Usability Evaluation Methods for Spatial Information Visualisation
Case Study: Evaluation of Tourist Maps

Pavel Sedlák, Jitka Komárková, Miloslav Hub, Stanislav Struška and Miroslav Pásler
Institute of System Engineering and Informatics, Faculty of Economics and Administration, University of Pardubice, Studentska 95, 532 10, Pardubice, Czech Republic

Keywords: Usability, Usability Testing, Evaluation of Maps, Maps.

Abstract: Many decisions are influenced by location. Geoinformation technologies together with digital data are today very often used to support spatially-oriented decisions. Another reliable way of spatial information presentation is represented by analogue maps. This contribution describes utilization of software engineering methods in cartography to evaluate and improve quality of maps. Authors have long experience with utilization of usability evaluation of Web based geographic information systems. They propose utilization of suitable methods to evaluate usability of analogue maps. Usability of analogue tourist maps of attractive areas of the Czech Republic was evaluated by means of proposed methods. Maps published by the most famous publishers in the Czech Republics, i. e. maps published by Kartografie Praha, a. s., SHOCart, spol. s r. o., Klub českých turistů o. s. and Geodézie On Line were evaluated. Usability User Testing and Heuristic Evaluation were used as methods for usability evaluation. The main results of case studies are briefly described in the paper. Results of one case study are processed by multi-criteria decision making methods. Benefits and weaknesses of used methods derived from author experience are stated in the end.

1 INTRODUCTION

Spatial information is very important for all human beings because thinks happen somewhere and objects are located somewhere. In the past, various analogue ways for storing spatial data were invented. Bones, stones and paper can be given as examples of used media. Nowadays, computer-based systems, like geographic information systems (GIS), are often used. But papers maps are still used too.

Usability of Web-based GIS applications and interactive maps on Internet has been recognized as a very important issue. Many articles describing case studies, usability evaluation procedures and usability problems can be found, e.g. on Web of Science. On the other side, it is difficult to find similar studies for paper maps. It is stated by Nivala et al. (2007) that there is still not enough attention paid to the usability of analogue maps, although they are irreplaceable in some situations, e.g. in crisis management or in situations when electricity and information systems are not available.

Authors carried out several case studies focused on usability evaluation of tourist analogue (i.e. paper) maps during several past years (Flamik et al., 2013, Hub et al., 2012, Sedlak et al., 2010). They introduced an idea of utilization of usability and software engineering methods into maps production and their quality evaluation. Used methods are briefly described at first. Then, overview of results is provided. It is followed by comparison of used methods and obtained experience.

2 USABILITY OF ANALOGUE MAPS

Analogue maps are an important information resource which enables visualization of various places and plays an important role in the cartographic area because maps are still not replaceable in some situations (Čapek and Komárková, 2009). Unfortunately all maps do not have the same quality, some maps are better and some worse from the quality point of view. For example, some maps contain a large amount of unnecessary information; some maps use a different scale, coordinate system or different symbolism. But users of these maps want to choose the right map that completely satisfies their needs.

Sedlák P., Komárková J., Hub M., Struška S. and Pásler M..
Usability Evaluation Methods for Spatial Information Visualisation - Case Study: Evaluation of Tourist Maps.
DOI: 10.5220/0005557904190425
In *Proceedings of the 10th International Conference on Software Engineering and Applications* (ICSOFT-EA-2015), pages 419-425
ISBN: 978-989-758-114-4

Quality of a product includes several various characteristics; usability is one of them. It is related to every product that has a user interface. Map can be though as a user interface by itself. Usability of a user interface is in accordance with ISO (1991) a set of attributes. Current usability definition selects certain usability attributes and is stated by (Bevan and Kirakowski, 1999, ISO, 1991, Law and Hvannberg, 2002) as: 'The effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment'. Next current definition of usability can be found in Ivory (2001): 'Usability is the extent to which users can use a computer system to achieve specified goals effectively and efficiently while promoting feelings of satisfaction in a given context of use'.

Usability engineering is a branch of software engineering. It provides structured methods for achieving usability in user interface design during whole product lifecycle. The usability evaluation methods are divided as follows (Scholtz, 2004):

- User-centered evaluations (usability testing methods)
- Expert-based evaluations (inspection methods)
- Model-based evaluations.

In the field of usability engineering there are many different methods available, such as Cognitive Walkthrough, Feature Inspection, Thinking Aloud Protocol, Field Observation/Ethnography, Coaching Method, Co-Discovery Learning, Retrospective Testing, Individual Interview, Performance Measurement (all Nielsen, 1994), Heuristic Evaluation (Nielsen 2007), Eye-Tracking, Question-Asking Protocol, Questionnaires (Hom, 2003), Remote Testing, Focus Groups, Logging Actual Use (all Usability Evaluation, 2013), and User Testing (Usability.gov, 2007).

As it was mentioned before, Nivala (2007) stated that there was almost no attention paid to usability of analogue maps. This study carried out in Finland used semi-structured interviews with companies that produce maps, map applications and GIS. They realized that several companies included representatives of users into the late development stages, several companies did not include them at all. Carto-graphic evaluation was carried out for paper maps only. Usability testing was rarely included at all, mostly for various GIS applications. Lately, several studies have been carried out using eye-tracking to evaluate usability of both analogue maps and graphical outputs from GIS (Brychtová et al, 2012a, Brychtová et al, 2012b, Popelka and Voženílek, 2012).

3 CASE STUDIES

When selecting appropriate usability testing and evaluation method for case studies, the knowledge gained from Budinská (2009) was used by the authors. In addition, a questionnaire survey where usability engineering experts were asked about important criteria for appropriate method selection was carried out. The following criteria are proposed by the authors to be taken into account when choosing usability testing and evaluation method for analogue maps:

- *Development stage*. There are usability evaluation methods available for each stage of system development life cycle. Only those methods that apply to evaluation of the final map products are suitable.
- *Place of testing*. Some methods allow remote testing which needs specialized software. Some methods require observation of a user in his/her real environment. These special requirements can increase costs and bring several complications. Methods which require testing in a simple testing room or which do not require any special places are more suitable.
- *Type of output data*. This issue is deeply connected to the aim of each usability evaluation and testing study. Quantitative (e.g. usability problems identification to improve the application) and/or qualitative (e.g. comparison of applications to support choice of the best one) data can be obtained from experimental measurements. Methods providing both quantitative and qualitative data are more suitable.
- *Number of participants*. It is important to involve at least one participant – a representative of users. It allows to test and verify an ability of real users to interact with evaluated maps. This step requires a clear description of potential users and correct choice of their representatives. On the other side, a high number of involved participants can increase costs.
- *Number of experts on usability testing*. It is important to involve at least one usability expert too to assure a quality of usability evaluation and testing process. On the other side, a high number of involved participants can increase costs.

Some of usability testing and evaluation methods provide only qualitative or quantitative outputs, some methods require special equipment, the presence of large number of experiments, etc. Case studies were primarily focused on quantitative data

and both experts and representatives of users should be involved. Based on the literature research we chose the methods of Usability User Testing and Heuristic Evaluation that meet the criteria formulated above. Both are experimental methods run in a laboratory environment.

3.1 Case Study – Heuristic Evaluation

According to Scholtz (2004), Heuristic Evaluation belongs to expert-based evaluations, i.e. inspection methods. Recognized usability principles are represented by a set of heuristics. Evaluators inspect whether an evaluated product meets particular heuristics or not.

Within the study Flamik et al. (2013) three analogue tourist maps were evaluated and 5 evaluators took part. A set of 109 heuristics was proposed in the beginning. Heuristics covered the following issues: a) technical implementation; b) content of a map – its completeness and visualisation; c) up-to-dateness of the content; d) readability and aesthetics of a map; e) geometric precision and conformity with the reality; f) help and additional and descriptive informations. Next, a form for evaluators was created to allow them to write their evaluations, i.e. if each particular heuristics was met or not.

After evaluation, all evaluators assessed severity of each heuristics (range from 0 to 4, where 4 means the severest usability problem). Normalized weights were calculated for all heuristics. Then, weighted scores were calculated for all evaluated products.

Evaluation resulted into identification almost 20 serious usability problems. Problems with readability and complicated orientation in maps were identified. Besides, not correct title of a map (it did not contain date) and not enough proper location of the date of thematic content were identified as usability problems too although they did not significantly obstruct users to get expected answers.

Table 1: Results of Heuristic evaluation (Flamik et al., 2013).

	Product A KP	Product B KCT	Product C SC
Evaluator 1	0.11350	0.13607	0.16806
Evaluator 2	0.16541	0.18109	0.19648
Evaluator 3	0.11103	0.15231	0.21203
Evaluator 4	0.08414	0.10472	0.18018
Evaluator 5	0.14630	0.18274	0.12436
Average value	0.12408	0.15139	0.17622

3.2 Case Study – Usability User Testing 1

Usability User Testing belongs to user-centered evaluations (Scholtz, 2004). It is based on observation of representatives of users while using an evaluated product to complete given tasks.

Within case study Sedlak et al. (2010) three analogue tourist maps were tested and 6 participants were involved. All the participants were required to evaluate all the maps. The order of evaluation was different to prevent participants from learning how to use evaluated map products.

A simple testing room was used. It was equipped by a web camera, camera, computer and software CamStudio 2.0 was installed in the computer. The proposed testing scenario contained set of particular tasks, e.g. identification of an object in the map, measurement of a distance, identification of coordinates, etc. All participants were required to fulfil the given tasks, i.e. to loudly state the answers. Simultaneously, there was a set of criteria proposed to measure participants' efficiency and effectiveness, e.g. time necessary to prepare map, time to provide answer, precision of answer. As the last step, participants filled a short questionnaire to provide some information about their background and skills concerning maps.

Testing was followed by collected data processing. Loudly stated usability problems were compared with the measured efficiency and effectiveness data to confirm an existence of the usability problem.

Again, more complicated orientation in a map caused by incorrectly placed map elements and readability of a map were the most often identified usability.

3.3 Case Study – Usability User Testing 2

The second Usability User Testing study includes again three tourist maps. All maps cover the same are of interest (Luzicke Mountains); the same coordinate system (WGS 84) is used; their producers are: Geodezie On Line, Kartografie Praha and SHOCart. All maps are folded.

Both performance data and subjective evaluation of maps by participants are collected. This case study proposes utilization of multi-criteria evaluation method to more objectively evaluate final quality of maps. It means that various criteria are considered to choose the best option from the given set of potential options. The options are three

analogue maps and the goal of multi-criteria evaluation is to determine the best map. Values of particular criteria are measured as performance and subjective data using Usability User Testing.

The overall result is based on the weighted sum of the criteria. The most important issue is to determine weights of particular criteria. It is necessary to notice that different groups of assessors could prefer different criteria and different weights and therefore there can be a conflict of opinions. A number of methods based on simple subjective information acquisition from users to finally construct estimation of the weights exist. All of these methods are based on the principle that the sum of the weights over all criteria is one.

To set weights of criteria, participants fulfilled the questionnaire after testing. Rating method (unconstrained) is preferred to point allocation method because it is easier for participants. The rating scale 1 – 10 was used in this case.

The following criteria are proposed to be used for maps evaluation (see Table for their final weights; *min* = minimizing criterion):

C1: Time of unfolding map and its preparation, *min*
C2: Time of identification of basic map elements, *min*
C3: Successful finishing of given tasks, *max*
C4: Deviation of measured value from the correct value, *min*
C5: Time necessary for solving given task, *min*
C6: Time necessary for estimation of required information/value, *min*
C7: Questionnaire – cartographic evaluation, *min*
C8: Questionnaire – usability evaluation, *min*

Table 2: Criteria and their weights (Struska, 2014).

Kind of data	Weight of group	Crite-rion	Weight of criterion	Final weight
Perfor-mance	0,568	C1	0.038	**0.022**
		C2	0.109	**0.062**
		C3	0.214	**0.122**
		C4	0.200	**0.114**
		C5	0.241	**0.137**
		C6	0.198	**0.112**
Subjec-tive	0,432	C7	0.427	**0.243**
		C8	0.573	**0.326**

Each participant evaluates all maps; different order of maps is used. At first partial benefit is calculated for each map and each participant using previously calculated weights (see Figure 1).

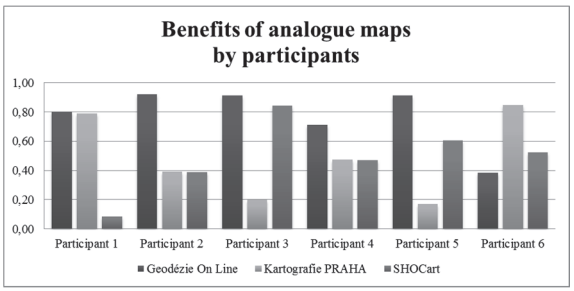

Figure 1: Partial benefits (Struska, 2014).

The equation (1) is used to calculate the benefit (Ramík, 1999):

$$U_j = \sum_{i=1}^{n} u_i\left(x_{i,j}\right) \cdot v_{i,k} \tag{1}$$

The final benefit is calculated by means of calculation of average of partial values. The final results follow (Struska, 2014):

1. Geodézie On Line – 0,775
2. SHOCart – 0,487
3. Kartografie PRAHA – 0,480

Size of fonts and not enough contrasting colours were identified as the most important usability problems of the winning map. Legend, scale bar and colours were problems of the other maps.

4 DISCUSSION OF RESULTS

The following text is based on authors' experience obtained during the previously carried out experiments – usability testing and evaluation of analogue maps (Flamik et al., 2013, Hub et al., 2012, Sedlak et al., 2010) and Web-based GIS (Komarkova et al., 2007a, Komarkova et al., 2007b, Komarkova et al., 2010a, Komarkova et al., 2010b). All the experiments were primarily focused on qualitative results (identification of usability problems). In some cases, quantitative data were collected and evaluated too.

4.1 Heuristic Evaluation

Heuristic Evaluation involves experts in usability. A combination of both experts in usability and cartography is the most suitable option to ensure correct proposal of a set of heuristics although it can increase costs.

In the beginning, it is very important to identify potential users of an evaluated map and describe them – create their profile (e.g. by means of

developing so called 'personas'). Profiles should describe users themselves (e.g. age, education, skill), and their environment (e.g. planning at home, using a map in terrain, car navigation or walking navigation, etc.). In the case of Web-based GIS and other Internet or mobile applications, technical equipment must be described too. Conceptual models of potential tasks solved by users by means of an evaluated map must be correctly prepared in the very beginning too. This information later helps to propose a suitable set of heuristics so it correctness is very important.

Experts have to get familiar with the evaluated map, its potential users and their ways of utilization of the map at first to be able to propose an appropriate set of heuristics.

A set of heuristics must be proposed and verified by experts for each particular evaluation according to the expected results, evaluated product and its potential users. This step is time consuming.

Although the set of heuristics is proposed by an expert (experts), there is a risk that some heuristics can foist usability problems on evaluators.

One evaluator is able to identify approximately 1/3 of problems. It is suitable to involve more evaluators although it increased costs as far as evaluators should be experienced.

Identification of many "cosmetic" usability problems is typical for this method (in comparison with usability User Testing). The next step is to set level of importance of usability problems to decide which problems should be focused at first. Again, involving of experts is required. Multi-criteria decision-making methods are suitable for this step.

Method itself does not require any specific testing laboratory and software.

It is not so expensive method in comparison to the Usability User Testing.

4.2 Usability User Testing

Contrary to Heuristic Evaluation, Usability User Testing method involves representatives of users of an evaluated product into the process of usability evaluation and testing.

Due to the participation of representatives of users, this method is focused on users' cognitive processes and their working memory content during utilization of an evaluated map. On the other side, obtained results are subjective points of view of involved participants.

This method requires well defined users' profile too. There is no difference in this point between Heuristics Evaluation and User Testing.

Set of experimental tasks (scenarios for testing) must be well proposed so it covers possible ways of utilization of the evaluated product in a reality as close as it is possible. It requires deep understanding of potential users and their behaviour. This part is time consuming.

Choice of representatives of users is another demanding step. A representative sample should be chosen based on previously defined users' profiles. This step highly depends on an availability of representatives of users.

This method requires a dedicated testing room but it is easy and reasonably cheap to prepare the simplest versions of it. More sophisticated testing rooms can include semi-transparent window and a complex camera system so participating persons are not disturbed by evaluators. In this case, the room is quite expensive.

Verbal reports, videos and written reports are outputs of this method. Especially think aloud procedure produces a high amount of recorded data which require further demanding processing. Possibility to measure efficiency of users' utilization of the evaluated product is a very important benefit. It is possible to measure how to which level participating persons are able to fulfil the given tasks and how quickly they are able to fulfil them.

Less number of usability problems is identified by means of the method but more severe ones are identified. It helps to focus on the severe problems which obstruct users while using the evaluated map.

5 CONCLUSIONS

Usability engineering methods have been widely used in many branches to improve quality of a product. Usability engineering itself is deeply connected to software engineering. It means its principles have been successfully applied on Web-based geographic information systems and interactive maps on Internet. Analogue maps provide spatial information as well but there was not so strong focus on their usability. Authors have focused on Web-based GIS usability testing and evaluation. Lately, they started to focus on usability of analogue maps too. Utilization of software engineering methods in maps production is reasonable as far as maps are today produced by means of specialized software tools.

After defining aim of usability testing and evaluation procedure, it is important to choose a proper usability testing and evaluation method. A chosen method can significantly influence obtained

results, e.g. if results are quantitative or qualitative, number of identified usability problems and their severity. Today, many methods exist. Different methods can be used in different product development stages and they require different tools, equipment, spaces and participating persons. All these facts can influence obtained results and, consequently, costs connected to the procedure of usability testing and evaluation.

Authors' approach: a combination of Heuristic Evaluation and User Testing allows to involve both experts and representatives of users. These methods require existence of a user interface to be evaluated, so they are suitable for analogue maps. There are several strengths and weaknesses connected to each method which are described in the text.

For future authors plan to improve Heuristic Evaluation by means of artificial intelligence methods, namely fuzzy logic, to make evaluation easier for evaluators letting them use their natural language during evaluation.

ACKNOWLEDGEMENTS

The paper has been completed with the kind support of SGS project of FEA, University of Pardubice.

REFERENCES

Bevan, N., Kirakowski, J., 1999. What is Usability? In: *Proceedings of the 4th International Conference on Human Computer Interaction*, pp 651-655.

Budinská, I., 2009. *Klasifikace a porovnání metod testování a hodnocení použitelnosti software*. Pardubice. Diploma thesis, University of Pardubice. Available in Czech only.

Brychtová, A., Popelka, S., Dobešová, Z., 2012a. Eye-tracking methods for investigation of cartographic principles. In: *SGEM 2012 Conference Proceedings* Volume II STEF92 Technology Ltd., Sofia, Bulgaria, pp 1041-1048.

Brychtová, A., Popelka, S., Voženílek, V., 2012b. The analysis of eye movement as a tool for evaluation of maps and graphical outputs from GIS. In: *Geography and Geoinformatics: Challenges for practise and education*, Masarykova univerzita, Brno, pp 154-163.

Čapek, J., Komárková, J., 2009. Windstorms and flood risk management. In: *CERS 2009 - 3rd Central European Conference in Regional Science: International Conference Proceedings*. Technická univerzita v Košiciach, Košice, pp 151-159.

Flamík, M. et al., 2013. Využití heuristického hodnocení pro zjednodušení systému hodnocení turistických analogových map. In: Křupka J, Vítek M (Eds.) *Systémy složité a zjednodušené*, 1. ed, Vydavatelství Univerzity Pardubice, Pardubice, pp 156-166.

Hom, J., 2003. *The Usability Methods Toolbox*. http://usability.jameshom.com/. Accessed 13 Dec 2013.

Hub, M., Sedlák, P., Kožuriková, M., 2012. Testing of Quality of Cartographic Products. In: *Latest trends in information technology*. Vienna, Austria, 10. - 12. 11. 2012, pp 209-214.

International Standards Organisation (ISO), 1991. *International Standard ISO 9126. Information technology: Software product evaluation: Quality characteristics and guidelines for their use*.

Ivory, M. Y., 2001. *An Empirical Foundation for Automated Web Interface Evaluation*. PhD Dissertation, UC Berkeley Computer Science Division.

Law, L., Hvannberg, E.T., 2002. Complementarity and convergence of heuristic evaluation and usability test: a case study of universal brokerage platform. In: *NordiCHI '02: Proceedings of the second Nordic conference on Human-computer interaction*. Aarhus, Denmark: ACM, pp 71-80.

Nielsen, J., 1994. *Usability inspection methods*. John Wiley & Sons, New York.

Komarkova, J. et al. 2007a. Usability of GeoWeb Sites: Case Study of Czech Regional Authorities Web Sites. In: Abramowicz W (ed.) BIS 2007, *Lecture notes in computer science*, vol 4439. Springer, Heidelberg, pp. 411–423.

Komarkova, J., Visek, O., Novak, M., 2007b. Heuristic Evaluation of Usability of GeoWeb Sites. In: J.M. Ware and G.E. Taylor (Eds.): W2GIS 2007, *Lecture notes in computer science*. vol 4857. Springer, Heidelberg, pp. 264–278.

Komárková, J., Jedlička, M., Hub, M., 2010a. Usability User Testing of Selected Web-based GIS Applications. *WSEAS Transactions on Computers* 9(1):21-30.

Komárková, J. et al, 2010b. Problems in usability of web-based GIS. In: *International Conference on Applied Computer Science - Proceedings*, pp 549-554.

Nielsen, J., 1994. *Usability Engineering*. Morgan Kaufmann, San Francisco.

Nielsen, J., 2007. *How to Conduct a Heuristic Evaluation* http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/. Accessed 08 Dec 2013.

Nivala, A. M., Sarjakostki, T., Sarjakoski., T., 2007. Usability methods' familiarity among map applications developers. *Int. J. Human-Computer Studies* 65:784-795.

Popelka, S., Voženílek, V., 2012. Specifying of Requirements for Spatio-Temporal Data in Map by Eye-Tracking and Space-Time-Cube. In: *Proceedings of International Conference on Graphic and Image Processing Singapore*, 5 p.

Ramík, J., 1999. *Vícekriteriální rozhodování: analytický hierarchický proces*. Vyd. 1. Karviná: Slezská univerzita v Opavě. 216 p. Available in Czech only.

Sedlák, P. et al, 2010. Nový přístup k testování a hodnocení kvality map. *Geodetický a kartografický obzor* 56/98(9):182-188. Available in Czech only.

Scholtz, J., 2004. *Usability evaluation*. Gaithersburg, USA: IAD National Institute of Standards and Technology, Encyclopedia of Human-Computer Interaction.

Struška, S., 2014. *Testování a hodnocení použitelnosti turistických analogových map vybrané příhraniční oblasti.* Pardubice: Univerzita Pardubice, bachelor thesis. Available in Czech only.

Usability Evaluation, 2013. *Usability Evaluation methods*. http://www.usabilityhome.com/. Accessed 12 Dec 2013.

Usability.gov, 2007. *Learn about usability testing*, http://www.usability.gov/refine/learnusa.html. Accessed 25 Mar 2010.