipulated using Adobe Photoshop to add effects such as fire and smoke.

3.2.3. The magnetic board for map-making

To allow the participants to rapidly create the situation map, and in an effort to ensure that the map can be consistently translated into quantitative data, the participants were not asked to draw their recollection of the events. Instead, they were provided with sets of icons of the objects, actors, confidence levels, and a map of the environment. As shown in Figure 1, participants were able to use these icons to illustrate their recollections of the events on a top view map of the disaster area. Since the map was fixed on a magnetic board, it was also possible to edit the locations of icons after they were placed on the map. This also gave the participants the ability to quickly edit the map if they wanted to. The board was light and simple to handle making it easy to hold up right to face the camera, photograph the map, and share it with the other participant.

The confidence level information was implemented using star icons that participants could place next to events, actors, and vehicles on the map. The confidence level information was presented with a red star for 'low confidence', yellow for 'medium confidence', and green for 'high confidence'.

Figure 1. A participant placing the icons on the magnetic board.

3.3. Participants

This study involved 32 participants that were grouped into pairs, thus totally 16 pairs. The pairs were arranged in such a way that each team consisted of unacquainted partners, to simulate that they never worked together before as is characteristic during a disaster. There were 7 female and 25 male participants, who were 22 to 42 years old ($M = 28$, $SD = 4.26$) with undergraduate to post-graduates level of education. The participants had a wide variety of different nationalities, and were recruited from the Faculty of Electrical Engineering, Mathematics and Computer Science at the Delft University of Technology. They all had normal or corrected-to-normal eyesight. Only two out of 32 participants had special training or experience as rescuers.

The participant received a token gift as an incentive to take part in the experiment which took approximately 2 hours to complete. The available gifts they could choose from had a value of about €15.

3.4. Design

A two-way repeated-measures design was used, the within-subject factors were the stages of collaboration (no collaboration, shared map collaboration, shared map with voice communication collaboration) and the availability of confidence level information (without confidence and with confidence). The order of the scenarios and the availability of the confidence level information were counterbalanced. It would however have been confusing to show the rescue slideshows before the accident slideshows, this aspect of the experiment was therefore not counterbalanced. Similarly, the stages of collaboration always followed the same sequence: individual maps with no collaboration, shared map collaboration, then shared maps with voice communication collaboration.

3.5. Procedure

Each experiment was conducted with a pair of participants. First, the procedure of the experiment was explained to the participants after which they were guided to separate rooms. Each participant read and signed a consent form that explained how the results of the experiment would be used. After completing a colour blindness test the first of four sessions started (scenario A: accident, rescue; scenario B: accident, rescue). In each session, the participants went through the task of constructing a situation map in three different stages of collaboration (no collaboration, shared map collaboration, shared map with voice communication collaboration) explained in further detail in the Tasks section below. After finishing all four sessions, the participants filled in a final questionnaire giving their impressions of the experiment.

3.6. Tasks

In each session, a slideshow was shown to the participants depicting the events for one of the scenarios. Each slideshow consisted of 21 slides, and each slide was shown for 5 seconds. Each of the two participants saw a slideshow of the same events but from a different point of view. After viewing the slideshow, for the first no collaboration stage, participants were given the magnetic board with the top view map of the disaster area and were asked to reconstruct the events they just saw on the map.

In the 'with confidence level information' condition, the participants were asked to add their confidence level for all events, actors, and vehicles involved in the scenario. Participants could place these stars next to the icons they placed on the maps, as shown in Figure 2.
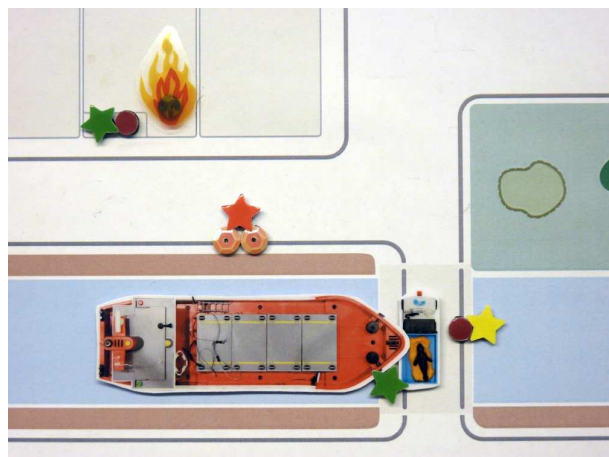

Figure 2. A participant's map with confidence level icons next to the objects in the map. This map shows, for example, that the participant was sure about the collision between the cargo boat and the small truck (green star), but has low confidence that there were two helping bystanders standing next to the water channel (red star).

A photo was then taken of the magnetic board and shown to the other participant, representing a shared map (shared map collaboration stage). The participants were given the chance to adjust their map if they felt necessary.

Finally, in the shared map with voice communication collaboration stage, the participants were again shown the map of the other participant and given the chance to have a voice discussion with the other participant for a maximum of 5 minutes. During or after the voice discussion the participants could adjust their map. A final photo was taken of the maps for evaluation purposes.

This was followed by three more sessions involving the rescue slideshow of that scenario, and the accident and rescue slideshows of the other scenario. In each of these sessions, the pair of participants went through the three above mentioned stages of collaboration. When starting to construct the map for the rescue session, the participants were given the choice to either modify the map they created for the accident slideshow or clear the map and start constructing a new one.

3.7. Measures

In order to examine the given hypotheses a set of measures was chosen. The first hypothesis involved the information sharing stages. To test this hypothesis, the quality of the map was measured after each stage of collaboration in order to see how it was affected. The quality of the maps was measured by comparing them to an ideal-map that contained all the events in the slideshows.

The second hypothesis involved the effect of explicit confidence information on the quality of the map. To test that, the quality of the produced maps was again used to see how it was affected by the availability of confidence information. To analyse the third hypothesis that focuses on the effect of explicit confidence information on the communication process, all voice conversations were recorded. They were examined for any effect on duration and behaviour in the discussion that was caused by the different conditions (with or without confidence information). Finally, the perceived usefulness of confidence information was collected by using a post-questionnaire at the end of the experiment.

3.8. Results

3.8.1. Data Preparations

For assessing the map quality, first an ideal-map was produced based on the ideal recreation of the events shown in the slideshows. The maps created by the pairs were evaluated by comparing them object by object to this ideal-map. Each object had two properties to be rated, namely detection (whether it was detected and placed on the map) and location (whether it was placed on the correct location). Each property received a score that could either be 0 (completely wrong), 0.5 (partially correct), and 1 (an exact match of the key-map). For example, an object on the map received a rating of 0 if the location did not correspond at all to the one in the ideal map, while a 0.5 rating was given when it was close to the correct location, indicating that the participant had an approximate idea regarding the location of the object.

Objects were then tagged into categories to facilitate further analysis of the data. For example, it was possible to analyse the quality of the mapped vehicles by looking at the score of all objects with the vehicle tag (police cars,

fire trucks, cars involved in the accident, etc.). The score for this category was calculated by taking the average score from all vehicles. When calculating the general quality of the entire map, the average score was taken of all categories of the objects on the map (there were 15 categories and 68 objects). This average score was a value ranging from 0 to 1.

In preparing the voice discussion data, a coding scheme tailored to the recordings was first developed. There were four sessions and 16 pairs of participants with a maximum of five minutes of discussion time, so totally around 320 minutes (around five hours) of discussion recordings. While listening to these recordings, the important keywords were noted down and were clustered to find the important phases and events in this specific discussion process. A phase is defined as a distinct period or stage in the discussion process that has a time duration. Only one phase can take place at the same time, meaning that one phase can only start once the other phase has ended. In other words, phases are mutually exclusive. An event is a single occurrence of a process that can take place within a phase. Events have no time duration. Events that took place during the discussion were grouped into five different types of events.

The six different phases were: (1) communication, (2) meta-communication, (3) my story, (4) your story, (5) bargaining, and (6) conclusion. The phase definitions, some utterance examples, the average discussion time (seconds), and the standard deviation of each phase can be seen in Table 2.

Definition and example of utterances	Duration(sec)	
	M(%)	*SD*
1 **Communication,** a phase where the participants greet each other, give compliments, and say goodbye. *"Hello, how are you?", "Can you hear me?", "Time is up, goodbye"*	35 (3 %)	36
2 **Meta-communication,** a phase where the pairs communicate on how they should communicate in this discussion, such as discussing their working procedure, suggesting procedure, and explaining what they are doing. *"So, how are we going to do this, shall we start by telling of what each of us saw, or shall we discuss the differences between our map?", "I'm looking at your map at the moment, so what I did, I changed the camping car to your location."*	99 (10 %)	58
3 **My Story,** a phase where a participants talk about their point of view of the story. *"I see that ...", "I have not seen that ..."*	670 (65 %)	201
4 **Your Story,** a phase where the participants talk about (what they think is) the other participant's point of view. *"I've seen that you put it on your map"*	144 (14 %)	69
5 **Bargaining**, a phase where the pairs discuss their findings, such as trying to convince the other participant and give suggestions of solutions. *"Can we remove that victim?", "Can you change the map then?", "We are now only focusing on the red car while there's more important things to*	26 (3 %)	27

discuss."

6 **Conclusion**, a phase when the pairs conclude and summarized the agreements. *"Ok, I think this way we are totally in agreement"*	39 (4 %)	33

Table 2. Six different phases in the discussion processes.

The five defined events were: (1) referencing the map, (2) certainty, (3) uncertainty, (4) agreement and (5) disagreements. The explanation of these events and some examples can be seen in Table 3. The table shows the Median instead of the Mean because the frequency of the events data was skewed, making the Median a better description of the central tendency of the data.

Events	Frequency	
	Mdn	*Range*
1 **Referencing the map**, an event that involves talking about the map itself. *"I saw it in your map that you put the fire truck", "Did you modify your map based on my map?", "Yes I modified it"*	0	0 - 6
2 **Certainty**, this event refers to the moments where any of the participants are talking about how confident they are of certain events. *"I'm sure about the fire truck", "I'm really sure about it"*	5	1 - 21
3 **Uncertainty**, this event refers to the moments where any of the participants are talking about how uncertain they are of events. *"I remember seeing it but don't know exactly where it is", "The thing is I can't be sure about that because I was standing in front of the building"*	13	5 - 21
4 **Agreement**, this event refers to the moments where the participants are in agreement. *"Okay, I will adjust my map then", "I put a victim on the map now, and we can agree about that one"*	0	0 - 4
5 **Disagreement**, this event refers to the moments where the participants are in disagreement. *"I don't completely agree about what you said"*	0	0 - 2

Table 3. Five events in the discussion process.

Using a custom built annotation program, a person who was not involved in the study was asked to annotate the 320 minutes of discussion recording with all the phases and events. To examine the reliability of the annotation, a second annotator rated a sample of 10 random recordings (out of 64 total recordings). The average case by case inter-rater Pearson correlation was 0.98 for phases, and the Spearman correlation was 0.60 for events. Additionally, the average phase by phase inter-rater Pearson correlation was 0.78, and 0.46 for the average event by event inter-rater Spearman correlation. The discussion duration and frequency for each session and scenario was then calculated. Durations were only calculated for phases and not the events. These were calculated by summing up the durations of all segments of the discussion spent on a specific phase. Furthermore, the duration data was logarithmically transformed, $\log_{10}(x+1)$ to decrease the effect of outliers and extreme values. The frequency (of the phases and the events) refers to the number of times they were initiated during the discussion.

To understand whether the confidence information was related to the frequency of mentioning the objects participants referred to during the discussions, the recordings in the sessions, which used 'confidence information', were further analysed. Depending on the combination of confidence information an object had from both participants, three different groups were defined: (1) both participants sure (green-green), (2) less sure (green-yellow, green-red, yellow-yellow, yellow-red, red-red), and (3) a confidence information was missing (green-missing, yellow-missing, red-missing, both missing). One missing referred to a situation where one of the participants forgot to put the confidence information in their map. The average frequency, by which an object in a specific category was mentioned during the discussion, was calculated.

To meet the independent sampling assumption, all analyses were done on a pair level. Therefore, all data, such as the map quality, the duration of the discussion, and the post questionnaires was averaged for each pair.

3.8.2. Statistical Analyses

The statistical analysis in this paper was divided into three parts. First, hypotheses H1 and H2 were tested by analysing the map quality. Second, the voice discussion was analysed to test hypothesis H3. Finally, hypothesis H3 was further tested by analysing the results of the post questionnaires focusing on the participants' experience.

Map quality

The quality of the map was analysed using a repeated-measures MANOVA. The independent variables were the availability of confidence level information and the stage of collaboration (no collaboration, shared-map collaboration, shared-map with voice communication collaboration), while the general map quality was the dependent variable. The results showed that the stage of collaboration had a main effect on the quality of the map with $F_{2, 14} = 57.13$, $p < 0.001$. This main effect was also found consistently in the analysis of the individual categories such as the victims, vehicles, etc., both on the accident map and the rescue map.
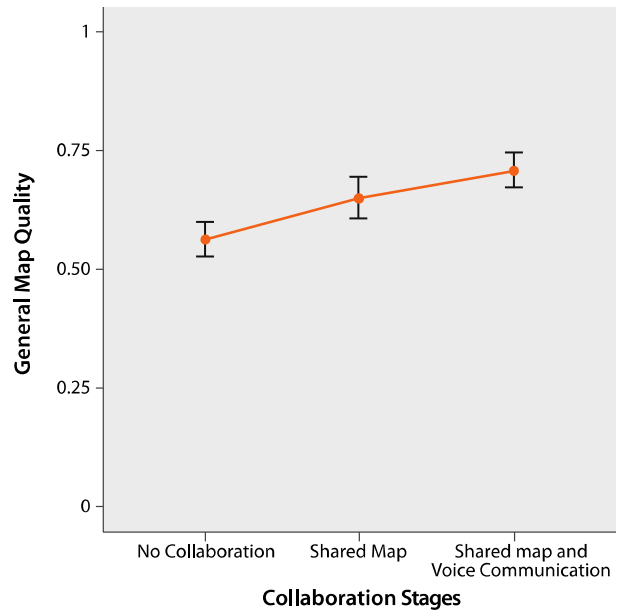

Figure 3. The mean map quality with 95% confidence intervals.

The effect for collaboration stages is illustrated in Figure 3, which shows that with more collaboration the quality of the map improved. A post hoc comparison analysis using a Sidak correction $\alpha_{PC} = 0.025$ showed that the indirect collaboration by sharing a map was better than no collaboration ($t_{15} = -6.08$, $p < 0.001$), and the collaboration of a shared-map together with voice communication was better than that with a shared map only ($t_{15} = -5.77$, $p < 0.001$). This therefore seems to support H1.

The analysis did not find a significant main effect for confidence level information availability $F_{2, 15} = 0.02$, $p = 0.884$ nor an interaction effect between collaboration and the availability of confidence level information $F_{2, 14} = 1.56$, $p = 0.244$. Therefore no support was found for H2.

Voice Discussion

To study hypothesis H3, the voice discussion data was analyzed in two ways: (1) the total duration for each of the phases, and (2) the frequency of the events.

To analyze the voice discussion duration, a repeated-measures MANOVA was used, with the type of phase (Communication, Meta-communication, My Story, Your Story, Bargaining, Conclusion) as an independent variable. The test showed a main significant effect $F_{5, 10} = 139.27$, $p < 0.001$ for the type of phase. Looking at Figure 4, the duration of the My Story phase seems to stand out from the rest of the discussion phases. Table 2 also shows that the My Story phase accounted for 65 % of the discussion time. Furthermore, the *t*-test comparisons (Table 4) among phases (Sidak correction $\alpha_{PC} = 0.003$) showed that it also significantly different from the other phases.
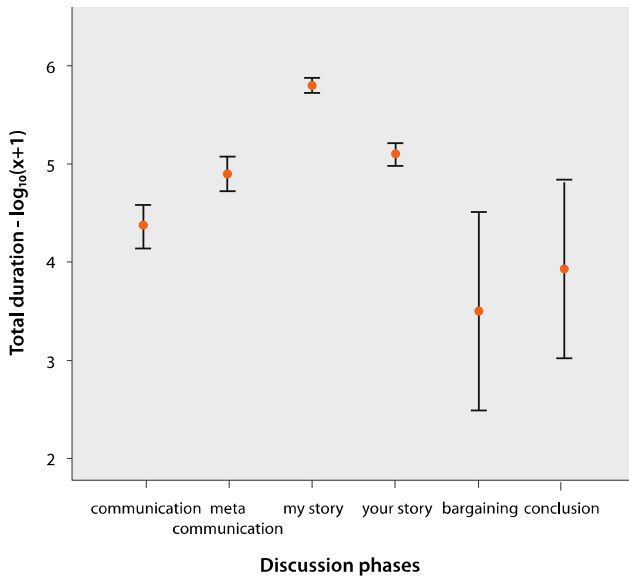
Figure 4. The total duration of discussion phases with 95% confidence intervals.

Phases comparison	df	t	p
Communication – Meta-communication	14	-3.79	0.002
Communication – My Story	14	-16.82	< 0.001
Communication – Your Story	14	-6.44	< 0.001
Communication – Bargaining	14	1.69	0.113
Communication – Conclusion	14	0.99	0.341
Meta-communication – My Story	14	-9.03	< 0.001
Meta-communication – Your Story	14	-1.93	0.074
Meta-communication – Bargaining	14	2.95	0.010
Meta-communication – Conclusion	14	2.15	0.049
My Story – Your Story	14	13.71	< 0.001
My Story – Bargaining	14	4.74	< 0.001
My Story – Conclusion	14	4.37	0.001
Your Story – Bargaining	14	3.29	0.005
Your Story – Conclusion	14	2.69	0.018
Bargaining – Conclusion	14	-1.40	0.184

Table 4. Phases duration comparison

Furthermore, a two-way repeated-measures MANOVA was conducted to analyze the effect of confidence level availability on the voice discussion duration. The session (accident and rescue session) and the availability of confidence level information were the independent variables. The two discussion phases (Bargaining, and Conclusion) were the dependent variables. The Bargaining and Conclusion phases were chosen since they were the phases during which the participants started to revise their maps. Although the test showed no significant effects (the availability of confidence level information main result $F_{2,13} = 2.94$, $p = 0.089$, interaction effect $F_{2,13} = 3.48$, $p = 0.062$), the result approached the significance level of $p = 0.05$. Furthermore, the univariate test for each phases revealed that only the availability of confidence level information had a significant main effect on the Conclusion phase of the discussion. The main effect

found, $F_{1,14} = 5.31$, $p = 0.037$, showed that the duration of the Conclusion phase in the accident session became shorter when the confidence level information was available, supporting therefore H3. Additionally the analysis also showed a significant two-way interaction effect between session and the availability of confidence information $F_{1,14} = 6.89$, $p = 0.02$, as shown in Figure 5. A detailed analysis of this interaction effect using a paired-sample *t*-test (Sidak correction $\alpha_{PC} = 0.0127$) showed that participants took less time in concluding when the confidence level information was available in the accident session ($t_{14} = 3.19$, $p = 0.007$), again supporting H3. However, this effect was not found in the rescue session ($t_{14} = -0.77$, $p = 0.455$). Furthermore, without the confidence information, the time duration spent on concluding the discussion showed no significant time decrease between the accident and rescue sessions ($t_{14} = 1.86$, $p = 0.085$). It also showed no significant increase in time when the confidence level information was available ($t_{14} = -2.62$, $p = 0.02$).
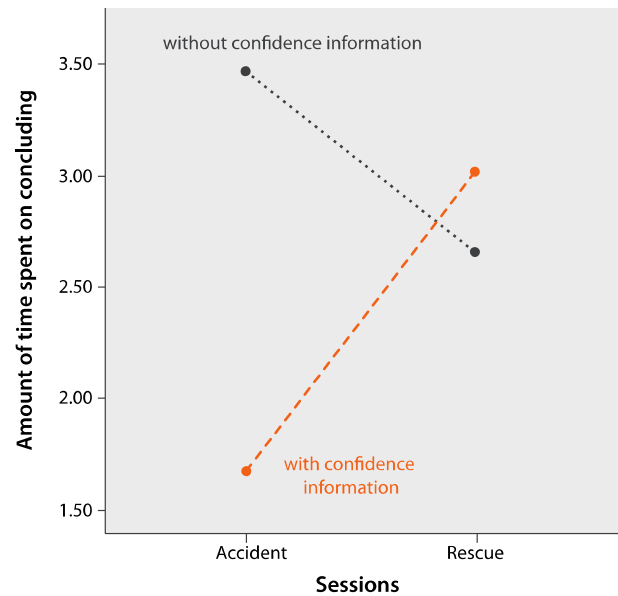

Figure 5. The effect of the availability of confidence level information on the duration of the conclusion discussion during the accident and rescue sessions.

Besides the phases, the effect of confidence level information on frequency of the certainty and uncertainty (Although the average of the event by event Spearman correlation was low, 0.46, these two events, certainty and uncertainty, have a high interrater correlation, 0.87 and 0.79 respectively) events (Table 3) were further analyzed using a Wilcoxon signed-rank test (Sidak correction $\alpha_{PC} = 0.025$). The uncertainty events frequency showed a tendency toward a significant increase ($Z = -2.142$, $p = 0.03$) as a result of the availability of the confidence level information ($Mdn = 8$) compared to when the confidence level information was not available ($Mdn = 4$). In other words, when the confidence information was explicitly shown, the uncertainty was mentioned more often. This may be because participants were more aware of their

uncertainties when it was represented on the map, which again supports H3.

The next step was to move on from an analysis between have and not having confidence information, to an analysis of what happened when confidence information was presented. The frequency when participants mentioned an object was analysed. The average frequency of mentioning an object in the discussion was compared using a paired sample *t*-test (Sidak correction α_{PC} = 0.025). The comparison between the frequencies of the group where both participants were sure about the object in the map (green-green) (*M* = 1, *SD* = 0.39), and when the participants was less sure about an object (green-yellow, green-red, yellow-yellow, yellow-red, red-red) (*M* = 3.87, *SD* = 3.69) showed a significant increase with t_{14} = -2.94, *p* = 0.011. Additionally, the comparison between total confidence (green-green) and when one participant missed the confidence level information (*M* = 2.93, *SD* = 1.74) also showed a significant increase with t_{14} = - 4.16, *p* = 0.001.

By exploring the voice discussions on a qualitative level, it was often noticed that: (1) when the participants were both sure (green-green) about an object or an event, they briefly pointed it out and then used it as a reference; (2) when one was less sure than the other, or when one completely missed an object, they had a longer discussion; and (3) when they both were not sure (red-red), they simply ignored those events. In some cases, it was possible through the discussion to recall missed objects from memory, but this occurred very rarely.

Here is an example of a conversation where both participants were sure about the events (green-green):

A: I think we are quite sure about two things
B: the fire
A: the location of the accident
B: yea
A: the car crash.

They later referred to these events to identify the timing of other events on the map:

B: did it happen after or before the fire start?

It therefore seems that the confidence level information sped up the Conclusion phase in the discussion process and made it more efficient by allowing the participants to only focus on the things that they were less sure of.

Post Questionnaires

At the end of the experiment, participants were asked to rate the perceived usefulness of the confidence level information during the process of collaborative map-making on a 7-point rating scale. A one-sample *t*-test with test value = 4 (we assume here that 4 is the neutral ground between positive and negative attitude) showed a significant (t_{15} = 2.93, *p* = 0.01) deviation from this middle rating. Looking at the means response, of 4.97, this suggested that participants leaned toward a positive

attitude with regard to this feature. From this, it seemed that participants on average were in favour of the confidence information. This again supports H3.

4. Limitations

One possible limitation of the study was the lack of a practice session, which may have had a negative influence on the performance of the participants in the first session (e.g. because of unfamiliarity with Playmobil forms and colours). In retrospect, it seems evident that a practice session would have helped reduce such an effect. On the other hand, by taking into consideration that each complete experiment took the pair of participants approximately two hours to complete, it would have been difficult to add extra components to the experimental setup.

5. Discussion

Although it was found that the confidence level information affected the discussion process and the participants were in favour of this feature (H3), the confidence level information was not found to affect the quality of the map (H2). This might be caused by confusion over how to use this confidence level information in the map-making process. In fact, it can be interpreted in two different ways, whether it was confidence about the type of the object, or the confidence level information about the location of the object. It is also possible that participants liked the confidence level information because it helped the discussion process run smoothly.

Additionally, the effect of the confidence level information was only found in a small parts of the discussion process (Bargaining and Conclusion phases). Therefore, the confidence level might not have a major impact on the duration of the entire discussion process.

6. Conclusion

This study showed that during collaborative map-making, additional stage of collaboration can improve the quality of the map (H1). It is useful to enable indirect collaboration of sharing a map made from different viewpoints since it improved the quality of the map. Next, this quality can be further improved by bridging the communication between the collaborators in addition to the shared map using voice communication. No support was found that providing confidence information lead to an improved map (H2). However, supporting the collaboration by providing confidence level information can shorten the conclusion phase of the discussion process (H3). Additionally, during the discussion, uncertainties are more often expressed when the confidence level information is available. This shows that expressing confidence level information explicitly coupled with events and objects can help to make the discussion be more efficient. Finally, also the confidence level information is perceived as useful by the users.

6.1. Contributions

The scientific contributions of this paper lie in understanding the collaboration process of situation map making. It shows that increasing the number of communication channels, by first making isolated maps, then sharing the maps, and finally sharing the maps together with voice communication, improves the quality of the produced map. Therefore, the suggestion is made to open more communication channels (especially visually shared maps) during collaborative map-making to complement the voice communication channel. This is because relying completely on voice communication to relay geo-spatial information has been reported in the literature to be inefficient and ineffective. Additionally, this study also shows how the confidence level information affects the efficiency of the discussion process. It does so by shortening the Conclusion phase of the discussion and by helping collaborators to better articulate their uncertainties since it is explicitly represented on the map.

6.2. Future works

As the collaboration stages and the confidence level information can enhance the process of situation-map making, both ideas can be implemented as a technological solution, especially with similar domain and usage. In addition, future research can integrate this study with some navigation technology in the case of a disaster where the affected population needs to be guided in a safe way to a designated destination. Considering that the participants spent 65 percent of the discussion time talking about their point of view of the story (namely the My Story phase), further improvements should address supporting this phase as here it seems most can be gained.

In regard to the novel use of toy sets, Playmobil, as quick prototyping tools for depicting disaster scenarios, served the purpose of the mentioned studies. Surprisingly, consultations with a fire-fighter commandant revealed that Playmobil was also used to train the fire-fighters during their exercises, where the toys were used to model disaster situations that fire-fighter trainees use to understand the scenario and plan their actions. It is therefore recommended to use this kind of method to simulate large-scale real life situations for similar future research.

This paper demonstrated that a distributed collaborative map-making mechanism can serve as a method to generate a situation map, thus in a disaster situations it might lead to a better situation awareness. This awareness is necessary to support the decision making process and help countless victims in disaster situations.

7. Acknowledgments

8. References

Allport, G. W., & Postman, L. (1946). An Analysis of Rumor. *The Public Opinion Quarterly, 10*(4), 501--517.

Buckner, H. T. (1965). A Theory of Rumor Transmission. *The Public Opinion Quarterly, 29*(1), 54--70.

Carey M.A., (1995) Comment: concerns in the analysis of focus group data. Qualitative Health Research 5, 487--495.

de Greef, T., Oomes, A. H. J., & Neerincx, M. A. (2009). *Distilling Support Opportunities to Improve Urban Search and Rescue Missions*. Paper presented at the Proceedings of the 13th International Conference on Human-Computer Interaction. Part IV. Retrieved from http://dx.doi.org/10.1007/978-3-642-02583-9_76

Drabek, T. E., & McEntire, D. A. (2003). Emergent phenomena and the sociology of disaster: lessons, trends and opportunities from the research literature. *Disaster Prevention and Management, 12*(2), 97--112.

Dynes, R. R. (1994). Community Emergency Planning: False Assumptions and Inappropriate Analogies. *International Journal of Mass Emergencies and Disasters*(275).

Fitrianie, S., Yang, Z., & Rothkrantz, L. J. M. (2008). *Developing concept-based user interface using icons for reporting observations*. Paper presented at the Proceeding of ISCRAM 2008.

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal, 69*(4), 211--221.

Guha-Sapir, D., Hargitt, D., & Hoyois, P. (2004). *Thirty years of natural disasters 1974-2003: the numbers*: Centre for Research on the Epidemiology of Disasters (CRED). Louvain-La-Neuve: UCL Presses, Universitaires de Louvain.

Gunawan, L. T., Alers, H., Brinkman, W.-P., & Neerincx, M. (2010). *Effect of map sharing and confidence information in situation-map making*. Paper presented at the Proceeding of European Conference on Cognitive Ergonomics 2010.

Gunawan, L. T., Oomes, A. H. J., Neerincx, M., Brinkman, W. P., & Alers, H. (2009). Collaborative Situational Mapping during Emergency Response, *Proceeding of European Conference on Cognitive Ergonomics 2009*.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making Sense of Sensemaking 2: A Macrocognitive Model. *IEEE Intelligent Systems, 21*(5), 88--92.

Landgren, J. (2007). *Designing Information Technology for Emergency Response*. IT-University in Göteborg.

n.d. (Artist). (2010). *Haiti Project In OpenStreetMap Wiki*.

Neal, D. M., & Phillips, B. D. (1995). Effective Emergency Management: Reconsidering the Bureaucratic Approach. *Disasters, 19*(4), 327--337.

Quarantelli, E. L. (1986). Research Findings on Organizational Behavior in Disasters and Their Applicability in Developing Countries. *DRC Preliminary Papers*.

Quarantelli, E. L. (1999). Disaster Related Social Behavior: Summary of 50 Years of Research Findings. *DRC Preliminary Papers*.

Vieweg, S., Hughes, A., Starbird, K., & Palen, L. (2010). Microblogging During Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness., *Proceeding of Computer Human Interaction*.

Wenger, D. E., Quarantelli, E. L., & Dynes, R. R. (1986). Disaster Analysis: Emergency Management Offices And Arrangements. *DRC Final Project Reports*.

Yang, Z., & Rothkrantz, L. J. M. (2007). *Dynamic Scripting in crisis environments*. Paper presented at the Proceeding of HCII 2007, LNCS.