

Assessing qualitative usability in life-threatening, time-critical and unstable situations

Abstract

We developed a heuristic for assessing the usability of mobile user-interfaces in life-threatening, time-critical and unstable situations in a qualitative way. The major advantages of our approach as opposed to standardized quantitative questionnaires is the independence from a baseline, the possibility to make absolute statements and the potential for adaptations. When creating a qualitative semi structured interview we adhere to the common *modus operandi* of the qualitative social research. On the basis of 17 common quantitative questionnaires on usability we identified the five major categories **Utility**, **Intuitiveness**, **Memorability**, **Learnability** and **Personal Effect**. We selected all questions from the questionnaires which are useful for assessing the usability of user-interfaces in emergencies. Furthermore, we rephrased the closed-ended questions to open-ended ones. The quantification of research results is possible by weighting the qualitative results in dependence to the research question.

Simon Nestler¹
 Eva Artinger¹
 Tayfur Coskun¹
 Yeliz Yildirim-Krannig²
 Sandy Schumann²
 Mareike Maehler²
 Fabian Wucholt²
 Stefan Strohschneider²
 Gudrun Klinker¹

1 Technische Universität München, Fakultät für Informatik, Fachgebiet Augmented Reality, Garching bei München, Germany

2 Friedrich-Schiller-Universität Jena, Interkulturelle Wirtschaftskommunikation, Jena, Germany

1 Introduction

Within the scope of the *SpeedUp* project [The project *SpeedUp* is funded by the German Federal Ministry of Education and Research (BMBF) within the programme “Research for Civil Security” (May 1st, 2009–April 30th, 2012, FKZ: 13N10175). Website: <http://www.speedup-projekt.de/>.] we found out that the sound evaluation of mobile user-interfaces for medical emergencies is challenging for three reasons: (1) Mobile user-interfaces replace paper based workflows, (2) evaluations take place in lifelike trainings and (3) stress is dominant in medical emergencies.

Paper based workflows: When mobile user-interfaces are compared to paper based approaches, this comparison is inhomogeneous. Although mobile user-interfaces increase the quality of information [Quality of information is enhanced by increased structuring.], entering high quality information is more laborious. By comparing the usability of the mobile user-interface to the usability of paper, the mobile user-interface is in an inferior position. It is essential to provide a possibility to evaluate mobile user-interfaces without needing a baseline (see A).

Real life scenarios: Due to the fact that evaluations of mobile user-interfaces for medical emergencies take place in lifelike trainings, the repeatability of the evaluation is limited. The trainer cannot completely control the set of parameters in these scenarios. Consequently, the sequential, quantitative comparison of different design alternatives is subject to restrictions. Evaluating user-interfaces without requiring a quantitative comparison of different alternatives is essential (see B).

Dominance of stress: The impact of stress on usability is not considered by the different standardized questionnaires. Questionnaires either focus on usability or on physical and mental demands. However, because usability depends on the task load and the mental demands are high in medical emergencies, considering the impact of stress is essential for gaining meaningful results. Furthermore, weighting the different categories is essential for evaluating mobile user-interfaces (see C).

Consequently, for the proper evaluation of user-interfaces in life-threatening, time-critical and unstable situations a new type of usability evaluation is required. We propose the conduction of a qualitative semi structured interview for three reasons:

- qualitative interviews **do not depend on a baseline** (A) [The qualitative properties of user-interfaces are absolute.]
- qualitative interviews **do not require different alternatives** (B)
- qualitative interviews **allow the weighting of categories** (C)

Furthermore, qualitative information is essential for improving the capabilities of mobile user-interfaces. By providing detailed qualitative information, the engineers and computer scientists can identify the weaknesses of mobile user-interfaces more easily and can improve them more effectively and efficiently. Consequently, the application of qualitative assessments in requirements analysis, interaction design, prototypical implementation and evaluation simplifies the overall process of developing mobile user-interfaces.

2 Related work

In literature different standardized questionnaires are used for the evaluation of user-interfaces. In the following we present the 17 most common quantitative questionnaires with focus on *Usability*, *Attractiveness*, *Satisfaction*, *Experience* and *Work Load*. Some of these questionnaires categorize the different questions. All categories which are transferred to our qualitative semi structured interview are written in bold.

The *After-Scenario Questionnaire* consists of three questions on the user's **Satisfaction** [1]. The *AttrakDiff* questionnaire consists of 21 pairs of antithetic adjectives. The *AttrakDiff* focuses on **Attractiveness**, *Hedonic Quality* and *Pragmatic Quality* [2]. The *Computer Literacy Scale* consists of different questions on the user's experience with computers. The *CLS* focuses on *Experience*, *Symbols* and *Terminology* [3]. The *Computer System Usability Questionnaire* consists of 19 questions on the system's usability. The *CSUQ* is unstructured and does not use categories [4]. The *IsoMetrics* focuses on usability in general. The *IsoMetrics* focuses on *Adequacy of Tasks*, *Ability of Self-Characterization*, *Controllability*, **Compliance with Expectations**, **Error Robustness**, **Customizability** and **Learnability** [5], [6]. The *IsoMetrics* questionnaire is based on the *Isonorm 9241-10*, so the structure is quite similar. The categories are identical to *IsoMetrics* – except *Fault Tolerance* (instead of **Error Robustness**) and *Ease of Learning* (instead of **Learnability**) [7]. The *Nielsen's Attributes of Usability* consist of 5 different categories: **Learnability**, **Efficiency**, **Memorability**, **Errors** and **Subjective Satisfaction** [8]. The *Nielsen's Heuristic Evaluation* consists of 10 questions which result in a heuristic guideline. The *NHE* is unstructured and does not use categories [9], [10], [11]. The *Practical Heuristics for Usability Evaluation* consist of a heuristic guideline with 13 questions. The *PHUE* focus on **Learning**, *Adapting to the user*, **Feedback** and **Errors** [12]. The *Perceived Usefulness and Ease of Use* questionnaire consists of 12 questions on *Usefulness* and **Ease of Use** [13]. The

Purdue Usability Testing Questionnaire consists of 100 questions in different categories. The *PUTQ* focuses on *Compatibility*, **Consistency**, **Flexibility**, **Learnability**, *Minimal Action*, *Minimal Memory Load*, *Perceptual Limitation* and **User Guidance** [14]. The *Questionnaire for User Interface Satisfaction* consists of 27 questions on **Satisfaction**. The *QUIS* focuses on *Overall Reaction*, *Screen*, *Terminology*, *System Information*, **Learning** and *System Capabilities* [15]. The *Software Usability Measurement Inventory* consists of 50 questions on usability in general. The *SUMI* is unstructured and does not use categories [16]. The *System Usability Scale* consists of 10 questions on usability. The *SUS* is unstructured and does not use categories [17]. The *NASA Task Load Index* consists of 6 questions on work load. The *NASA-TLX* is unstructured and does not use categories [18]. The *User Experience Questionnaire* consists of 26 pairs of antithetic adjectives. The *UEQ* focuses on **Attractiveness**, **Perspicuity**, **Novelty**, *Stimulation* and **Dependability** [19]. The *USE Questionnaire* consists of 30 questions on general usability. The *USEQ* focuses on *Usefulness*, **Ease of Use**, **Ease of Learning** and **Satisfaction** [20].

When taking a closer look on the categories from these questionnaires it becomes clear, that these various questionnaires are not selective. Several aspects, however, such as **Stress**, **Expericence** or **User Guidance**, are only considered by one questionnaire.

3 Method

For the qualitative evaluation of user-interfaces we make use of a method which comes from social science: qualitative interviews. We combined this method with a structured literature review. According to [21] these qualitative interviews take the perspective of the subjects into consideration. Besides the parameters the underlying causes of decisions are subject of the evaluation. The used methods are legitimated by their contribution to the solution of a research question. The semi-structured interviews base on a set of problems, each of these problems consists of a set of questions. These questions are orally answered by the subject and the interview is documented on a voice recorder as described by [22], [23]. The interviewer uses open-ended questions and avoids interrupting the subject. The aim of the interview is to discuss all problems with the subject. [Usually it is not necessary to ask the complete set of questions to cover all problems.]

Although this method is quite well known and successful in social sciences, it has not found its way into usability research. Therefore, we transferred this method from social sciences to usability research. Our qualitative interview on usability was developed on the basis of an extensive brainstorming process in combination with creating a detailed associagram. During the brainstorming we made use of the quantitative questionnaires on usability. Furthermore, the categories of the quantitative questionnaires were utilized in our associagram. Due to the fact

that we started from the scratch we could design a method which is independent from a baseline (see *A*) and does not rely on different alternatives (see *B*). The adaptation of the questionnaire to various fields of application can be guaranteed by a flexible weighting of different categories (see *C*). Consequently, this method can flexibly be adapted to specific requirements.

Qualitative evaluation is not limited to an effectivity control, assessing the value of components or the overall object is of equal importance. The evaluation has to conform with critical-rational demands as well as with ethic-moral standards. In general an evaluation can have four different aims [24]: (R.1) facilitating insights, (R.2) reaching decisions, (R.3) legitimating decisions and (R.4) optimizing objects. In usability engineering the aspects (R.1) and (R.4) are of special importance during the prototyping phase.

[21] describes the general principles for preparing qualitative interviews: (1) The research question has to be concretized, (2) the questions have to be selected by a team of experts and (3) questions have to be formulated open-ended. Furthermore, the evaluation design has to consider the following questions: (D.1) How many subjects should be interviewed, (D.2) how are these subjects chosen, (D.3) when should the evaluation take place, (D.4) how are the interviews recorded and (D.5) how will the transliteration be done. The most popular literature on the number of subjects was published by [25], [26], [27] and [28]. In these publications a mathematical model is developed which enables the user-interface designer to calculate the optimal sample size: $U=1-(1-p)^n$. The probability to detect a problem (*U*) depends on the probability (*p*) that a subject identifies a problem and the number (*n*) of subjects. Due to the fact that the probability (*p*) is not known for qualitative usability evaluations, we have to rely on their general heuristics: (1) most usability problems are detected with three to five subjects, (2) it is unlikely that additional subjects reveal new information and (3) most severe usability problems are detected by the first few subjects. Consequently, we use three to five subjects for the qualitative usability evaluations (D.1). We choose this set of subjects randomly from the group of our end users (D.2). The evaluations take place in each iteration (D.3). [An iteration consists of requirements analysis, interaction design, prototypic implementation and evaluation.] The interviews are documented by voice recorders (D.4) and are manually transliterated (D.5).

After the transliteration of the interviews, the different statements are categorized according to the following rules: (1) categories are terms, (2) categories are deduced from the aims and research questions, (3) categories should neither be slender nor extensive and (4) categories have to be selective [29]. According to [24] the following requirements have to be taken into consideration in order to get methodically dependable and valid evaluation results: (1) The individual cases are part of the research process, (2) research process is open for revision and extension, (3) the general procedure is lead by a set rules,

(4) research processes are seen as an interaction, (5) the objects are analyzed holistically and (6) generalization is demonstrated by arguments. Consequently, the presented set of problems and questions is always subject of further research.

4 Results

The combination of existing quantitative questionnaires on usability and qualitative research methods leads to semi structured interviews on usability. In the following the resulting categories and questions of the qualitative interview are described. The process of generating categories and open-ended questions was performed according to the process from [21] as presented above.

4.1 Categories

On the basis of these questionnaires three usability experts conjointly identified five main categories for the qualitative interview: (**U**) **Utility**, (**J**) **Intuitiveness**, (**M**) **Memorability** [8], (**L**) **Learnability** [5], [6], [7], [8], [12], [14], [15], [20] and (**P**) **Personal Effect**. These main five categories are subdivided in four to five sub-categories each: The (**U**) **Utility** is classified in: (U-I) Dependability [5], [6], [7], [19] (U-II) Ease of Use [13], [20], (U-III) Effectiveness, (U-IV) Efficiency [8] and (U-V) Productivity. The (**J**) **Intuitiveness** is classified in: (J-I) Affordance [5], [6], [7], (J-II) Expectations [5], [6], [7], (J-III) Conventions and (J-IV) Transparency. The (**M**) **Memorability** is classified in: (M-I) Consistency [14], (M-II) Customizability [5], [6], [7], [14], (M-III) Complexity and (M-IV) Perspicuity [19]. The (**L**) **Learnability** is classified in: (L-I) Error handling [5], [6], [7], [8], [12], (L-II) Feedback [12], (L-III) Help and (L-IV) User Guidance [14]. The (**P**) **Personal Effect** is classified in: (P-I) Attractiveness [2], [19], (P-II) Novelty [19], (P-III) Satisfaction [1], [8], [15], [20] and (P-IV) Stress.

These categories fulfill the major requirements from [21]: The categories are terms, are deduced from the research question and are selective. These terms are directly taken from the different questionnaires. Due to the fact that these usability questionnaires deal with our research question, we could prove that the categories are deduced from the research question. Furthermore, the concordant identification of the five main categories by three usability experts indicates the selectiveness of these categories. The question whether these categories are neither slender nor extensive has to be proven within the scope of an evaluation. Table 1 gives an overview on all categories and sub-categories.

4.2 Questions

In the next step three usability experts conjointly assigned all questions from the qualitative usability questionnaires to the different categories. When the assignment was ambiguous, the question was removed from the qualitative semi structured interview. [An unambiguous assign-

Table 1: Categories for quantitative summarization

Utility	Intuitiveness	Memorability	Learnability	Personal Effect
Dependability	Affordance	Consistency	Error handling	Attractiveness
Ease of Use	Expectations	Customizability	Feedback	Novelty
Effectiveness	Conventions	Complexity	Help	Satisfaction
Efficiency	Transparency	Perspicuity	User Guidance	Stress
Productivity				

ment requires at least the same categorization by two of the three experts.]

The resulting semi structured interview is shown in the appendix (Attachment 1). We followed the principles from [21]: The concrete research question is **Usability**, the questions were selected by a team of three experts with regard to the research question and finally the closed-ended questions were rephrased to open-ended questions. This rephrasing is necessary to be able to use these questions as a guideline for the semi-structured interview. In each of the categories several different questions are available and the interviewer is free to choose a subset of questions which fits best for the concrete user-interface. As an alternative the team of interviewers can pre-select a subset of questions in the run-up to the evaluation.

5 Discussion

The qualitative evaluation provides a detailed assessment of the quality of a mobile user-interface. According to [30] a quantification of research results is important. A quantitative score, however, is not directly deduced by the proposed method. [24] proposes a quantitative analysis of the qualitative evaluation to receive the required quantitative data. In this analysis the transliterated statements from the interview are sorted by experts into the categories and sub-categories. This quantitative summarization of the qualitative evaluation is the basis for the quantification of the research results according to [30]. The summarized data is adjusted on a 3-point scale: (a) positive comment (1.0), (b) neutral comment (0.5) and (c) negative comment (0.0). The mean value over all statements in the same sub-categories is calculated. As a result we receive a quantitative rating of all sub-categories on a scale from 0.0 to 1.0.

After calculating these scales, an usability scale which is application specific can be calculated by weighting these categories. The categories for the quantitative summarization are shown in Table 1. The balanced weighting of the categories and sub-categories from Table 1 is shown in Table 2. When performing general evaluations of user-interfaces this weighting leads to a quantitative one-dimensional usability value – besides the qualitative results.

Table 2: Balanced weighting (general user-interfaces)

20 ^a	20 ^b	20 ^c	20 ^d	20 ^e
4	5	5	5	5
4	5	5	5	5
4	5	5	5	5
4	5	5	5	5
4				

^a Utility; ^b Intuitiveness; ^c Memorability;
^d Learnability; ^e Personal Effect

We argued that weighting the different sub-categories is essential due to the dominance of stress in emergencies. Therefore, we included a weighting of the qualitative results in dependence to the research question. For the evaluation of mobile user-interfaces for emergencies the experts changed the weighting with regard to the research question as shown in Table 3. The **Utility** and the **Personal Effect** is of special importance in emergencies [31]. Therefore, the weights for these categories were increased. **Stress** is dominant in emergencies, whereas **Attractiveness** and **Novelty** is of lower importance. Therefore, the weight for **Stress** was increased as well. From previous requirements analyses we know that **Customization** is difficult in emergencies [31], [32]. Consequently, customization is not considered in the emergency specific usability value.

Table 3: Focused weighting (user-interfaces for emergencies)

35 ^a	10 ^b	15 ^c	10 ^d	30 ^e
7	3	5	3	2
7	3	0	3	2
7	2	5	2	6
7	2	5	2	20
7				

^a Utility; ^b Intuitiveness; ^c Memorability;
^d Learnability; ^e Personal Effect

Finally, the usability score is calculated by multiplying all weights $w(s)$ with the quantitative scores $v(s)$ of the sub-categories s : see Figure 1.

$$\begin{aligned}
 U &= \sum_{c \in C} \left(U_c * \frac{w_c}{100} \right) = \sum_{c \in C} \left(U_c * \sum_{s' \in S(c)} \frac{w(s')}{100} \right) = \\
 &= \sum_{c \in C} \left(\sum_{s \in S(c)} \frac{w(s) * 100}{\sum_{s'' \in S(c)} w(s'')} * v(s) * \sum_{s' \in S(c)} \frac{w(s')}{100} \right) = \sum_{c \in C} \left(\sum_{s \in S(c)} w(s) * v(s) \right) = \sum_{s \in S} (U_s * w_s)
 \end{aligned}$$

Figure 1

Besides the general usability score U , more specific scores U_c for each category c can be calculated – as described above. Each category c consists of a set of sub-categories $S(c)$. For the calculation of these specific scores the weights w have to be normalized. The score U as well as the scores U_c go from 0 to 100, due to the fact that the sum of all weights is 100.

6 Conclusion and future work

The qualitative usability evaluation leads to a benefit for developers, because detailed qualitative information is provided. This qualitative information helps to identify the weaknesses of mobile user-interfaces more easily. Nevertheless a quantitative usability score can be provided as well and the qualitative usability evaluation has no disadvantages – but many benefits – as opposed to the quantitative usability evaluation. Furthermore, the quantitative score can flexibly be adapted to the concrete research question [in our case: mobile user-interfaces for emergencies].

We will use our qualitative interviews to get deeper insights in the human-computer interaction in emergencies (R.1, see above). Furthermore, the existing research method is iteratively improved by all these qualitative evaluations due to the fact that a re-categorization of the answers is performed subsequent to every interview. The questionnaire can be simplified by reducing the number of questions and can be customized by weighting the different categories. Due to the fact that a small set of subjects (three to five) is used in qualitative evaluations, our evaluation efforts are reduced significantly.

In the future we expect an intensive use of qualitative usability evaluations in the ubiquitous computing domain because of the following reasons:

- ubiquitous applications **are new and innovative**
- ubiquitous applications **are frequently evaluated in real-life scenarios**
- ubiquitous applications **are developed domain-specific**

When building new and innovative ubiquitous applications, the comparison with existing applications is often difficult. On the one hand innovative applications exceed the capabilities and functionalities of existing ones and on the other hand subjects are more familiar with existing applications. Consequently, these qualitative, comparative evaluations are often inhomogeneous in the ubiquitous computing domain. In our impression qualitative assessment of the attributes is more promising with regard to the effective improvement of the ubiquitous application.

When evaluating in real-life or lifelike scenarios, the comparability of successive runs is limited. When the ubiquitous application focuses on a specific domain, the usage of standardized questionnaires is complicated. Consequently, the flexible customizability of the qualitative evaluation is a strong argument for its future, more intense application.

Notes

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from

<http://www.egms.de/en/journals/mibe/2011-7/mibe000115.shtml>

1. GMS-mibe-Nestler-Att1.pdf (108 KB)
Appendix: Questionnaire

References

1. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM SIGCHI Bulletin*. 1991;23(1):78-81. DOI: 10.1145/122672.122692
2. Hassenzahl M, Burmester M, Koller F. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Ziegler J, Szwillus G, eds. *Mensch & Computer 2003: Interaktion in Bewegung*. Stuttgart: B.G. Teubner; 2003. p. 187-196.
3. Sengpiel M, Dittberner D. The computer literacy scale (CLS) for older adults – development and validation. In: Herczeg M, Kindsmüller MC, eds. *Mensch & Computer 2008: Viel Mehr Interaktion*. München: Oldenbourg Verlag; 2008. p. 7-16.
4. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum Comput Interact*. 1995;7(1):57-78. DOI: 10.1080/10447319509526110
5. Gediga G, Hamborg LC. IsoMetrics: An usability inventory supporting summative and formative evaluation of software systems. In: Bullinger HJ, Ziegler J, eds. *Human-Computer Interaction: Communication, Cooperation, and Application Design, Proceedings of HCI International '99; 1999 August 22-26; Munich, Germany*. Mahwah, NJ: Lawrence Erlbaum; 1999. p. 1018-1022.
6. Gediga G, Hamborg KC, Düntsch I. The IsoMetrics usability inventory. *Behaviour & Information Technology*. 1999;18(3):151-64. DOI: 10.1080/014492999119057

7. Prümper J. Der Benutzungsfragebogen ISONORM 9241-10: Ergebnisse zur Reliabilität und Validität. In: Software-Ergonomie '97 – Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung. Stuttgart: Teubner; 1997. p. 253-262.
8. Nielsen J. Usability engineering. Boston: AP Professional; 1993.
9. Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: Carrasco Chew J, Whiteside J, eds. Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people; 1990 April 1-5; Seattle. New York: ACM; 1990. p. 249-256.
10. Nielsen J. Enhancing the explanatory power of usability heuristics. In: Adelson B, Dumais S, Olson JS, eds. Proceedings of the SIGCHI conference on Human factors in computing systems: Celebrating Interdependence; 1994 April 24-28; Boston. New York: ACM; 1994. p. 158. DOI: 10.1145/191666.191729
11. Nielsen J. Usability inspection methods. In: Plaisant C, ed. Conference on Human factors in computing systems; 1994 April 24-28; Boston. Conference Companion. New York: ACM; 1994. p. 413-414.
12. Perlman G. Practical usability evaluation. In: Plaisant C, ed. Conference on Human factors in computing systems; 1994 April 24-28; Boston. Conference Companion. New York: ACM; 1994. p. 407-408.
13. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly. 1989;13(3):319-40. DOI: 10.2307/249008
14. Lin HX, Choong YY, Salvendy G. A proposed index of usability: a method for comparing the relative usability of different software systems. Behav Inf Technol. 1997;16(4):267-77. DOI: 10.1080/014492997119833
15. Chin JP, Diehl VA, Norman KL. Development of an instrument measuring user satisfaction of the human-computer interface. In: Soloway E, Frye D, Sheppard SB, eds. Proceedings of the ACM CHI 88 Human Factors in Computing Systems Conference; 1988 June 15-19; Washington. New York: ACM; 1988. p. 213-218.
16. Kirakowski J. The software usability measurement inventory: background and usage. In: Jordan P, Thomas B, Weerdmeester B, eds. Usability evaluation in industry. London: Taylor and Francis; 1996. p. 169-178.
17. Brooke J. SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, eds. Usability evaluation in industry. London: Taylor and Francis; 1996. p. 189-194.
18. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv Psychol. 1988;52:139-83. DOI: 10.1016/S0166-4115(08)62386-9
19. Laugwitz B, Held T, Schrepp M. Construction and Evaluation of a User Experience Questionnaire. Lect Notes Comput Sci. 2008;5298:63-76. DOI: 10.1007/978-3-540-89350-9_6
20. Lund A. Measuring Usability with the USE Questionnaire. Usability & User Experience. 2001;8(2).
21. Kuckartz U, Dresing T, Stefer C, Rädiker S. Qualitative Evaluation: der Einstieg in die Praxis. Wiesbaden: VS Verlag für Sozialwissenschaft; 2007.
22. Holtzblatt K, Jones S. Contextual inquiry: A participatory technique for system design. In: Schuler D, Namioka A, eds. Participatory design: Principles and practices. Mahwah, NJ: Lawrence Erlbaum; 1993. p. 177-210.
23. Sarodnick F, Brau H. Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung. Bern: Verlag Hans Huber; 2006.
24. Kinast EU. Evaluation interkultureller Trainings. Lengerich: Pabst; 1998.
25. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Ashlund S, Mullet K, Henderson A, Hollnagel E, White TN, eds. Human-Computer Interaction, INTERACT '93, IFIP TC13 International Conference on Human-Computer Interaction, jointly organised with ACM Conference on Human Aspects in Computing Systems CHI'93; 1993 April 24-29; Amsterdam. New York: ACM; 1993. p. 206-213.
26. Virzi RA. Refining the test phase of usability evaluation: How many subjects is enough? Hum Factors. 1992;34(4):457-68.
27. Lewis JR. Sample sizes for usability studies: Additional considerations. Hum Factors. 1994;36(2):368-78.
28. Lewis JR. Usability testing. In: Salvendy G, ed. Handbook of human factors and ergonomics. New York: Wiley; 2006. p. 1275-1316. DOI: 10.1002/0470048204.ch49
29. Mayring P, Gläser-Zikuda M, eds. Die Praxis der Qualitativen Inhaltsanalyse. 2nd ed. Weinheim: Beltz, Psychologie-Verl.-Union; 2008.
30. Mayring P. Einführung in die qualitative Sozialforschung: Eine Anleitung zu qualitativem Denken. 2nd ed. Weinheim: Beltz, Psychologie-Verl.-Union; 1993.
31. Nestler S. Design, Implementation and Evaluation of User-Interfaces for lifethreatening, time-critical and unstable Situations [dissertation]. München: Technische Universität München; 2010.
32. Nestler S. Einsatzorganisation in Katastrophen: Unterstützung von Einsatzkräften durch mobile User-Interfaces. Saarbrücken: VDM Verlag Dr. Müller; 2008.

Corresponding author:

Simon Nestler
Technische Universität München, Fakultät für Informatik,
Fachgebiet Augmented Reality, Boltzmannstraße 3,
85748 Garching bei München, Germany
simon@nestlers.de

Please cite as

Nestler S, Artinger E, Coskun T, Yildirim-Krannig Y, Schumann S, Maehler M, Wucholt F, Strohschneider S, Klinker G. Assessing qualitative usability in life-threatening, time-critical and unstable situations. GMS Med Inform Biom Epidemiol. 2011;7(1):Doc01. DOI: 10.3205/mibe000115, URN: urn:nbn:de:0183-mibe0001150

This article is freely available from

<http://www.egms.de/en/journals/mibe/2011-7/mibe000115.shtml>

Published: 2011-10-17

Copyright

©2011 Nestler et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>). You are free: to Share – to copy, distribute and transmit the work, provided the original author and source are credited.