EXPLORING THE USABILITY OF DIFFERENT BASE MAPS FOR EMERGENCY RESPONSE CENTRES: A CASE STUDY IN THE CZECH REPUBLIC

Zbyněk Štěrba, Milan Konečný

Laboratory on Cartography and Geoinformatics
Department of Geography, Masaryk University
Kotlářská 2, 611 37 Brno, Czech Republic
zbynek.ste@post.cz, konecnymilan3@gmail.com

Abstract

The proposed contribution is focused on evaluation process of cartographic visualizations for emergency response centres in crisis management. We evaluated the suitability of three versions of background topographic maps, two versions that have been used in emergency centres in the Czech Republic, and one newly proposed visualization. The experiment consists of comparative multi-criteria test that measures the level of usability of selected crisis maps used by emergency response operators. During this specific test, users were supposed to find the required cartographic symbols on the different background maps, while response time and accuracy of the solution in particular tasks were stored to database. Eventually the results of the evaluation process were interpreted with an emphasis on the usability of the selected maps, where significant differences were found. Results of the experiment showed the importance of the particular cartographic visualization that is used in crisis management.

1. INTRODUCTION

The law requires that crisis management respond to crisis situations arising on all levels as a result of emergencies such as natural disasters or industrial accidents. A means to this end in the Czech Republic is the Integrated Rescue System, which has three basic components: the Fire Service of the Czech Republic, the Police of the Czech Republic and the Emergency Rescue Service. An essential tool in responding quickly and effectively to emergencies is a suitable geoinformation system (Staněk et al, 2011, Konečný and Reinhardt 2010). It is similarly important to effectively visualize and communicate this information to the end user, who is most often an employee of an emergency response centre.

At present, the individual branches of the Integrated Rescue System employ various data bases deriving from various sources, both state and private. Although the situation has been improving rapidly in recent years, the mutual sharing of data among the services is rare, with negative consequences for interoperability. The usability of these information systems is further reduced on the presentational level due to a lack of coherence in cartographic notation, which is often created or adapted for ad hoc use with little regard to conceptual unity. Emergency response maps created in such a manner may lead to ineffective or even erroneous decisions by the responding authorities (especially operators, incident commanders and members of response units).

In recent years, observations at emergency response centres have led to the discovery not only of certain technological deficiencies but also considerable imprecision in the cartographic symbology in use (background maps and figure symbols), which has resulted in less effective and slower response. Some of the represented items, for example, were depicted in a manner that significantly complicated the process of finding them. Consistent with these findings are complaints by the operators themselves regarding the cartographic visualization, which is difficult to read and from which, especially in a stressful environment, it is difficult to obtain and then convey precise and complete information (Dror et al 1999). Meanwhile, it is generally true that a map's readability is largely determined by the possibility of distinguishing *figure* information from *background* layer (Stachoň et al, 2013; Robinson et al 1996)).

Various symbology sets and methods of cartographic visualization have already been proposed within the context of several studies for use in emergency response centres for the dispatch of responders. The model's basic problem, however, was the suitability of various types of background maps for concrete tasks, which depend on the extent to which individual cartographic means of expression are used. The use of certain visualization instead of others results in more effective perception by the user, and the resulting map then has greater potential to solve a particular problem. It was determined that in many cases certain visualization methods are inadequate to the task of effectively communicating needed information and they inhibit workers' performance during emergency situations. For this reason, special maps were proposed at the Department of Geography of Masaryk University which would allow the effective dispatching of responders during emergencies (Konečný et al 2011).

Every cartographic work must be not only created but also critically evaluated so we may ascertain its usability to its purpose and to the demands of the user. This paper addresses the problem of empirically evaluating the usability of crisis management maps with an emphasis on use of suitable topographic underlay, the contents of which are often suppressed for the sake of the thematic overlay (figure). Before every evaluation it is essential to know what we are evaluating and what characteristics the evaluated subject (in this case, a map) is supposed to have. Given that crisis management presents very complex problems, composed of many different approaches to cartographic visualization, evaluating the usability of individual cartographic products is a problem with its own rather specific character. A map's usability must in every case be deduced only by examining its relation to its user.

2. METHODOLOGY AND STUDY DESIGN

For the presented case study, a newly created cartographic background topographic map (called *Basetopo*) was compared to existing maps used in crisis management operational centres. The usability of these maps was tested empirically as users performed practical tasks while using the maps. The study used objective methods while employing the interactive testing software MUTEP (see Stachoň et al 2013), which facilitates such testing. While a method for the empirical testing of cartographic symbols for usability has not yet been developed and standardized in detail, the design of this experiment was based on earlier experiences (Stachoň et al 2013; Štěrba and Šašinka 2012; Brychtová et al 2012).

2.1. Experiment description

The test's object was the evaluation of the usability of three background maps created for a single purpose. These underlays were a visual variable and the main subject of the experiment. The experiment entailed a comparison of these three equivalent topographical underlays for the purpose of demonstrating the influence of a given map's cartographic treatment on the effectiveness of operators' decision-making in crisis situations. The primary motivation was determining and demonstrating differences among individual underlays in their negative effects on users' perception when focusing attention on the figure symbols. Several map segments selected for this purpose were used successively for individual tasks in the test. The selection included maps of both urban and rural landscape types. The experiment was composed of several portions, in which respondents performed map-related tasks that gradually increased in their complexity.

2.2 Stimuli

The case study evaluated two versions of topographic background maps either previously or currently in use (known here as *variant 1* and *variant 2*) and a newly created map known by the working title *Basetopo* (see examples in Figure 1). The main characteristic distinguishing the underlay *Basetopo* from the two variants is the suppression of colour saturation and the simplification in the depiction of certain point and line elements. It may be assumed that such a reduction in visual complexity will increase the readability of items belonging to the figure – that is, the thematic overlay. The figure, which in this case represented a visual constant (its elements were the same for all three underlays), was made up of points-of-interest objects and aspects of emergency infrastructure, represented in most cases by discrete symbols on the map. The principles when creating these map symbols were, among other things, that the size of individual symbols should be uniform, that their shape and structure should be simple and that they employ a minimum number of colours (Tajovská, 2011). A set created in this manner is to a certain extent not dependent on the underlay used. All of the chosen map symbols used in the experiment thus correspond to usual qualities imposed on items belonging to the figure (use of intense colours, clear and pronounced bordering, a high level of associativeness; MacEachren 2004; Kraak and Ormeling et al 2003). It may thus be assumed that they will create a strong contrasting effect with respect to the items belonging to the map background.

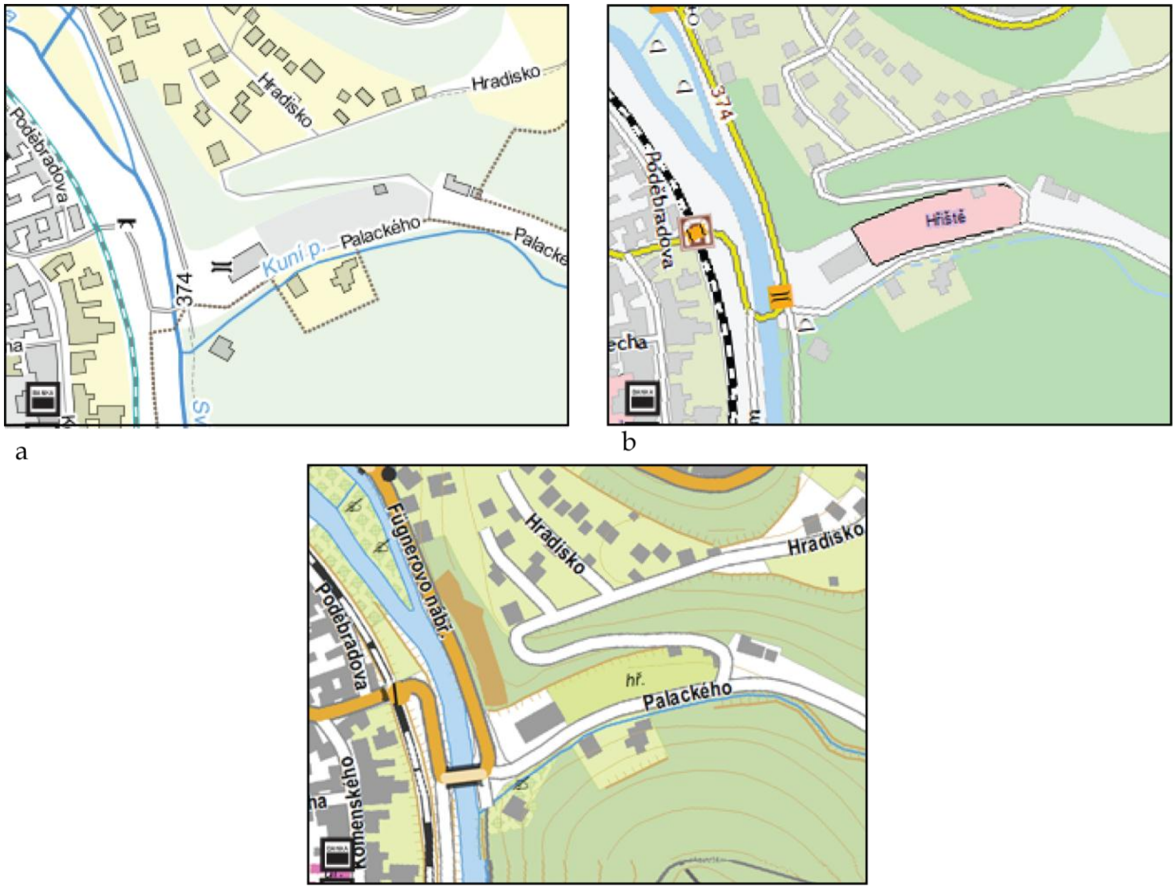

Figure 1. Examples of topographic background map variants for emergency response centres used in the experiment: (a) variant Basetopo; (b) variant 1; (c) variant 2; all map segments are of a scale 1:5,000.

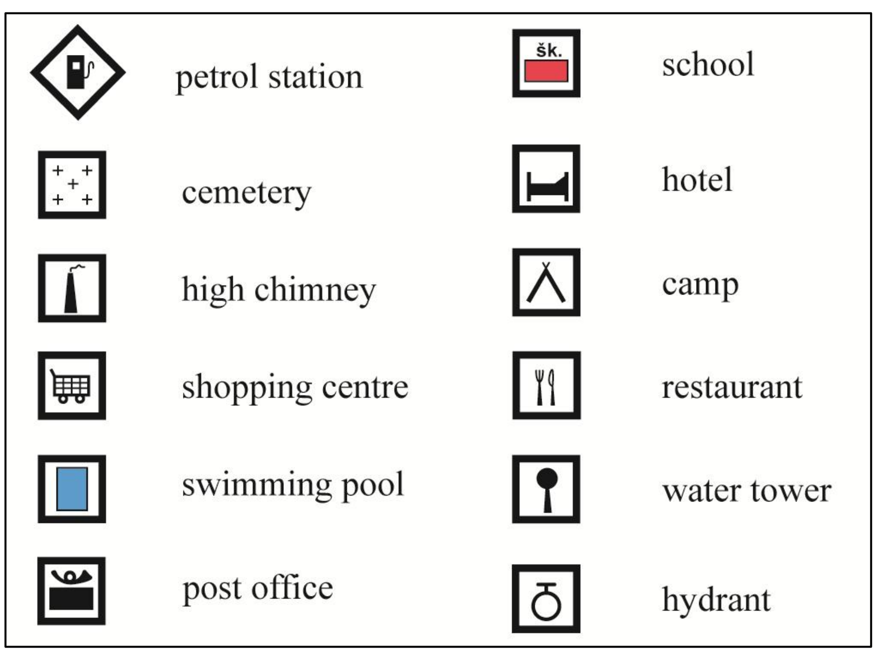

Figure 2. Example of figure symbols used in the experiment

2.3 Procedure

The entire experiment followed an inter-subject plan in which each respondent was exposed to only one level of the independent variable (the topographic map). The dependent variable was effectiveness and efficiency in performing the given tasks with the map variant, measured by the success of each respondent (measured response time and error rate).

Altogether, participants in the experiment were comprised of 137 students from Masaryk University with an average age of just under 22 years. Only 25 respondents claimed to use maps very rarely. Thus we can consider the overall map literacy of the group in question to be average. An important (if expected) piece of information was the fact that 129 respondents claimed to use a computer every day. Given that the experiment was conducted on a computer, an assumption of the ability to use one was essential. We may thus state that any errors caused by poor or slow control over the test module were negligible and had no influence on the results. All respondents were randomly assigned to three groups.

For the sake of validity, the experiment was conducted in a computer laboratory with near uniform conditions with respect to lighting and ambient noise. All respondents also used the same type of monitor. The functionality of all computers was checked before the test itself in order to minimize technical errors during the experiment. Each respondent completed the test on a computer, after which a short interview was conducted in which each respondent was asked about problems encountered during the test and his or her questions about the test were answered.

A detailed text was presented before each portion of the test, explaining the principle governing the tasks that followed. Given that the more complex tasks were presented toward the end of the test, we may assume that all respondents were already able to understand and control the test module without problems and that the overall results were therefore not significantly influenced by this factor. The first portion of the test was composed of 10 tasks in which a respondent had to search for and mark one map symbol by clicking on it. After the description of each task, the respondent was shown an image of the desired map symbol which he or she then had to locate on a map field. The order of map segments was produced randomly. Individual map symbols were successively placed in various locations on the map field, which was meant to increase the ecological validity of the experiment. The response time required to locate the desired item was automatically measured and recorded from the moment the particular map appeared on the computer monitor. This information, together with the precision of the location, was later entered into the database. The task description provided to the respondent and the corresponding map sample are shown in *Figure 3*. In the same manner, the respondent completed all 10 tasks presented. The test given to each group of respondents was assembled in identical manner, the difference always lying in the independent variable (that is, in the topographic background map employed). Individual elements of the figure layer were always placed in an identical manner.

The second portion of the test proceeded in manner analogous to the first. The only difference was in the number of symbols to be located and marked in the map sample. In every case, the instructions called for finding *all* symbols on the map but did not indicate just how many there were. The respondent was required to carefully examine the entire map field so that no item would be overlooked. An additional point of interest when interpreting the results for this portion was the error rate. For the purpose of evaluating the test, information regarding precision of identifying symbols was recorded too. The second portion also consisted of 10 tasks in a sequence identical to that in the previous portion.

The final portion of the cartographic test contained the most complex of the tasks. Solving these tasks demanded more complex cognitive processes than merely finding and distinguishing shapes, as the other portions demanded. Here it was required that the respondents perceive information from the background, i.e. from the topographical map. In the course of this portion, respondents were also presented with tasks involving various map underlays in a simulation of the transitions among differing visualisations that is often required of crisis operators. A second background visualisation employed was an orthophoto map (with some thematic items added), which is often used at operations centres to find information in situations on a large scale requiring detailed depiction of the terrain in question. This portion was composed of eight tasks alternating between a topographic and an orthophoto map. Each task using a topographic map was followed by a task using the same segment from an orthophoto map, thus simulating this very transition among various visualizations of a single terrain. Individual tasks consisted above all in finding the requested objects using both background and figure layers. In this way, we determined the usability of particular topographic visualizations, using all tested variants (e.g. “Find the intersection of Jehnický Brook and the road,” etc.).

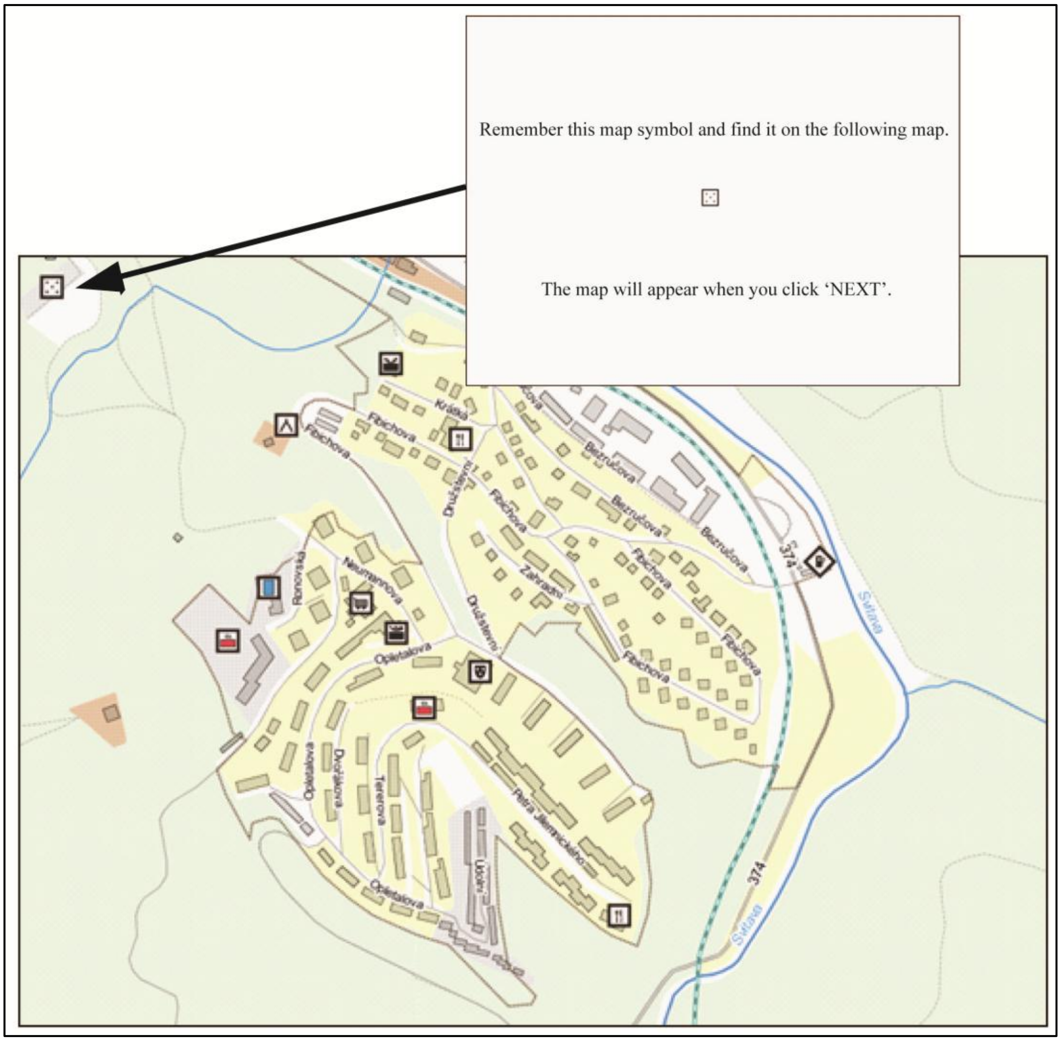

Figure 3. An example of a test task (a reduced image of the test task description, above right) and a map sample used in the first portion of the test; map symbol to be found by respondents is indicated by arrow (the example shows the Basetopo variant).

3. RESULTS

In the course of testing, all results were successively stored into the database from which they could later be easily exported with the help of the external module. In the first portion of the test there was no statistically significant difference discovered in the reaction times among individual groups, the average overall values not varying by more than one second (M_1=37.8 seconds, SD_1=9.5 sec.; $M_{basetopo}$=38.3 sec., $SD_{basetopo}$=7.5 sec.; M_2=38.5 sec., SD_2=10.9 sec.). Similarly, neither was there any statistically significant difference discovered in any individual task.

A decisive aspect of the results of the test's second portion was the time taken to complete the task, an expression of respondents' level of confidence regarding its correct completion. It may be assumed that the respondent proceeded to the following stage (task) of the test only after the subjective completion of the task and through the conscious transition to the following task by clicking the appropriate tab. Table 1 shows statistically differing values for four individual tasks. Overall times for completing all tasks show results for the group with the visualization *Basetopo* (87.3 sec.) differing significantly from those for the other two groups, *variant 1* (98.3 sec.) and *variant 2* (102.3 sec.). Similarly, the values of the standard deviation show a higher level of variability among the data for these two groups, reaching as much as two times that of the group with the variant *Basetopo*.

Table 1. Average response times for completing tasks on individual map segments in second portion of test: locating two or more items on a map (asterisks indicate significantly lower values, while boldface indicates significantly higher values).

task nr.	**map segment**	**basetopo**		**variant 1**		**variant 2**		**analysis of variance**	
		time [s]	***stan. dev.***	***time [s]***	***stan. dev.***	***time [s]***	***stan. dev.***	***F test***	***p***
1	C	*9.7	2.4	***11.2***	4.1	***11.9***	3.8	4.7	0.01
2	J	10.3	3.5	10.4	3.4	11.3	3.8	1.2	0.30
3	I	*9.6	2.7	9.9	3.1	***11.7***	4.8	4.3	0.02
4	D	*6.8	1.6	7.2	1.9	***8.1***	2.8	3.5	0.03
5	H	6.2	2.4	7.0	2.5	6.6	1.0	1.4	0.24
6	F	9.3	3.4	9.7	4.6	10.0	3.3	0.4	0.67
7	A	6.3	1.3	6.5	1.6	6.9	2.2	1.6	0.21
8	E	*9.1	2.2	9.7	3.4	***10.8***	2.9	3.2	0.05
9	G	15.1	3.9	16.9	4.8	16.4	4.1	1.8	0.18
10	B	6.5	1.4	6.6	1.6	7.0	2.2	1.1	0.35
∑	**A-J**	*87.3	12.6	***98.3***	24.9	***102.3***	21.7	3.8	0.03

In the final portion of the cartographic test, in which respondents solved tasks of greater complexity on a map, the trend of the results from the preceding portion (see *Table 2*) was confirmed. The results obtained show a statistically significant difference in average reaction times for tasks on the topographic map underlay (tasks no. 1, 3, 5 and 7): the best results were again achieved by respondents using the *Basetopo* variant (overall time for these tasks was 43.2 seconds), whereas the average response time with *variant 1* was 65.9 seconds and with *variant 2* 65.8 seconds. However, for tasks solved on the orthophoto map underlay (tasks no. 2, 4, 6 and 8), it is possible to say that the observed results are similar among all the groups. This is demonstrated by the fact that the overall average time for completing all tasks never shows statistically significant differences across the groups: for the first group (*Basetopo*) 57.8 seconds, for the second group (*variant 1*) 62 seconds and for the third group (*variant 2*) 58.5 seconds.

Table 2. Average response times on individual map segments in third portion of test: complex tasks with map (asterisks indicate significantly lower values, while boldface indicates significantly higher values).

task nr.	**map segment**	**map type**	**basetopo**		**variant 1**		**variant 2**		**analysis of variance**	
			time [s]	***stan. dev.***	***time [s]***	***stan. dev.***	***time [s]***	***stan. dev.***	***F test***	***p***
1	F	topo	*7.9	3.9	***11.1***	5.9	***16.5***	9.7	17.5	0.00
2	F	ortho	13.2	4.4	14.3	3.9	14.3	3.9	1.1	0.35
3	D	topo	*9.6	4.3	***20.3***	11.2	***15.1***	8.4	13.7	0.00
4	D	ortho	12.8	5.1	15.6	7.3	14.1	5.2	2.4	0.09
5	A	topo	*19.7	8.2	***24.8***	9.9	***24.9***	9.4	3.4	0.04
6	A	ortho	8.7	4.8	10.5	6.5	10.4	6.2	1.1	0.32
7	E	topo	*6.3	2.7	***10.1***	5.8	7.2	2.7	10.9	0.00
8	E	ortho	21.7	14.3	23.1	17.3	17.9	21.7	0.8	0.45
∑	**A-H**	ortho	57.8	20.0	62.0	25.8	58.5	28.9	0.3	0.78
∑	**A-H**	**topo**	*43.2	12.5	***6.9***	18.1	***65.8***	20.7	16.0	0.00

A second aspect that was followed during the testing was the level of efficiency, expressed by error rate. Errors consisted both in marking an item that was not requested in the prompt and in not marking all of the requested items (for tasks in the second and third portion of the test). Evaluation of the test's first portion yielded almost no errors, so it was not necessary to conduct a comparison among the groups. The second portion (finding two or more items on the map), however, showed a certain level of errors. Here it was noted when a respondent failed to mark all requested items (particularly for tasks with multiple items in the map field). It cannot, however, be said that any of the tasks showed a statistically significant difference among individual groups. This is evident from the overall average error rates for all three groups (the error rate for the group with the underlay *Basetopo* was 4.6 %, for *variant 1* 3.5 % and for *variant 2* 5.8 %).

Error rates showed greater differences in the results for tasks from the third portion of the test. In this portion, statistically significant differences were discovered for two tasks using the topographic underlay (tasks no. 3 and 5), both tasks showing the lowest error rate when working with the *Basetopo* background map and the worst results when respondents were working with *variant 1*. As for the overall error rate for this portion, the lowest error rate was shown for the group working with the *Basetopo* map (2.7%); the error rate was statistically significantly higher ($p<0.001$) for work with *variant 1* (21.7%). The error rate for *variant 2* was not significantly different (6.7%). For tasks when using the orthophoto map, however, there was no statistically significant difference observed among the individual tested groups; given the use by all groups of the same independent variable, this is not a surprising result.

4. DISCUSSION AND CONCLUSION

The experiment's results delivered some rather interesting discoveries regarding the usability of the selected maps for the purposes of crisis management response centres. Results reveal differing effectiveness when working with all three of the selected background maps. It was demonstrated that the variant *Basetopo* is, given the defined purpose of the maps in question, the most suitable map underlay. The experiment confirmed our hypothesis that this is overall the most usable of the variants. We can assume therefore that, when using more complex maps with less pronounced differentiation between *figure* and *background*, increased cognitive difficulty associated with difficult tasks will result in significantly poorer performance.

In the case of simpler tasks concerning the location of a single requested symbol within the map, no significant differences were observed in performance effectiveness and efficiency among the tested groups, i.e. neither regarding response times when solving problems, nor regarding error rates. This fact is clearly a function of the character of these tasks, which consisted in merely finding a single item and therefore make no great demands on a subject's attention. Thus we may assume that any differences revealed during this type of task won't be so much dependent on the cartographic visualization employed. More difficult tasks, however, show rather different results, which may indeed be attributed to the particular visualization. In this case, the performance by the respondents using the map variant *Basetopo* was significantly better, which is illustrated by the average overall times required to complete all assigned tasks. While the group working with the map variant *Basetopo* required approximately 87 seconds to complete all assigned tasks without any errors, the third group, using *variant 2*, had an average time of more than 17 % longer. The group using map variant *1* showed a time of more than 12% longer than the group using the *Basetopo* background map. This difference, which is the result of nothing more than different way of visualization of the same elements of the map field, is a very significant result which testifies to the respective successes in fulfilling the purposes of a given map. This part revealed no significant differences in error rate. This is clearly due to the character of the particular tasks, which called for the location of items in the figure; the background, then, shows a negative influence only on the speed with which items are located, but not on accuracy when locating relatively obvious and unmistakable map symbols. Thus the error rate, which was for each group approximately 4 %, expresses in most cases a failure to mark all of the requested symbols.

In the last part of the cartographic test, the usability of the variants was tested through more complex tasks the purpose of which was to examine higher cognitive processes. Respondents were required to find and use a greater amount of information in the map using the topographic background. This portion fully revealed the differences among the individual map variants, both in response times and in the rate of error. Differing performances were observed in individual tasks and in the average overall time. From the results obtained we may say that, when completing tasks in the maps, response times on *variants 1* and *2* were on average 50 % longer than on the variant *Basetopo*, which is illustrated by the overall time for completing all tasks on the topographic map. From the character of these tasks, which require information also from the map background, we may clearly presume easier transfer of the user's attention between *figure* and *background* when using maps with less visual complexity, i.e. those making lower cognitive demands on the user. A particularly interesting observation from this part of the test concerns error rates, which reach their lowest point with the *Basetopo* map variant (2.7 % overall), the highest error rate of which for a single task is 6.5 %. The error rate for the other two map variants is higher and also reveals greater extremes (e.g. *variant 1*'s 63 % error rate for one task). We may presume that certain elements are represented erroneously, which in turn leads the user to respond erroneously. The tasks in the third part of the test, using the orthophoto map underlay, which were uniform for all groups, showed no differences, neither in response time, nor in error rate. For this reason, we cannot assume that the visual complexity of the field in question would have a significant effect on perception of information from the next represented image (that is, on a momentary change in visualization).

Given the breadth of the experiment and with it the high level of ecological validity, we may consider its conclusions regarding usability of individual map variants to be relevant to real situations in which operators dealing with crisis events perform tasks beyond merely finding certain symbols on maps. Based on our findings, we may claim that for the purposes attendant to crisis response the *Basetopo* map is more suitable than the other variants used in the experiment. We may also say that this fact is due to the lower visual complexity of the cartographic visualization employed and the more suitably chosen colour composition. When creating new cartographic visualizations it is always necessary to ask which elements are necessary and which expressive means are most relevant to their depiction. Further studies will surely have to address personal and intercultural differences among users, e.g. the influence of particular cognitive styles on the perception of visual information. Such factors must be considered, given the increasing use of global systems in support of crisis management decision-making processes. We may claim that the quality of cartographic production is every bit as important as the technological solutions governing these systems.

This work was supported by the project „Employment of Best Young Scientists for International Cooperation Empowerment“ (CZ.1.07/2.3.00/30.0037) co-financed from European Social Fund and the state budget of the Czech Republic.

REFERENCES

Brychtová, A., Popelka, S., Dobesova Z. (2012): Eye-tracking methods for investigation of cartographic principles. In *SGEM 2012 Conference Proceedings Volume II*, STEF92 Technology Ltd., Sofia, Bulgaria, 1041-1048 p. ISSN 1341-2704

Dror I. E., Busemeyer J. R., Basola B. (1999): Decision making under time pressure: An independent test of sequential sampling models. In *Memory* & *Cognition* [online], 1995, vol. 27/4, pp. 713-725.

Konečný, M., Kubíček, P., Šašinka, Č. & Stachoň, Z. (2011): Usability of Selected Base Maps for Crises Management - Users Perspectives. Applied geomatics, 3(4), 189-198.

Konečný, M., Reinhardt, W.(2010): Early Warning and Disaster Management: The Importance of Geographic Information (Part A). In *International Journal of Digital Eart* (editorial), 2010, vol.. 3/ 3, pp. 2017-220, ISSN 1753-8947.

Kraak, M. J., Ormeling, F. (2003): *Cartography: Visualization of Geospatial Data*, Great Britain: Prentice Hall, 2003, 205 p. ISBN 0-13-088890-7.

MacEachren, A. M. (2004): *How Maps Work: Representation, Visualisation and Design.* New York: The Guilford Press, 2004, 513 p., ISBN 0-89862-589-0.

Robinson, A. H., Morisson, J. L., Muehcke, P. C., Kimerling, A. J., Guptill, S. C. (1995):. *Elements of Cartography*, 6. vyd., Wisley&Sons, 1995, 664 p., ISBN 0-471-55579-7.

Stachoň, Z., Šašinka, Č., Štěrba, Z., Zbořil, J., Březinová, Š., Švancara, J. (2013): Influence of Graphic Design of Cartographic Symbols on Perception Structure. In *Kartographische Nachrichten*, 4/2013, pp. 216-220.

Staněk, K., Friedmannová, L., Kubíček, P., Konečný, M. (2011): Selected Issues of Cartographic Communication Optimization for Emergency Centers. In *International Journal of Digital* Earth, UK: Taylor&Francis, 2011, vol. 3/ 4, pp. 316-339, ISSN 1753-8947.

Štěrba, Z., Šašinka, Č. (2012): Evaluation of Maps for Crisis Management in the Context of User's Cognitive Features. In *Risk Models and Applications CODATA-Germany*, Berlin, 10 p., ISBN 978-3-00-036336-8.

Tajovská, K. (2011): *Mapová symbolika v krizovém řízení*, dissertation, Masaryk University, Brno, 2011, 166 p..

Zbyněk Štěrba, Ph.D. is a post-doctoral researcher in the Department of Geography at Masaryk University in Brno, Czech Republic. His research interests focus mainly on cognitive aspects of cartography, the role of particular graphic variables on map usability and differences among individual users in map perception processes with an emphasis on crisis management applications.

Prof. Milan Konečný is a professor in the Department of Geography at Masaryk University in Brno and Head of Laboratory on Geoinformatics and Cartography. During 2003-7, he was President of International Cartographic Association (ICA (contemporary past president), now serves as chairman of ICA Working Group "Cartography on Early Warning and Crises Management". His research interests include the role of geoinformatics and cartography in Early Warning and Crises Management, dynamic geovisualization, creation of SDI in Europe (he is expert of INSPIRE and GMES), design of Information and Knowledge-based Society. He is a co-author of first GIS book in Europe in 1985 and editor of EW and CM book in Springer Verlag (.in 2010).