# COMPARISON OF USABILITY EVALUATION OF PUBLIC ADMINISTRATION WEBPAGES BY USER TESTING AND BY ANALYTICAL MODELS

**Miloslav Hub, Barbora Musilová**

***Abstract:*** *The usability of user interfaces especially of webpages is currently a very hot topic. The existing scientific literature contains various methods for testing and evaluation of the usability of various user interfaces. The benefit of various usability evaluation methods is described in this article, especially evaluation of usability by models and user testing. The most commonly used methods for these purposes are usability testing and heuristic evaluation. However, there are also suggestions of usability evaluation through analytical models that allow us to evaluate usability without having to involve end-users or experts in this evaluation process. These methods and their applicability are described in the article as well. For the purposes of the research as representative websites of public administration were chosen websites of selected municipalities. Next, the paper deals with evaluation of the web pages of selected statutory cities usability where the usability is evaluated by both suggested and validated analytical model and usability testing. The main goal is to compare the results of usability evaluation by analytical models and usability testing that use real end users and to check the analytical model's ability to act like in a real situation.*

***Keywords:*** *Analytical Models, Usability, Usability engineering, Usability evaluation, User interfaces, User testing.*

***JEL Classification:*** *C88, H83, L86.*

## Introduction

Usability is an approach to product development which includes user feedback within the development cycle in order to achieve cost reduction and to create such products and tools that comply with user's needs [15]. Usability of user interface as a accessibility [13] is important for both commercial [11] and public administration [14] software. The term usability is built on five attributes:

1. learnability - how difficult it is to learn,
2. errors - minimum errors when working with a system,
3. efficiency - achieve the highest possible efficiency,
4. satisfaction - user likes to work with the system,
5. memorability - ability to work with the system even after a longer period of not using it [12].

The goal of usability evaluation is to specify the requirements on user interface, to identify problems connected to usability or to improve the user interface [6]. Another term related to usability is usability engineering which represents so called usability evaluation within the product life cycle [12].

Usability might be tested in several ways; this article will focus on two of them - user testing and testing with the use of models with minimum utilization of users [6]. The main goal of this article is to compare these methods and the results achieved by using models and users. The goal of this comparison is to find out how much the methods using models differ from real user behaviour; the real user will be represented by an average user consisting of a few cooperating participants.

# 1 Statement of a problem

## 1.1 Comparison definition

The goal of this paper is to introduce the results of comparing two methods targeted on usability evaluation. Both methods will be used to evaluate web sites of five statutory cities on five tasks. The tasks will be the same for both methods. The comparison will consist of the most basic approach to usability evaluation - user testing and analytic modelling techniques which are not often used for usability testing.

## 1.2 Further process definition

The first step will be creating a research of already existing models which focus on usability evaluation; one model out of existing seven will be selected based on comparing models features. The selected model will be used to model created tasks and to evaluate the usability of statutory cities web sites. In the next step, the same web sites usability will be evaluated by using already existing user testing.

The results of both methods, especially time necessary to carry out presented tasks for individual statutory cities, will be compared and future process will be decided.

# 2 Methods

## 2.1 Methods for usability evaluation based on user testing

Usability testing with real users is a basic method of testing as stated in [12]. It is an irreplaceable method which provides direct information about how users work with computers. It is important to clarify the reason of testing - determining correct and wrong aspects of the interface together with interface improvement or evaluation of the overall interface quality [12]. Further, it is necessary to select a corresponding sample of users with the right knowledge, abilities etc. and a place for the testing.

User testing is a method based on monitoring the user during task fulfilment using particular hardware and software. From a certain viewpoint it is an irreplaceable method which is only complemented by the other methods. Usability testing consists of five phases: planning, trial test, live test, control test, results interpretation and presentation [12]. User testing offers a whole scale of options to scrutinize usability. It is possible to monitor participants' behaviour in their regular environment or in a carefully selected laboratory. Task fulfilment by participants may be monitored from a great distance using software which enables the participant and tester to be at different places; or they may be at the same place. The activity of participants is recorded for further research. Different types of interviews and sessions are organized with the participants to uncover their viewpoints, approaches, needs and reactions [6].

### 2.2 Usability evaluation methods based on models

Analytical usability evaluation represents using models created by an analyst. This method excludes utilization of users or experts for direct usability evaluation. Uncertainty can be modelled for example by a fuzzy apparatus [9]. Usability is evaluated only based on models describing user behaviour in selected environment. Modelling methods often enable predicting usability at low cost. They are usually used to complement other methods, e.g. user testing [6]. This type of usability testing, unlike user testing, does not have problems with slow speed as well as high time and financial requirements, although this might not be true in all cases. The methods of user and expert testing are limited by time, costs and user selection.

Usability evaluation might use one of the following models selected in accordance with the particular problem, available knowledge and available time and financial resources:

- Design analysis [6]
- Task environment analysis [10]
- Knowledge analysis [6]
- GOMS analysis [7]
- Cognitive task analysis [3]
- Programmable user models [1]
- UIDE analysis [5]

## 3 Statutory cities web sites usability

### 3.1 Goals of models

The main goal of the whole project described in this paper was evaluation of statutory cities web sites usability first using models and then using participants. In order to be able to use models, it is necessary to know clear goals which the models should reach, in other words the knowledge of concrete processes and information that users carry out and search at the websites. This whole process or path to the sought information might by described by a model in detail. Feasibility of the goal of sought information is a crucial element. Therefore, before creation of the model, cities were sent an email asking them to supply information on real traffic of users at their websites. The cooperating cities include Hradec Kralove, Olomouc and Pardubice, the information from these cities served to create 5 tasks for usability evaluation by both methods. The list of statutory cities was completed by adding Brno and Plzen.

Example of a created task:

Task 2 - find available jobs.

### 3.2 Usability evaluation – models

Comparison of descriptions and features of models found within the research led to selecting the NGOMSL model belonging to GOMS modules [6] (GOMS means “a set of Goals, a set of Operators, a set of Methods for achieving the goals, and a set of Selections rules for choosing among competing methods for goals” [2]). These models examine usability especially from the point of prediction. They deal e.g. with determining required

time for task accomplishment by the user, which holds true also for NGOMSL. NGOMSL uses external operators to determine task duration; the operators are taken from a model of lower level: KLM-GOMS. These external operators are actually keyboard operators providing information on time duration of pressing key by an average-experienced user, see Tab 1. Further it works with duration of waiting for system response and number of declarations and mental operators.

***Tab. 1: Time necessary for keyboard typing - average user***

| Typing on keyboard | Words per minute | Time [s] |
|---|---|---|
| Average user | 55 | 0.20 |

*Source: [4]*

NGOMSL model was used to record all 5 tasks for each statutory city. The record represents an example of a created model. This represents so called rule of selection which represents a signpost for further progress in the particular model. The example below under the selection rules represents the lowest step of hierarchical decomposition of solved task where keyboard operators are used with an assigned time of processing.

Example of a created model:

```
Rules of selection for goal: identify key words on home page.
If one of the key words exists on the screen, carry out the
goal: work with 1 key word.
If no key word is on the screen, carry out the goal: work
with array.
Return with accomplished goal.

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

Method for goal: Select 1 key word
Step 1. Determine stating position 1 key word     (1.20 s)
Step 2. Decide: If hands are not on mouse,
        move them onto mouse                      (0.40 s)
Step 3. Move cursor at the beginning
        of 1 key word                             (1.10 s)
Step 4. Click mouse button.                       (0.10 s)
Step 5. Return with accomplished goal
```

In order to determine task duration for all 25 models, following steps are followed [8]:

NGOMSL declarations time = number of processed declarations x 0.1 seconds,

primitive operators time = total time,

analyst defined operators time [4] = total time,

waiting time = total time when user is waiting for system response.

***Tab. 2: Time of tasks processing***

| Time to process task [s] | Pardubice | Hradec Králové | Olomouc | Brno | Plzeň |
|---|---|---|---|---|---|
| Available jobs | 20.30 | 13.00 | 12.40 | 7.30 | 9.00 |
| Dog age | 56.40 | 17.70 | 25.60 | 66.60 | 1000.00 |
| Driving license | 24.80 | 19.80 | 24.60 | 21.10 | 44.50 |
| Online reservation | 21.10 | 18.80 | 20.10 | 20.50 | 25.20 |
| Magistrate department | 37.90 | 26.60 | 19.60 | 26.90 | 15.70 |
| Total time | 160.50 | 95.90 | 102.30 | 142.40 | 1094.40 |

*Source: Authors*

Table 2 represents total times of task processing. The duration depends on complexity of goal accomplishment; only the time for statutory city Plzen needed to find the dog age threshold for paying dog fee is extreme compared to the others. Process of searching this information at Plzen website is difficult and depends on knowledge of further clues.

The results of usability show the winner is Hradec Kralove which achieves shorter time of tasks accomplishment in comparison to the other 4 cities.

### 3.3 Usability evaluation – user testing

Seven participants called Subject 1 to Subject 7 (five women, two men) were invited for testing. The age ranges between 24 and 69. All participants took part in a trial user testing, as the website of statutory city Pardubice had been used by minority of them, which even more the case for other websites of other cities. Trial user testing included providing a paper with key words which helped the participants in accomplishing the given tasks. The final user testing ran with no help to the participants and was recorder using Camtasia Studio 8 software.

All durations of task accomplishing were recorded and compared with the times obtained using models, see table 3 for a shortened overview.

***Tab. 3: Subject_1 – Comparison of testing time and model time***

| Time to process tasks [s] | Pardubice | Hradec Králové | Olomouc | Brno | Plzeň |
|---|---|---|---|---|---|
| Available jobs – subject | 18.06 | 7.26 | 8.17 | 4.23 | 16.00 |
| Deviation available jobs | 2.24 | 5.74 | 2.83 | 3.07 | -7.00 |
| Dog age – subject | 67.15 | 16.12 | 13.17 | 158.18 | 1000.00 |
| Deviation dog age | -10.75 | 1.58 | 12.43 | -91.48 | 0.00 |
| | | | | | |
| Total time[s] – subject | 153.32 | 73.80 | 78.73 | 216.59 | 1109.34 |
| Total time [s] – model | 160.50 | 95.90 | 100.90 | 142.50 | 1094.40 |
| Deviation [s] | 7.18 | 22.10 | 22.17 | -74.09 | -14.94 |

*Source: Authors*

Each of the seven users differs in behaviour during task fulfilment from the other users; therefore, an average user was created by calculating the average time of recorded values for users. The total time of this user was compared with the time obtained by the model and a total deviation of -517.63 seconds was found. This means that the model works app 8 minutes faster compared to an average user; such a deviation is not acceptable, therefore, the model needs to be adjusted to match the user better.

### 3.4 Model modification

With such a large deviation of the model from the user, it was necessary to modify the model, more accurately to modify the model time. This means mostly modification of mental operators which represent user's thinking. It is assumed to be the greatest problem. The time of keyboard operators was not changed as these have been proven in practice many times and their values are considered correct.

By repeated work with user records with websites their "reaction" times to determine starting positions were obtained. The times were enquired at some tasks, in total 21 times were collected for each subject. Further, an average time of 10.70 seconds was generated which increased the time of the new model in comparison with the old model significantly. As this new time was not created from the full set of tasks, therefore, it need not correspond with the reality, experimental determination of real time was carried out. The task is to find a greatest similarity of the model with time of an average user. The first step was to decrease the time of 10.7 seconds to a half. The course of the experiment is recorder in table 4; the new deviation represents the difference between total model time and average user time.

***Tab. 4: Experiment – finding real time***

| New deviation | Time [s] |
|---|---|
| -56.43 | 5.40 |
| -45.53 | 5.50 |
| -12.83 | 5.80 |
| 1.93 | 5.90 |
| -8.97 | 6.00 |

*Source: Authors*

## Conclusion

The paper dealt with comparison of user testing methods and methods working with models. Both of these methods are used to evaluate usability, each of them uses different means. In order to compare these methods, a research was carried out to map existing models used for usability evaluation. Feature and process study of these models led to selecting NGOMSL. This model was used to model five tasks and evaluate usability of five statutory cities websites. The results were tasks processing durations, the sum of which was app. 1600 seconds. Further part of the process was to carry out user testing with the use of sever participants, five women and two men. The result of user testing is so called average user which can process all five tasks at all web sites within app 2100 seconds. The comparison of results showed deviance of app. 520 seconds, therefore, the model did not reflect the reality. Repeated review of user testing results and experimental definition of mental operators' duration led to achieve deviation of only app. 2 seconds. In conclusion, it is possible to state that usability evaluation by models is possible, however, the results need to be verified by e.g. here used user testing which is based on behavior of real participants - users.

The next phase we would like to conduct in future is to suggest a generic methodology that could be used for every type of public administration information system user information. So we would like to focus to other types of graphical user interface for example to mobile equipment.


### Acknowledgement

This contribution was supported by project of University of Pardubice SGSFES_2015001 with the title "Economic and social development in the private and public sectors".


### References


[1] BLANDFORD, A., GOOD, J., YOUNG, R. M. Programmable User Modelling Analysis for usability evaluation. In *UCL INTERACTION CENTRE: Division of Psychology & Language Sciences and Department of Computer Science* [online]. 1998. [cit. 2015-05-11]. Available at WWW: <http://www.uclic.ucl.ac.uk/annb/docs/armpack/wp11a.pdf>

[2] CARD, S., MORAN T. P., NEWELL, A. *The Psychology of Human Computer Interaction*. Lawrence Erlbaum Associates. 1983. ISBN 0-89859-859-1.

[3] CRANDALL, B., KLEIN G. A., HOFFMAN, R. R. *Working minds: a practitioner's guide to cognitive task analysis*. Cambridge, Mass.: MIT Press, 2006, xii, 332 p.

[4] FRANK, A. U., CAMPARI, I. Spatial information theory: a theoretical basis for GIS. In *European conference, COSIT'93*. Marciana Marina, Elba Island, Italy, September 19-22, 1993: proceedings. New York: Springer-Verlag, c1993, xi, 477 p. ISBN 0387572074-. Available at WWW: <http://books.google.cz/books?id=BlWlDxRZ05MC&pg=PA409&lpg=PA409&dq=average+typist+%2855+wpm%29+pointing&source=bl&ots=HyRCRT-viL&sig=sNNj40dx6SBi9YSeR2yhU1XVh28&hl=en&sa=X&ei=OKpjVJqbFY7nata6gIAN&ved=0CCgQ6AEwAQ#v=onepage&q=average%20typist%20%2855%20wpm%29%20pointing&f=false>

[5] FOLEY, J., GIBBS, C., KOVACEVIC, S. A knowledge-based user interface management system. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '88* [online]. 1988 [quote 2015-05-11]. DOI: 10.1145/57167.57178.

[6] IVORY, M. Y. *An Empirical Foundation for Automated Web Interface Evaluation*. UC Berkeley Computer Science Division. Available also at WWW: <http://webtango.berkeley.edu/papers/thesis/>. Dissertation thesis. University of California at Berkeley. 2001.

[7] KIERAS, D. E. *EECS University of Michigan* [online]. [cit. 2015-05-11]. 2010. Available at: http://web.eecs.umich.edu/~kieras/

[8] KIERAS, D. GOMS modeling of user interfaces using NGOMSL. In *Conference companion on Human factors in computing systems - CHI '94* [online]. 1994 [cit. 2015-05-11]. DOI: 10.1145/259963.260467. Available at WWW: <http://www.idemployee.id.tue.nl/g.w.m.rauterberg/lecturenotes/GOMS96guide.pdf>

[9] KULICKA, J. Estimation of uncertainty for problem solving by fuzzy mathematic tools. In A. *Bilsel, M. U. Garip. Frontiers in Mathematics and Science Education Research. Proceedings of the Frontiers and Science Education Research Conference* 01-03 May 2014, Famagusta, North Cyprus. Famagusta, North Cyprus: Science Education Research Group at Eastern Mediterranean University, pp. 144-152. 2014.

[10] MORAN, T. P. Getting into system: ETIT analysis. In *Carnegie Mellon University: University Libraries, Digital Collections* [online]. [cit. 2015-05-11]. Available at WWW: <http://digitalcollections.library.cmu.edu/awweb/awarchive?type=file&item=359668>

[11] MYSKOVA, R. Economic Information Systems for Small and Medium Businesses and Evaluation of Return. In *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS*, Issue 3, Volume 8, March 2011, p. 119-128, ISSN: 1790-0832

[12] NIELSEN, J. *Usability Engineering*. Boston: AP Professional, 1993, 362 s.,pic., tb. ISBN 0125184069.

[13] ŠIMONOVÁ, S. E-Inclusion and disabled-people-friendly web. *Scientific Papers of the University of Pardubice Series D*. Pardubice: Univerzita Pardubice, 2006. Series D, Faculty of Economics and Administration, 10 (2006). pp. 164-168. ISBN 80–7194–851–9, ISSN 1211 – 555X.

[14] ŠIMONOVÁ, S., KOPÁČKOVÁ, H. eDocument in eGovernment. *WSEAS Transactions on Information Science and Applications*, 2010, vol. 7, no. 1, s. 92-101. ISSN: 1790-0832.

[15] TULLIS, T., ALBERT, B. *Measuring the user experience: collecting, analysing, and presenting usability metrics*. Burlington: Morgan Kaufmann, 2008. 317 s. ISBN: 978-0-12-373558-4.

## Contact Address

**doc. Ing. Miloslav Hub, Ph.D.**
University of Pardubice, Faculty of Economics and Administration, Institute of System Engineering and Informatics
Studentska 84, 532 10, Pardubice, Czech Republic
Email: miloslav.hub@upce.cz
Phone number: (+420) 466 036 071

**Ing. Barbora Musilová**
University of Pardubice, Faculty of Economics and Administration, Institute of System Engineering and Informatics
Studentska 84, 532 10, Pardubice, Czech Republic
Email: st32701@student.upce.cz

Received: 29. 02. 2016
Reviewed: 05. 07. 2016, 02. 09. 2016
Approved for publication: 08. 09. 2016