Finding the best devices for emergency responders in Norway – an empirical study

Linda Katrine Andresen
University of Oslo
lk.andresen@creuna.no

Erik G. Nilsson
SINTEF ICT
Erik.G.Nilsson@sintef.no

ABSTRACT

In this paper we present the results from four empirical studies where we investigate the most appropriate devices for high-ranking emergency leaders – termed *commanders* in this paper – working close to the scene of incident, usually outside. The studies apply different methods including interviews, questionnaires and usability tests to investigate the technology used by commanders in the agencies today, their motivation for using new technology, as well as their needs for new devices. Three of the studies involve commanders from the police, the ambulance service and the fire and rescue agency in Norway. Devices with different screen sizes ranging from 4,3" to 40" were used or discussed in the studies. The main conclusion from all the studies is a very clear preference for tablets. We did however identify difference between the agencies regarding the preferred size of an "ideal" tablet, as well as wishes for using more than one device.

Keywords

Emergency response, Norway, device types, screen sizes, local commanders.

INTRODUCTION

Even though technology, and especially mobile devices like smartphones and tablets, is increasingly being used within other work domains as well as in our everyday lives, the implementation of new technology within emergency response (ER) in Norway has been comparatively slow. This is especially the case for support tools for emergency response leaders working at or near the scene of incident. In this paper, we focus on the needs of these leaders, who we refer to as commanders, which are high-ranking leaders working outside (or in a van or tent) and in charge of ER operations. Commanders are central in the cooperation and coordination between the involved agencies (primarily the police, the fire and rescue agency and the ambulance services), as well as within their own agency. These leaders therefore spend a lot of time communicating over radios and cooperating with other leaders at a local command post (LCP). ICT-based support tools for these users should include a combination of presenting information from sensors and other systems, and easy to use functionality for entering key information about the situation at hand, both contributing to a common operational picture.

Today there exists a wide range of devices with different sized screens, from mobile phones, via tablets and laptops, to table solutions with 40" screens or larger. Thus, there are several screen sizes available which can be exploited by commanders in emergency response agencies. Initiatives for developing support tools for these users tend to choose device types to use rather arbitrarily. The goal of the research presented in this paper (Andresen, 2013) has been to try to find out which type of devices – and implicitly how large screens – our target users want, and through this identify the "ideal" type of equipment/screen size for commanders.

RELATED WORK

Even though we found no literature exploring which screen size(s) emergency commanders need, different devices, ranging from mobile phones to large table solutions, are included in some literature regarding ER. When it comes to mobile devices, like mobile phones and tablets, some papers even include more than one type of device (e.g. Kim, Jang, Mellema, Ebert and Collins, 2007; Luyten, Winters, Coninx, Naudts and Moerman, 2006; Smirnov, Levashova, Shilov and Kashevnik, 2009). A common denominator for these papers, however, is

that the devices are not compared. Luyten et al. (2006), for example, include both a tablet and a mobile phone in their study regarding fire fighting, but the devices are not compared as it was assumed that they would be operated by different user groups. Most studies where mobile devices such as mobile phones and tablets are included, however, usually include just one type. These studies have little focus on screen size and devices, as they mainly concern technical solutions related to for example ad-hoc networks and sensors. Jiang, Chen, Hong, Wang, Takayama and Landay (2004) used a PDA in their study regarding visualizing sensor-information from an ad-hoc network for fire fighters, but neither the device used nor the screen size is really addressed.

The users addressed here, namely commanders, are to a higher degree addressed in studies where larger screens, such as Microsoft PixelSense, Smartboard and DiamondTouchTable, are included. These devices are not mobile, but have the advantage of having a large screen or surface which often can be used by several actors at the same time. Scotta, Pleizier and Scholten (2006) for instance investigated how the use of a tactical user interfaces on a large multi-touch table could enhance cooperation and coordination during an ER, and points amongst other things to the importance of providing spatial information such as maps and photos. The need for geo-information such as maps when cooperating on large screens in ER is also identified by others, e.g. by Busher and Mogensen (2007), Dubé, Kramer, Vachon and Tremblay (2011), Neuvel and Zlatanova (2006), and Rauschert, Agrawal, Pyush and Sharma (2002). Rauschert et al. (2002) to a higher degree than Scotta et al. (2006) address this need by exploring the use of multimodality, such as speech and gestures, when interacting with large screens displaying geospatial information.

Different devices and different sized screens are therefore included in some papers within the domain of ER. A common characteristic for these papers, however, is that the devices included in the studies are selected beforehand. There are generally few studies where *needs* related to device and screen size within ER are addressed. An exception is Way (2009), who provides a framework for finding the appropriate device for crisis responders, even though he does not test devices himself. The framework is furthermore constrained to mobile devices and is based on an expert finding the appropriate device on behalf of the user instead of exploring screen sizes with the intended end-users. Research where different devices are compared can nevertheless be found within other domains such as health care and architecture, but the differences between these domains and ER make the transfer of the results difficult. Some of these papers do however indicate that it is not given beforehand which screen size is the most suitable for the addressed user group, implicating the importance of investigating screen sizes with the intended end-users.

METHOD

In this research we have applied a triangulation approach (Taylor, Kermode and Roberts, 2006) and conducted four studies to investigate which type(s) of devices that could fit the needs of ER commanders from different angles. The studies build on results from a number of empirical studies of emergency response work (Nilsson and Stølen, 2011). Below we present the methods used and the set-up of each of these four studies, as well as the devices used in Study 3 and 4.

Study 1: Interviews with emergency response commanders

This study consisted of semi-structured interviews (Crang and Cook, 2007) with one commander from each of the addressed agencies. The interviews were held at the interviewees' places of work, providing a more natural and probably more comfortable setting for them. The interviews lasted between 45 minutes and an hour and a half. Each interview was recorded through audio recording as well as note taking. The audio recordings were transcribed in full and a summary of each of the interviews were written.

These interviews had a wider scope than the research presented in this paper, but parts of the interviews addressed the technology used by commanders in the agencies today, as well as their needs for new devices.

Study 2: Questionnaire

As Study 1 had only one subject from each agency and did not have its prime focus on the questions addressed in this paper, we conducted a second study using a questionnaire. This method was used to verify some of the findings from the interviews as well as to gather more information about the user group in general. Based on the interviews we particularly wanted to get a deeper understanding and verify data related to three topics: *Experience with different kinds of equipment*; *needs and thoughts around the use of equipment during an ER*, and *preferences in relation to screen size*. In addition to these topics, the questionnaire also contained demographic questions.

The questionnaire included 25 questions which were mainly closed, but open questions were used where we wanted more comprehensive answers. The closed questions used different answering formats, including restricted, partially open-ended or rating scale. The type of rating scale used was a five-point Likert scale which was used when opinions and attitudes regarding IT-equipment and ER were measured. The questionnaire was web-based, created in SurveyMonkey, but email was used to reach targeted users. The mail was sent to commanders that we previously had been in contact with, and contained a link to the questionnaire and a request to answer the questionnaire and forward the mail to other commanders they knew.

The analysis of the results from the questionnaire can be described as a simple quantitative analysis focusing on frequencies, percentages and measures of centre, in addition to cross tabulation (Bordens and Abbott, 2005). The analysis of the data gathered through the questionnaire was performed using the analysis tools in SurveyMonkey, as well as in Microsoft Excel by exporting the data from SurveyMonkey to a spreadsheet. The open questions were analyzed by doing a form of frequency count of recurring topics, while the quantitative data were analyzed using simple statistics.

In the analysis of the gathered data, it was found that a total of 89 people from the different addressed agencies answered the questionnaire, and 71 completed it. However, it turned out that several people not part of the user group also participated, including rescue workers working both at a lower and higher level than commanders. Of the 89 who participated, only 50 of the respondents can be said to be in the user group. Thus, only the data collected from these 50 participants was included in the analysis.

Study 3: Interview using a prototype as boundary object

For Study 3 and 4, a simple high fidelity prototype was developed focusing on the cooperation and coordination need identified in Study1. The prototype has two main parts: a map and a photo browser. The map part shows the incident and resources involved in the response. It allows the user to add zones and location, as well as filter and find detailed information about single resources. The photo part allows the user to browse pictures and zoom into single pictures. The prototype was implemented as a responsive web page, presenting the UI in an optimized way depending on the screen size and orientation. This means that although all devices ran the same software, the screen size, the actual layout of elements on the screen, as well as the size of the elements on the screen were different on each device – all handled by the responsive design features.

In Study 3 we used a hybrid method between an interview and an evaluation. This combination was used to gather feedback about the developed prototype and to discuss different screen sizes on which the prototype was running with actual end-users. The different devices running the prototype can be seen as props. Props are often used to enrich the data gathering session by providing the participants with a context and to help him or her with grounding their answers (Sharp, Rogers and Preece, 2007, p. 308). Such a prop may be viewed as a boundary object, which is used to span the gap between different competencies (Brandt, 2007). The devices included in this hybrid method were therefore used as props to help in the communication between the interviewees and interviewer, and as a way of providing some context to the users.

As the table solution is cumbersome to transport, the interviews were conducted in our offices (partly in a meeting room and partly in our lab). The data recording techniques used were tape recording and note-taking. One commander from each of the addressed agencies took part separately, and each session lasted between 40 minutes and an hour. The sessions started with a walkthrough of the prototype on a laptop. After the walkthrough, the commanders were asked some prepared questions regarding the functionality of the prototype in relation to their needs.

In the second part of the sessions, the participants were shown the same prototype on the other included devices, and these were then used as props to discuss screen sizes as well as other aspects of a device. The participants did not interact much with the prototype, but they took up and held the devices to get a feel of them. Also in this part, some prepared questions were asked as a starting point, but there were also room for more exploration based on the answers of the participants.

As the questions were quite focused, the answers were mostly about given issues and contained several similarities. Thus we did a simple qualitative analysis of the gathered data based on the identification of recurring patterns and themes.

Study 4: Usability test

To investigate whether the different devices running the prototype (and implicitly the different screen sizes) influence task performance and user satisfaction, we conducted a usability test to compliment Study 3. This

usability test involved controlled tasks where for example the ease of searching for and locating information was examined (Sharp et al., 2007, p. 644). This was chosen for several reasons, but mostly because we were more interested in exploring different aspects and gathering rich data than to test hypotheses and to be able to make predictions. As there in addition were several variables involved, and as some of these were difficult to separate from each other in terms of effect, we assumed that the use of a less experimental form of usability testing would be more informative.

The usability tests were conducted at a meeting room at our offices, and were conducted with five researches and five students. Most of the researches had ER domain knowledge. All participants may be considered usability experts. The lack of domain knowledge for some of the participants was not a big problem as the prototype is simple and the user tasks were easy to explain and resembled tasks the participants had conducted in different domains.

Because of the relatively few participants, all the participants used every device in the evaluation. To avoid the effects of learning, which could bias the results, both the order of the devices as well as the set of tasks performed with the devices were randomized. The following types of tasks were used: 1) find information about a given car; 2) place a location; 3) set a zone; 4) find what time a given picture was added.

To reduce the learning effect from the first device used to the others, we had a short walkthrough of the prototype on the laptop to present the user with different functionality, the terms used, etc. After the walkthrough, the participants were handed one of the devices and asked to perform one of the prepared sets of tasks. While using the device, the participants were asked to 'think aloud', so that we could get insights into their thoughts while interacting with the device and prototype. The participants' performance was measured by counting and noting down the number of errors as well as completion time for each task. After having gone through a set of tasks on one type of the device, they were handed the next device and asked a similar set of questions and so on until all the devices had been used. To assess satisfaction, the sessions were ended with an informal interview with a few questions about the design and screen sizes. Each session lasted between 20 and 40 minutes.

A limitation compared to Study 3 was that the table solution was not included. Even though responsive web design was used as an approach to development in order to avoid issues related to versions and platforms, it turned out that the PixelSense could display the prototype just fine, but that it was too difficult to interact with. It is also a limitation that the study was not conducted with real user in a realistic context. The former is not a major problem as the user tasks in this study are generic and simple, meaning that domain knowledge would not influence task performance. The latter is also a limited problem as the tasks are performed in a context where the user is standing still at an LCP. Light conditions etc. would of course be different in a realistic context, but this is independent of device size.

The quantitative performance data were analyzed in a similar way as the data from the questionnaire in relation to frequency distributions and measures of centre. As the quantitative data were gathered with a ratio scale, however, completion time was also calculated in relation to measure of spread (Bordens and Abbott, 2005). Qualitative data (to some degree gathered during the user tests, but mainly in the interviews) were analyzed in a similar manner as the data gathered in Study 3 by identifying recurrent patterns and themes. This part of the analysis was based on notes taken during the sessions as well as audio recording.

The quantitative data gathered with the usability tests were analyzed using Excel and concerned mainly two categories of data: completion time and errors. After having created frequency distributions based on the different devices, the mean of both time and errors were calculated for each of the four types of tasks. This was done to be able to compare the use of the different devices and the prototype in general, and to see whether there existed significant differences between the use of devices for specific tasks. For completion time, also a measure of spread was calculated. Even though few, the errors were divided into one of two classes of errors. The first class, called Error1, includes actions done intentionally, but which were wrong and/or made it difficult to complete the task, such as selecting the wrong resource or zooming too much out of the map. The second class, called Error2, includes errors which were unintentional, such as accidentally pushing a button which opened another window, or errors which were caused by the device or prototype. The errors were divided up to separate between errors caused by the design from more accidental errors. The unintentional errors are still included as they affected the completion time.

Equipment used in Study 3 and 4

In Study 3 the five devices presented in Table 1 were used. In Study 4, the first four of these devices were used.

Type of device	Product name	Screen size	Screen resolution	Pixel density	Aspect ratio
Smartphone	Samsung Galaxy S2	4,3”	480 x 800	218 ppi	16:9
Tablet	iPad Mini	7,9”	1024x768	163 ppi	4:3
Tablet	Samsung Galaxy Tab	10,1”	1280x800	149 ppi	16:10
Laptop	Asus K53S	15,6”	1366x768	100 ppi	16:9
Table solution	Microsoft PixelSense	40”	1920x1080	55 ppi	16:9

Table 1. Devices used in Study 3 and 4

The first three devices can be characterized as mobile devices, all using touch as input mechanism. The laptop has a larger screen and is operated though a keyboard and a touch-pad instead of mouse. The last device was the Microsoft PixelSense, which is operated through multi-touch and can support cooperation on a single screen, but which on the other hand is large, heavy and immobile.

FINDINGS

In this section we present the findings from each study separately. In the next section (Discussion), these findings are discussed in combination.

Study 1: Interviews with emergency response commanders

The participants in Study 1 expressed a clear need for better equipment when cooperating and coordinating with others during ER operations. In the analysis of the data gathered in Study 1, we identified the following types of equipment that are in use by commanders at the scene of incident today: radio (analogue or digital depending on district), mobile phones, paper maps as well as pen and paper. In addition, most health and fire vehicles and some police vehicle have support systems (map and task support). This equipment is attached to the vehicle and does not give specific support for the tasks performed at the scene of incident. The mobile phones, or smart phones, are mainly used to have direct dialogs with others, but is also used to gather information through different commercial apps and sometimes to capture and share pictures from the scene of incident with others involved.

All participants expressed needs for new devices – all having a preference for a tablet type of device. The participant from the ambulance service pointed out that the device should not be too large, while the participants from the police and fire agencies were concerned about the devices should not be too small. The fire department representative wanted a large tablet, (size of an A4 paper), as having a large display was considered more important than having the hands free. All participants expressed needs for having equipment that can be operated while using gloves.

Study 2: Questionnaire

The results from Study 2 showed that the responders have much experience in using ICT in their spare time, and a strong motivation for using new technology in their work. 100% of the respondents use a laptop or a desktop computer, 83,7 % use a smartphone, and 68,8 % use a tablet in their spare time. 90 % of the responders are furthermore positive towards using new equipment in their jobs. Over half of the respondents answered that they totally disagreed or disagreed a little to the statement "I am afraid that new equipment may take up too much attention while I’m working". However, 34 % agreed a bit with this statement, and 2 % totally agreed. As the motivation is as high as 90 % for using new equipment, we infer these numbers to mean that they would like to use newer equipment, but that there are terms to the use of this equipment. 42 of the 50 respondents replied that they need to have their hands free during a response.

Regarding the responders preferences for screen sizes, 56 % answered large tablet, 26 % answered small tablet and 8 % answered smartphone. Laptop was preferred by only 6%, although the preference more than doubled when combined with tablet. One of the respondents thought that radio covered all of his or her needs, and another participant wanted a table solution but only when used in combination with tablets. From the cross tabulation, it was furthermore found a slight difference in preference based on agency, as several of the commanders from the ambulance service seemed to prefer a smaller tablet compared to other commanders. 68 % of the respondents did however see a need to be able to switch between devices of different size.

Study 3: Interview using a prototype as boundary object

In this section we summarize the reaction to each of the five types of devices used as props in the interviews, followed by a discussion about the use of more than one device.

Mobile phone. There were some mixed feelings about the use of mobile phones, but generally the participants thought that the phone would be too small during an ER. One of the commanders said that it was impractical both because of the small, virtual keyboard, and because the commanders were "grown-ups with eye-sights which are not getting better". The screen was thus perceived by this participant as too small to be used out in the field. Another participant also found the mobile phone to be too small, both to interact with and to see the necessary details in a map, etc. The third participant confirmed that the mobile phone was too small, but would like to have it when not working outside.

Small and large tablets. We present the findings regarding large and the small tablets together, as these often were discussed in relation to each other. All of the participants liked the tablets, and especially the large one, because of the size of the screen and the touch interface. The use of touch was perceived by one of the participants to be more intuitive and efficient compared to other input mechanisms. The commander from the police in particular favoured the large tablet, as it allowed him to operate it while driving alone. The small tablet was, compared to the large tablet, perceived as being more suitable for other police personnel working closer to the incident. The commander from the ambulance service however had some trouble deciding between the tablets. He liked the large tablet because of the large screen, but found it very practical to be able to put the small tablet in his pocket. A tablet was regardless preferred to the other types of devices.

Laptop. None of the participants thought that the use of laptops would be suitable during an ER. The commander from the police said that an incident commander could use it for other tasks, maybe when searching for missing people, but mostly when working at an office. One of the participants said that the laptop would be directly impractical, much because of the use of keyboard and mouse pad, and that it probably would be left in the car.

Table solution. There were strong opinions about the use of the table solution. The commander from the fire and rescue agency liked it and thought that the table solution had a reasonable size for use at an LCP, and that it could be placed in a bus his fire and safety agency has. In contrast, the two other commanders said that the table solution would be impractical to use at an LCP, and that it would be more suitable for those working at the operational centre. As a device used to support cooperation and coordination, they preferred more mobile devices. A table solution needs to be located in a vehicle or a tent, but these commanders would like to spend less time at an LCP if it is inside, and more time outside with the other personnel and closer to the scene of incident. In that way they have better overview and contact with the personnel, requiring less communication over the radio.

The use of several devices. The possibility to switch between different screen sizes was also discussed. The commander from the fire and rescue agency would have liked to have the table solution in the bus, but a large tablet when not at the LCP. The commander from the police would have liked to have a mobile phone, a small tablet, and a large tablet, dependent on the situation. As an example, he said that if he was in a meeting and received a message about an incident, he would like to have a mobile phone just to get a quick overview. When he later got to the car, he would then switch to the large tablet. The participant from the ambulance services did not see a need to switch, but thought that it would be best to use the same tool at all times to be familiar with it. This participant therefore thought that the best solution would be to have one kind of device which both could be fastened in the car and be brought with him when he left the vehicle.

Study 4: Usability test

In this section we present the results from the usability test for each of the four types of devices used. The quantitative analysis addressed the performance measures for each task, i.e. 1) finding information about a given car, 2) place a location, 3) set a zone, and 4) find what time a given picture was added.

For each task on each of the included devices the completion time, measured in seconds, and errors, divided into actions which were either intentional (Error1) or unintentional (Error2), were noted. Here we present the results of the analysis of these measures.

Table 2 shows the calculated mean of the time used for each task, a measure of spread, and the calculated means of the type of errors done while using the mobile phone.

Task	Completion time: Measure of center	Completion time: Measure of spread	Error1 (average)	Error2 (average)
Task 1	64,2	21	0,6	0
Task 2	10,5	7,03	0,2	0
Task 3	40	40	0,8	0,2
Task 4	10,26	9,35	0	0

Table 2. Task completion and errors made with the mobile phone

There were some issues in interacting with the phone which were largely due to the size of the icons and that the map loaded slowly. These issues particularly affected the results regarding question 1 and 3, which both involved having to select an icon. Some of the participants had few difficulties with these tasks, completing them in 3 to 4 seconds. Still, one of the participants had large difficulties selecting an icon, and ended up zooming in and out. He then lost his orientation as the map was loaded too slowly. This participant thus eventually gave up on task 1 and 3. For these tasks the interquartile range was therefore calculated instead of the standard deviation. Besides the problems with selecting the icons there were as such generally few errors, but the mean time used on the tasks are relatively high.

Compared to the mobile phone, the participants generally used less time and made fewer errors with the small tablet (results are presented in Table 3).

Task	Completion time: Measure of center	Completion time: Measure of spread	Error1 (average)	Error2 (average)
Task 1	6	3,6	0,1	0
Task 2	8,15	6,5	0,4	0
Task 3	8,3	2,9	0	0
Task 4	13,6	2,9	0,1	0,7

Table 3. Task completion and errors made with the small tablet

Type of errors in the Error1 class were usually related to interaction with the map, where for example wrong parts of the screen was selected resulting in the wrong dialogue box to appear.

The large tablet (see Table 4) had some different issues compared to the smaller devices.

Task	Completion time: Measure of center	Completion time: Measure of spread	Error1 (average)	Error2 (average)
Task 1	7,4	6,5	0,3	0
Task 2	5,95	3,5	0	0
Task 3	9,4	6,7	0,1	0
Task 4	16,2	12,9	0,2	0,1

Table 4. Task completion and errors made with the large tablet

The three errors categorized as Error1 for the first question were all about selecting the wrong resource. When asked to tell for example the ID of the police car nearest to the scene of incident, three of the participants selected an ambulance or fire truck instead.

The laptop was both the most efficient device to use, and the device used where the lowest number of errors were made (see Table 5).

Task	Completion time: Measure of center	Completion time: Measure of spread	Error1 (average)	Error2 (average)
Task 1	5,1	2,4	0,1	0
Task 2	6,4	5	0	0
Task 3	11,1	6,2	0	0,1
Task 4	8,2	2,7	0,1	0

Table 5. Task completion and errors made with the laptop

This includes the fact that one of the participants had problems pressing the mouse buttons, and therefore used

more time to complete the tasks.

Even though the laptop was found to be the most efficient device in terms of use, the completion time is not that much different from the use of tablets. Especially if the Error2 issues related to the tablets, such as the browser hiding the tabs on the small tablet, and the resulting additional time to complete the task could be discarded, the completion time on the tablets would probably be more similar to the laptop. The results of the use of the mobile phone are very different compared to the other devices, with higher task completion times and a higher number of Error1 type issues. These results are as mentioned heavily influenced by an 'outlier', and if his scores are removed from the results, the completion time for question 1 is approximately 22,4 compared to 64,2 sec. Furthermore, if we compare the results of the individual questions based on the different devices, especially for question 2 and 4, i.e., the tasks not including the selection of an icon, there are fewer differences across the devices compared to the other questions. The task completion for the tasks and the errors made on the mobile phone are regardless higher compared to the other devices.

The results of the qualitative analysis are to a high degree in compliance with the results of the quantitative analysis, which is not that strange as the issues encountered while performing the tasks also were commented upon. The mobile phone was the least preferred device. Some of the participants did not like the laptop, and some mentioned that they found it less efficient to use compared to the tablets where they interacted through touch. Which device the participants preferred was a bit more mixed and several mentioned more than one device, but they generally all preferred the tablets.

DISCUSSION

Even though some districts within the different agencies do have some IT support in their cars, this can only be used while inside a car. Study 2 shows that most commanders use smartphones in their spare time, and Study 1 shows that several of the commanders use smartphones at work, both to have direct dialogues with others and to gather information through different commercial apps. Some also use their phones for capturing and sharing photos during an ER. As the use of a mobile phones, besides from the radio, was found to be the most frequently used information and communication tool amongst commanders during an ER, one might assume that this would be the preferred device.

Yet, even though the mobile phone included in the evaluations had a relatively large screen compared to other mobile phones on the market today, and even though being a phone and as such have other functionality than other types of devices, the screen size was deemed to be too small to be practical during an ER. This was related to several factors, such as eye sight and interaction with a small keyboard. One of the interviewees from Study 1 also pointed to difficulties in operating a mobile phone in cold weather. In relation to the results of Study 4, one can arguably say that the problems related to interacting with the prototype on the mobile phone mostly were a result of the design, especially the small icons. On the tasks where the users were not required to select an icon to complete a task, the differences between using the different devices are smaller. Regardless, from the interviews that followed the usability tests, it was found that several of the participants perceived the screen size of the mobile phone as generally too small, as it amongst other things required them to zoom in and out to see both general and specific information.

The devices which can be characterized as having large screens, referring here to the laptop and the table solution, were on the other hand deemed to be too large. Even though having larger screens and thus the ability to counter some of the issues related to the mobile phone, the size of the device per se was perceived as being too big and impractical. The laptop was the least discussed device, as the commanders early discarded the device as impractical due to size and weight, and due the input mechanisms used. Seven out of the 50 respondents in Study 2 nevertheless answered that they would like to use a laptop when it could be used in combination with a tablet. The participants in the Study 4 mostly did not prefer the laptop, and some mentioned that they felt that it was more difficult to operate the laptop compared to the other devices which was operated through touch. Yet, the laptop was found to be the device which resulted in the fewest number of errors and the lowest completion time in the evaluation setting.

The table solution was not included in the Study 4 due to difficulties in operating the prototype with this device. There were however generally few commanders who wanted this device in Study 2 and 3. Even though one of the participants in Study 3 perceived the table solution as a good tool for supporting cooperation and coordination at an LCP, the two other participants perceived the device as something that rather should be used by those working indoors, and not at an LCP. This is an interesting finding in relation to previous work regarding the use of large screens within ER, as these sometimes have been based on the assumption that the leaders who are responsible for the cooperation and coordination during an ER need large screens which can be operated by several users simultaneously. Our findings, however, indicate that there is a larger need for mobile

devices. As pointed out by one of the commanders participating in Study 3, the cooperation and coordination can also happen over smaller screens which display the same information.

The most preferred type of device of the commanders was the tablet. This was very clear in Study 1, 2 and 3. Tablets were preferred due to the size of the screen, which was perceived as being able to provide enough information and a good overview when displaying a map. Even though the mobile phone also was operated through touch, the use of this input mechanism was in addition to size one of the reasons why the tablet was preferred. Especially the participants in Study 3 mentioned this as a positive feature, as it was perceived as a more intuitive and faster way of interacting with the device. Even though a tablet was generally agreed to be the most fitting type of device, there were different preferences when it came to the size of the screen. Even though most of the commanders who took part in Study 1 and 3 liked the tablet with the large screen due to information needs, some also liked the small tablet due to the possibility to store it in a pocket. Nevertheless, a large tablet was generally perceived as the best device for use by commanders during an ER. Even though the tasks performed on the laptop were conducted in both less time and with fewer errors compared to other devices, all the participants in Study 4 preferred interacting with the tablets, due to both size and input mechanism.

Even though a tablet generally was preferred, there may nevertheless be a need for flexibility when it comes to screen sizes. A majority of the respondents in Study 2 wanted to be able to switch between the use of various devices, and this was also requested by two of the commanders in Study 3, where different devices were thought to cover different needs whether they were inside a car, at an LCP, etc. The last participant on the other hand thought that the same device should be used at all times to be familiar with it. In a similar way as the need for a system to be used regardless of the magnitude of an ER, the device used can thus be argued should be the same in relation to the principle of similarity, proximity, subsidiarity and cooperation. Still, being constrained to one size may be limiting in certain situations. Providing the user with the flexibility to choose for him or herself what is most suitable in any given situation may be a way of enhancing the use of the system.

Storage is another aspect. Almost all the participants of Study 2 replied that they needed to have their hands free during a response. Even though commanders often work at some distance from the scene of incident, with less contact with victims, etc., their main task is not to operate a device. Instead, they are more concerned with moving around the scene of incident and maintain contact with their crew. They also have to operate multiple radios and their mobile phones. Hence, for a device to be used, there should be a way to both store and retrieve the device efficiently. One of the reasons for why the commander from the ambulance service liked the small tablet was as mentioned that he easily could store it in his pocket. If a large tablet is to be used, the storage of this device ought to be investigated as this potentially can hinder the use of the device. Surprisingly, none of our participants noted a possible benefit from using a small tablet, i.e. that it is easier to use than a large tablet when standing. The reason for this is that small tablets may be held by one hand and operated by the other. Using a large tablet in this way is more difficult.

CONCLUSIONS AND FUTURE RESEARCH

In this research we have conducted four studies to investigate the most appropriate devices for high-ranking emergency leaders (commanders) working outside, close to the scene of incident. In Study 1 we interviewed commanders in the three main emergency agencies, namely the police, the fire and rescue agency and the ambulance services. In Study 2, 50 commanders from the same agencies answered a questionnaire focusing on use of technology and preferred devices with respect to screen size. In Study 3 we interviewed commanders in the same emergency agencies using a high fidelity prototype as a boundary object to facilitate a discussion on device types and screen sizes. Finally, in Study 4 the same prototype was used in a usability test to measure task performance and to investigate user satisfaction on different device types with different screen sizes. In Study 3 and 4, a mobile phone (4,3”), a small tablet (7,9”), a large tablet (10,1”), a laptop (15,6”), and a table solution (40”) were used (the latter was excluded in Study 4).

The main conclusion from all the studies was a very clear preference for tablets. Participants from the police and the fire and rescue agency had a clear preference for large tablets, and saw clear benefits in being able to switch between different devices and screen sizes, while the participants from the ambulance services had a slight preference for small tablets and would like to use one type of device rather than switching. All participants from all the involved agencies expressed needs for a device to be operable in the environment in which it is to be used, e.g. that is should be possible to operate the device using gloves, and the need to store and retrieve the device efficiently.

As Study 4 was performed in an artificial setting using participants that are not the intended users, a clear direction for future research is to perform usability tests with real users in a realistic context applying emergency specific tasks. This could give different results than our studies, particularly regarding the difference between

large and small tablet. Such studies could also address issues like the weight, robustness and storage of devices, as well as use of ruggedized devices.

Another interesting topic, which is relevant also in other domains than ER, is the relationship between resolution and screen sizes. A clear trend in the development of screens on different devices is that there will be larger differences in pixel size between different screens. The effect of this should thus be explored.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°261817, the BRIDGE project (www.bridgeproject.eu) within the Security Programme SEC-2010.4.2-1: Interoperability of data, systems, tools and equipment.

REFERENCES

1. Andresen, L. K. (2013) Exploring screen sizes and design of GUIs within emergency response handling, Masters thesis, University of Oslo.
2. Bordens, K. R. and Abbott, B. B. (2005) Research Design and Methods: A Process Approach, 6. ed, McGraw-Hill.
3. Brandt, E. (2007) How tangible mock-ups support design collaboration, *Knowledge, Technology & Policy* 20(3), 179–192.
4. Busher, M. and Mogensen, P. H. (2007) Designing for material practices of coordinating emergency teamwork, *Proceedings of the 4th International ISCRAM Conference*.
5. Crang, M. and Cook, I. (2007) Doing Ethnographies, SAGE publications.
6. Dubé, G., Kramer, C., Vachon, F. and Tremblay, S. (2011) Measuring the impact of a collaborative planning support system on crisis management, *Proceedings of the 8th International ISCRAM Conference*.
7. Jiang, X., Chen, N. Y., Hong, J. I., Wang, K., Takayama, L. and Landay, J. A. (2004) Siren: Context-aware computing for firefighting, *Proceedings of the Second International Conference on Pervasive Computing (Pervasive 2004),* Lecture Notes in Computer Science Volume 3001, Springer, 87–105.
8. Kim, S. Y., Jang, Y., Mellema, A., Ebert, D. S. and Collins, T. (2007) Visual analytics on mobile devices for emergency response, *Proceedings of the 2007 IEEE Symposium on Visual Analytics Science and Technology, VAST '07*, IEEE Computer Society, Washington, DC, USA, 35–42.
9. Luyten, K., Winters, F., Coninx, K., Naudts, D. and Moerman, I. (2006) A situation-aware mobile system to support fire brigades in emergency situations, *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, Lecture Notes in Computer Science Volume 4278, Springer, 1966-1975.
10. Nilsson, E. G. and Stølen, K. (2011) Generic functionality in user interfaces for emergency response, *Proc. 23rd Australian Computer-Human Interaction Conference (OZCHI'11)*, 233-242.
11. Neuvel, J. M. M. and Zlatanova, S. (2006) The void between risk prevention and crisis response, *Proceedings of UDMS'06*.
12. Rauschert, I., Agrawal, P., Pyush, I. R. and Sharma, R. (2002) Designing a human-centered, multimodal GIS interface to support emergency management, *Proceedings of the 10th ACM international symposium on Advances in geographic information systems,* 119 – 124.
13. Scotta, A., Pleizier, I. D. and Scholten, H. J. (2006) Tangible user interfaces in order to improve collaborative interactions and decision making, *Proceedings of 25th Urban Data Management Symposium*, 15–17.
14. Sharp, H., Rogers, Y. and Preece, J. (2007) Interaction Design: Beyond Human-Computer Interaction, 2. ed, Wiley, Indianapolis, IN.
15. Smirnov, A., Levashova, T., Shilov, N. and Kashevnik, A. (2009) A hybrid technology for operational decision support in pervasive environments, *Artificial Intelligence Applications and Innovations III,* IFIP International Federation for Information Processing Vol. 296, Springer US, 3–12.
16. Taylor, B. J., Kermode, S. and Roberts, K. L. (2006) Research in Nursing and Health care: Evidence for Practice, 3. ed, Thomson, South Melbourne, Vic.
17. Way, S. C. (2009) Criteria for evaluating mobile technologies for crisis response, *Privacy, Security, Trust and the Management of e-Business*, 38–46.