Influence of mobile map size and user capacities on situation awareness tested in a virtual environment

Nanja J.J.M. Smets
TNO Human Factors
Soesterberg, the Netherlands
nanja.smets@tno.nl

Guido M. te Brake
TNO Human Factors
Soesterberg, the Netherlands
guido.tebrake@tno.nl

Jasper Lindenberg
TNO Human Factors
Soesterberg, the Netherlands
jasper.lindenberg@tno.nl

Mark A. Neerincx
TNO/ Delft University of Technology
Soesterberg, the Netherlands
mark.neerincx@tno.nl

ABSTRACT

In the near future, first responders may become equipped with mobile devices providing navigation, decision and communication support. Because of the complex and chaotic circumstances in which these devices will be used, the devices should support the creation and maintenance of adequate situation awareness. Extensive testing of such devices for crisis management in real-life is expensive, complex, risky and only possible for specific settings. Therefore, we developed a synthetic task environment that is suited for developing and evaluating new concepts. In this paper, we present the results of the first experiment in this environment. Participants had to rescue victims in a synthetic world, and were supported by a map of the area showing the location and orientation of the participant and the victims. The experiment focused on the effects of map size and user's spatial ability on the quality of the situational awareness that was developed by the first responders. Besides the results of the experiment, experiences with the use of a synthetic environment for evaluation and development purposes are presented.

Keywords

Situation awareness, emergency response, mobile maps, game-based simulation, synthetic task environment.

INTRODUCTION

First responders (firemen, medics or policemen) are the first emergency workers on the scene of a crisis, and have important tasks to perform: urban search and rescue, triage, exploration and informing new parties who arrive at the scene. The environment of first responders is constantly changing and the development of new mobile support tools can help first responders with communication, information sharing and maintaining good situation awareness (Reichenbacher, 2003). Such tools need solid evaluations to ascertain that the critical tasks are supported adequately and that possible risks for unforeseen negative effects, such as tunnel vision or distraction, are minimized. However, testing of mobile support tools in critical, dynamic and complex task environments is difficult with classical evaluation methods. For the evaluation of mobile devices and services that are developed for such environments, Streefkerk et al. present a framework for the selection of appropriate evaluation methods (Streefkerk, van Esch-Bussemakers, Neerincx and Looije, in press). This framework distinguishes seven constraints that affect the choice of a specific method: stage, purpose, complexity, participants, setting, duration and cost. When we look at the constraints for the evaluation techniques wizard-of-Oz, field test, game-based, and focus groups, game-based evaluation proves to be a good match for our research.

For the *stage* in development process, the framework states that game-based evaluation is suited for both the design and analysis stage. It can support the *purpose* of both formative and summative tests (i.e. respectively, to generate a design solution and to measure the usability of this solution). Integrated logging procedures make the measurements of performance data very accurate. Game-based evaluation can be used in medium and highly *complex* designs. It can be used with both representatives and "real" end-users (experienced professionals) as *participants*. It is often hard to get a lot of experienced professionals involved for a substantial period of their working time; representatives can participate in addition to answer some research questions. The *setting* of a game-based evaluation is artificial and under control of the researcher. The *duration* is relatively short (you do not have to wait for the incidents to occur) and the *cost* in time and resources are relatively low when you have a software environment available. For an introduction to games for scientific research, see Lewis and Jacobson (2002).

We use Unreal Tournament 2004 for the game-based evaluation. The choice for this synthetic task environment (STE) is described in de Greef et al. (de Greef, Keeris and te Brake, 2006), where it was found that our STE can be used to validate new user interface concepts. Te Brake et al. describe the rational behind the general cognitive engineering method that we apply for the generation and evaluation of user interface concepts (te Brake, de Greef, Lindenberg, Rypkema and Smets, 2006). This paper presents the first experiment with the STE to investigate effects of mobile support and user characteristics on situation awareness. We distinguish three research questions. The first question is whether there is a relation between map size and the quality of the situational awareness that is developed by the first responders. A better rescue performance and situation awareness is expected with a "tablet" map size that shows most of the emergency environment than with a "PDA" map size that shows a limited part of the environment. The second question is whether spatial ability has a substantial positive effect on the number of rescued victims, and whether this effect is dependent on the map size. The third question is whether the STE is a good environment to evaluate a mobile support tool.

In this paper we will first discuss situation awareness and the goal directed task analysis we used to define questions to assess situation awareness of participants. Then the experiment that was conducted will be explained. The Results sections will discuss the outcomes of the experiment and in the discussion we will interpret them. In conclusion a short summation of our findings is listed.

Situation Awareness and Goal Directed Task Analysis

According to the model of Endsley (Endsley and Garland, 2000), SA consist of three levels:

- Perception (level 1). Perception of cues in the environment is fundamental to situation awareness. Humans perceive the environment through one or more of their five senses (visual, auditory, tactile, taste or olfactory senses). Without good perception, situation awareness is incomplete and errors will arise.
- Comprehension (level 2). Comprehension of the perceived information is the second level of situation awareness. It encompasses how people combine, interpret, store, retain, and retrieve information, how they integrate different pieces of information, and how they relate the information to their goals.
- Projection (level 3). The ability to project from the current situation into the (near) future allows for timely decision-making.

SA was measured using Endsley's Situation Awareness Global Assessment Technique (SAGAT) (Endsley, Bolté and Jones, 2003), performance measures and a questionnaire that was filled in after the experiment. The SAGAT method consisted of suspending the task at randomly selected times (the screens of the participant went blank). Then questions were asked about the situation, after which the task was resumed. The questions that were asked at these interval points in the experiment were drawn from a list of specific questions concerning different levels of SA. The whole set of questions addressed the number of victims they had saved, indication on the map of landmarks or task significant events (a fire or an accident) and indication of participants' own location on the map.

The list of questions used for the SAGAT was composed after a goal directed task analysis was conducted of search and rescue tasks of firemen. A goal directed task analysis identifies the goals of the operator, the decisions that the operator must make to achieve these goals and the information needed to make this decision. We used a disaster plan model of the Ministry of Internal Affairs that distinguishes four clusters with 25 processes. These clusters are divided into: fire brigade, police, health services and government. We looked at the fire brigade cluster, whose main tasks are:

1. Fire fighting and preventing emission of dangerous chemicals
2. Rescuing people
3. Disinfecting humans and animals
4. Disinfecting vehicles and infrastructure
5. Observing and measuring
6. Warning civilians
7. Making accessible and clearing of areas, streets, etc,

In our goal directed task analysis we focused on rescuing people, with a special interest in coordination. Rescuing people was the overall goal for the task analysis. For this goal nine sub goals were identified. These resulted in 23 items of information needed for decision making. From the derived information items nine questions were formulated. A part of the goal directed task analysis is shown in Table 1. The goal of the fireman is: rescue all victims. Sub goals of rescue all victims are find a victim and find the exit of a building. To satisfy the goal and sub goals the fireman has to decide if the path to the next victim/exit is known. The information needed to decide if the fireman knows how to find the exit of the building are: place of fireman, place of exit etc.

Goal directed task analysis	**Goal:** Rescue all victims **Sub goal:** Find victim **Sub goal:** Find exit of building Verify path to next victim/ exit Place of myself Place of the victim/ exit Obstructions in path Check if information you have is last updated information
SA questions sample set	Indicate on the map which areas are obstructed. Indicate on the map where there was fire. Indicate your location at this moment?

Table 1. In the upper row is a part of the goal directed task analysis that was conducted to formulate the SA questions asked during the experiment. In the lower row is a subset of the SA questions that were asked during the experiment.

METHOD

Participants

Twenty participants participated in the experiment as paid volunteers. There were nine male and eleven female participants. Average age of the participants was 22 and they were all college students. None of the participants had any previous experience with search and rescue tasks. All participants had sufficient computer experience to be able to perform the task in a STE.

Task

The task of the participants was to search and rescue 30 victims that were distributed over the synthetic environment and shown on the map of the support tool. Participants were allowed to rescue victims in any order and take any route to that victim (more routes were possible). As secondary task the participants had to remember important landmarks and events (like a burning car) in the environment, because they could be asked about these by the commander. The participants moved through the environment using a game controller. They could move forward, backward, sideways, and turn. When the participant walked the support screen turned black, when they stopped the tool became available again. This was necessary to prevent the participants from navigating through the world solely by looking at their support map. In real life, it is not possible either to walk around and pay attention to your support tool (with the exception of head-up support tools), hence this appears to be a realistic solution. Victims in the game were rescued by walking up to them, after which they disappeared from the simulated environment and from the map.

Design

The experiment was within subjects. There was one indepent variable, map size, with two levels: a big map (tablet) and a small map (PDA). The small map showed a sub-section of the big map. To exclude any learning effect we used a Latin square design where we balanced map size with two comparable urban synthetic environments, participants were assigned to one of the four orders shown in Table 2.

1	Environment 1/tablet	Environment 2/PDA
2	Environment 2/tablet	Environment 1/PDA
3	Environment 1/PDA	Environment 2/tablet
4	Environment 2/PDA	Environment 1/tablet

Table 2. Latin square design of the experiment, with independent variable map size (PDA or tablet) and two synthetic environments (environment 1 or environment2).

A map of the area was available on the support screen. It showed a number of buildings (pink), office buildings (grey), streets (yellow/white) and green for parks, grass or trees. The colors of the map were taken from the colors of maps on www.routenet.nl. This way the map had a natural look and feel. In the task, victims on the map were indicated by a green circle and a unique number. The fireman himself was shown with a red or blue circle and a bar indicating the viewing direction of the fireman, so the fireman knew which way he was facing on the map.

Figure 1. Left: screenshot from one of the environment screen. Right: the support screen, tablet and PDA conditions, of which one was provided to the users.

The participants could see the simulated environment, on the environment screen, in the form of a 3D visualization. An impression of the environment is visualized in Figure 1. Victims and other firemen were not rendered realistically, because this is complex and not required for our purposes. They were taken from the Unreal Tournament 2004 library.

Dependent variables

Below is a list of dependent variables that were measured for the search and rescue task:

- Performance data

  This includes the distance participants walked in the environment, the time it took to complete the pre -and post-test (see procedure) and the number of victims rescued in the game.

- Situation awareness

  The perception and comprehension (level 1 and 2) of SA were covered by the SAGAT method, the questionnaire, and the performance measures. The SAGAT method focuses on global SA and has the advantage that it avoids retrospective recall like in the questionnaire. The questionnaire is less intrusive and does not interrupt the scenario played by the participants. Performance measures are an indirect indication of the quality of SA, and can be obtained in a non-intrusive way.

- Subjective data gathered in questionnaires

Procedure

At the beginning participants were given a general, written instruction about the experiment. Then participants had to fill in a general questionnaire containing questions about computer and game experience. Subsequently a spatial test was conducted. After instruction participants did a pre-test, to asses their basic skill, in an environment similar to the real test environments but smaller. Then they were given instruction of the task. Participants completed two blocks of 30 minutes, in between these blocks there was a break of ten minutes. Each block consisted of the same task but in a different environment and a different map condition (PDA or tablet condition). Afterwards the participants conducted a post-test to asses practice effects and filled in a questionnaire about the experiment. The experiment took approximately three hours.

Material

In the experiment each participant used two computers, both with a 17 inch computer screen. The support screen displayed either a small or a large map of the environment; the environment screen displayed the simulated environment from a person's perspective, see Figure 1. The support screen was situated on the right, in front of the participant. The environment screen was situated on the left, in front of the participant. To navigate in the virtual environment the participants used a Logitech game controller. The experiment was modeled in Unreal Tournament 2004 editor and additional programming was done in C#.

RESULTS

Performance Data

T-tests were conducted on the performance data. There was no significant difference in the distance that the participants walked in the simulated environments in the PDA condition or the tablet condition. There was a significant difference in the total number of seconds participants walked their route when they walked the route for the first time or the second time, $t(17) = 2.56$, $p < 0.05$. Participants were slower the first time than the second time.

A T-test was performed on the number of victims that were rescued. There was no significant difference in the number of victims that were rescued with different map sizes. The results were also analyzed according to their score on the spatial test that was conducted. The results of the participants were divided in two nominal scores; one group had better spatial abilities than the other group. Eleven participants had a high spatial ability and nine participants had a low spatial ability. A MANOVA was performed to determine whether the number of rescued victims saved in the tablet and PDA condition differed for participants with a high or low spatial ability. A non-significant trend was observed of an interaction between spatial ability and map size, $F(2, 17) = 2.49$, $p = 0.11$. Participants with a high spatial ability tended to rescue more victims in the tablet condition than did participants with a low spatial ability. Spatial ability had no effect on the number of victims rescued in the PDA condition.

To further investigate this trend a multiple linear regression analysis was performed, see Table 3. A regression analysis shows that spatial ability explains 27 % of the variance in the number of saved victims in the tablet and PDA condition together. The number of saved victims was higher when the spatial ability of the participant was high. A regression analysis shows that spatial ability explains 50 % of the variance in the number of saved victims in the tablet condition alone. In the tablet condition the number of saved victims was higher when the participant had a high spatial ability.

Criterion variable	Explained variance R^2 (%) by the predictor variable	Regression equation
Number of victims rescued (NVR) in PDA and tablet condition together	Spatial ability (SP) 27%	NVR = 9.244 + 0.562*SP
Number of victims rescued in tablet condition (NVRtablet)	Spatial ability (SP) 50%	NVRtablet = 1.640 + 0.874*SP

Table 3. The percentage explained variance the different predictor variables add for the different criterion variables and the regression equation.

Situation Awareness

T-tests were conducted on the SA questions data. Only SA questions answered within fourteen minutes have been analyzed, this left five out of nine SA questions for analysis. Some participants ended the rescue of all victims in fourteen minutes. To be able to compare the data among all participants we cut off the data at fourteen minutes for all participants. There was a significant difference for question two: how many victims did you rescue? Participants were better at answering this question in the tablet condition than in the PDA condition, $t(19) = -4.68$, $p < 0.05$. For the other questions no significant effects were found (indication on the map of landmarks or task-specific events (fire, accident) and indication of participant's own location on the map).

Subjective Data

Subjective data was gathered by a question list that had to be filled in after the experiment. In Figure 2 is shown how participants navigated through the STE: by map, by looking in the STE or both. In the tablet condition the map was used more than in the PDA condition. In the PDA condition participants looked more in the STE or used both.

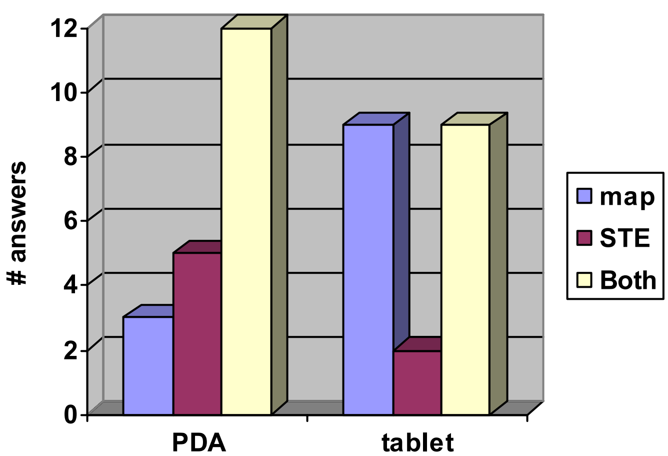

Figure 2. Answers to the question how the participants navigated in the virtual environment in the PDA condition and tablet condition, by looking at the map, the STE or both.

DISCUSSION

We were interested in the impact of map size on situation awareness and route planning efficiency. Participants in the tablet condition had a better overview of the amount of victims they rescued than participants in the PDA condition. But further no significant effects for performance data were found. This result may be due to the number of victims that was so large and the walking speed so high that no efficient route planning was required to find victims quickly. The participants could just walk around in a Pac-Man way and rescue a victim when they saw it appear on the map. Maybe an effect would have been found if there were not so many victims to be rescued and the walking speed in the environment was lower. This way planning is needed to rescue all victims. No significant effects were found for SA, although it turned out that the SA of the participants was rather low. This might be due to the task description where rescuing victims was said to be the most important task and maintaining SA a secondary task. Sayers (Sayers, 2004) summed up factors associated with being disoriented or lost in a virtual world. Below is a subset of these factors that applied to this experiment and could have added to the low level of SA from the participants in this experiment.

- Lack of support for speed control: Disorientation can be caused by a navigation speed that is too fast.
- Restricted field of view available on the screen: The field of view is different on a screen than in the real world.
- Lack of identifiable cues (landmarks): Were there not enough landmarks for the participants to navigate by and build up spatial knowledge?

These factors need to be taken into account with a following experiment.

A non-significant trend seems to suggest that participants with a high spatial ability rescued more victims in the tablet condition than participants with a low spatial ability. Spatial ability had no effect on the number of victims rescued in the PDA condition. This might mean that participants with a high spatial ability were able to look on the map and plan a route in the tablet condition. In contrast, in the PDA condition where spatial ability was of no effect, all participants could just walk a bit and look on the map. These results were also supported by the subjective data, where it was found that participants in the PDA condition navigated by looking mainly on the STE or both map and STE. In the tablet condition participants looked mainly on the map or both map and STE. It might be interesting to investigate in future research if people with less spatial ability can improve their performance by an adaptive interface. An adaptive interface can change the content, presentation or dialogue real time without explicit control by the user. If an adaptive interface is properly designed it can facilitate the performance of the user. If not properly designed it will however degrade performance.

During the experiment and analysis some important advantages of synthetic environments became evident. The software makes it possible to log events in a scenario that would be difficult to log in a real life environment, for instance the route walked by the participants, pauses, and the time they used the support tool. Logging in Unreal Tournament 2004 can be done extensively, because every interaction with the environment is captured. It is easy to extract the data from the game environment and then process it into useful visualizations to interpret the results, in Figure 3 is an example shown of routes participants walked in the environment.

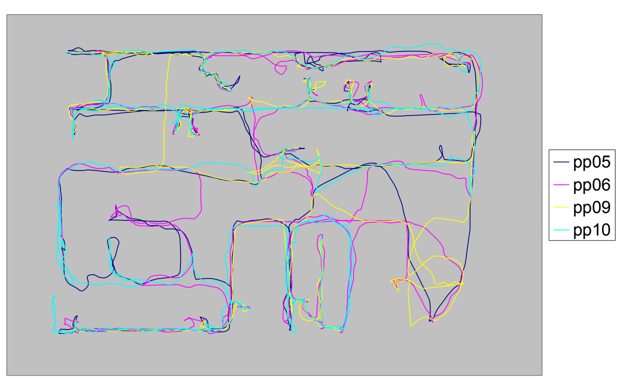

Figure 3. Routes of four participants in the tablet condition.

Participants' navigating performance improved during the experiments, as was shown by the decrease in time required by the participants to walk the trial run after the experiments. This shows that practice before the experiments is required to eliminate this effect. An alternative is to correct the measurements based on the speed difference right before and after the experiment. It was also found that speed differences between participants were considerable, mainly correlated with gaming experience and age, which should be taken into account when designing experiments. Use of the platform did not cause problems for the participants. This can be explained by the fact that most were computer experienced university students, but participants without gaming experience also found use of the virtual environment and the game controller to be easy. No problems with simulator sickness were encountered during the experiment, although in pilot runs when the turning speed was higher and people were sitting closer to the screens, one person had to stop because she did not feel well. Hence, simulator sickness is an issue that one should pay attention to when designing this type of experiments.

CONCLUSION

We were interested in the relation between map size and the quality of the SA that is developed by first responders. Participants did have a better overview of the number of victims they had rescued in the tablet condition than in the PDA condition, but no other significant effects were found for map size and SA. Results suggested that spatial ability had an effect on the number of rescued victims, which can be attributed to the tablet condition. The STE was considered a good environment to evaluate a mobile support tool. The software makes it possible to log events in a scenario that would be difficult to log in a real life environment. It is easy to extract the data from the game environment and then process it into useful visualizations to interpret the results.

ACKNOWLEDGMENTS

The research reported here is part of the Interactive Collaborative Information Systems (ICIS) project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024. The ICIS project is hosted by the DECIS Lab (http://www.decis.nl), the open research partnership of Thales Nederland, the Delft University of Technology, the University of Amsterdam and the Dutch Foundation of Applied Scientific research (TNO).

REFERENCES

1. Endsley, M.R., Bolté, B and Jones, D.G. (2003), in Taylor and Francis, Designing for situation awareness.
2. Endsley, M.R. and Garland, D.J. (2000). Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates (147-173). Mahwah, USA.
3. Te Brake, G., de Greef, T., Lindenberg, J., Rypkema, J., Smets, N. (2006). Developing Adaptive User Interfaces Using a Game-based Simulation Environment. In Proceedings of the 3rd International ISCRAM Conference. Newark, NJ (USA).
4. De Greef, T.E, Keeris. E., and te Brake, G.M. (2006) A Game-based Experimentation Environment. ISAGA 2006, St. Petersburg.
5. Lewis, M., and Jacobson, J. (2002). Game Engines in Scientific Research. Communications of the ACM, 45, 1, 27-31
6. Reichenbacher, T. (2003). Adaptive Methods for Mobile Cartography. The 21th International Cartographic Conference, pp. 1311-1321.
7. Sayers H (2004). Desktop virtual environments: a study of navigation and age. Interacting with computers 16
8. Streefkerk, J.W., van Esch-Bussemakers, M.P., Neerincx, M.A., and Looije, R. (in press). Evaluating context-aware mobile interfaces for professionals. In: Lumsden, J. (ed.), Handbook of research on user interface design and evaluation for mobile technology. Idea Group