

Specific-Domain Usability Heuristics: Are they really necessary?

Cristhy Jimenez^{1,2}, Pablo Lozada³, Pablo Rosas³

¹ Pontificia Universidad Católica de Valparaíso

² Universidad Nacional de Chimborazo

E-mail: cjimenez@unach.edu.ec

³ Escuela Superior Politécnica de Chimborazo

E-mail: {plozada, prosas}@epoch.edu.ec

Abstract. The quality of interactive systems is defined by several attributes. One of the main quality attributes is Usability. Evaluating the usability of interactive systems can help to ensure that applications will be satisfactorily accepted by users. One of the most commonly used evaluation methods is heuristic evaluation. In this method, the set of heuristics can significantly influence its performance. Although, the Nielsen's heuristics have been widely applied, the development of new specific-domain usability heuristics is increasing nowadays. The real need for providing new specific heuristics is the question that we pretend to solve through this work. This paper synthetize the results of a literature review conducted to analyze the progress in the development and use of usability heuristics for both, determining heuristics that could serve as basis for developing new ones, but also identifying activities that could be collected to establish a methodology to develop specific-domain usability heuristics. The review included studies from journals published between 2008 and 2015. It was possible to summarize the current knowledge concerning usability heuristics and identify gaps and new knowledge about this topic. Results showed the need for developing specific-domain usability heuristics and also served as basis to formalize a methodology which was named PROMETHEUS.

Keywords: Software Evaluation, Usability, Usability Heuristics, Usability Evaluation Methods, Literature Review.

1. Introduction

Since the term usability appeared almost three decades ago (1980s), several techniques and methods have been proposed for designing and evaluating interactive software systems. Currently, usability evaluation process plays a main role in human-centered design activities. Ensuring usability enables and facilitates the design of successfully accepted software systems. In fact, the benefits of usability evaluations are well established in the Human-

Computer Interaction domain. Research has shown that early detection of problems in the development of a software product can help to ensure its quality, reduce post-release service costs, and hence save money.

There are several methods for assessing the usability of interactive software systems. One of the most commonly accepted methods is the heuristic evaluation (Nielsen & Molich, 1990). Evaluators use a set of well recognized usability design principles called “heuristics” as a guide for discovering usability problems. The selected heuristics could significantly influence the results.

As literature reveals the existence of several sets of usability heuristics, a systematic literature review about this topic is necessary. The main goal is focused on to identify, evaluate, interpret and synthesize all available relevant research, regarding the progress in the development and application of usability heuristics. The identification of different sets of usability heuristics could serve as basis for developing new ones.

The paper is organized as follows. Section 1 presents the theoretical background with concepts regarding usability and usability evaluation methods. Section 2 details the methodology that we used for conducting the systematic literature review of usability heuristics. Section 3 summarizes the analysis of extracted information and the obtained results. Finally, Section 4 presents the conclusions and future work.

1.1 Usability

It was in 1985 when the term “usability” was used for first time. Gould and Lewis (1985) were the pioneers in the use of this term during the presentation of their work entitled “Designing for usability: Key Principles and What Designers Think”. The work was presented at the SIGCHI conference, focused on the relationship between human factors and computer systems. Since then, usability field and its main related concepts (attributes, paradigms, principles, etc.) started to be developed and widely applied.

The concept of usability has been formulated and/or evaluated by several researchers along the years. There are many usability definitions that have been proposed by diverse authors (Bevan, Kirakowski, & Maissel, 1991; Krug, 2005; Nielsen, 1993; Preece, Rogers, & Sharp, 2007). Standardization Organizations such as ISO (International Organization for Standardization), IEC (International Electrotechnical Commission) and ANSI (American

National Standards Institute) have developed standards in which usability is recognized as a very important part of the quality software. These organizations have also proposed their own usability definitions.

The standard ISO/IEC 9241 defines usability as the extent in which a product can be used by specific users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use (ISO, 1998). The standard ISO/IEC 9126 emphasizes usability as “Quality in use” and defines it as the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions (ISO/IEC, 2001). Both standards are focused on the relevance for assessing the usability degree of a software product, taking into account the user environment and its specific needs. A relatively more recent standard is the ISO/IEC 25062. The relevance of this standard, points toward the definition of one standard format to present in a clear and effective way the results of usability evaluations and the employed methods (ISO/IEC, 2006).

According to Nielsen (1993), usability can be defined in quantitative and qualitative terms. The set of five *usability attributes* (learnability, efficiency, memorability, error avoidance, and subjective satisfaction) can be measured for assessing the usability degree of a software system. The concept of usability goes beyond the general or beauty design of the software interface. Usability goes to determine how well the intended users can interact with a technology to successfully carry out the assigned activities and accomplish their intended goals. Therefore, usability is a relevant attribute to measure the quality of a software product.

1.2 Usability Evaluations

Usability evaluations are increasingly used to detect problems in the users’ interaction with the system. According to Lorés (2002), usability evaluations can be applied in every development software stage in order to ensure the accomplishment of the following objectives:

1. Providing feedback for getting a better design.
2. Verifying compliance with the user and organizational goals
3. Controlling the long-term use of a product or system.

Usability evaluations are considered as relevant activities that include a set of methods for analyzing the easiness to use of an interactive system. The usability evaluation tasks can be performed from early to last stages of

the software development lifecycle. This process can be carried out by people with different skills and knowledge. Usability evaluation should involve a multidisciplinary team of representative users, usability specialists, developers, designers, among others. The usability evaluation can help to assure the quality of software products and hence influence in the satisfaction of its users.

Literature reveals that there are several methods for evaluating the usability of a software product. A simple classification proposed by Holzinger (2005) divides the usability evaluation methods in two basic categories: Inspections and Tests.

- Inspections are typically performed by usability experts. Experts analyze the system trying to identify errors and usability design problems. Proposed by Nielsen (1990). This method involves a group of 3 to 5 usability evaluators, who analyze the system using a set of design principles (*usability heuristics*) as guide for discovering usability problems in interactive software systems.
- Tests are empirical evaluation methods which are carried out by representative users. These tests are based in the real experience of users who interact with the system. The main characteristic of this kind of evaluation methods is that they must be performed under controlled conditions and simulating as well as possible the real environment of use.

Several sets of usability heuristics have been proposed to be used in the heuristic evaluations. The selected set of heuristics could significantly influence the results (usability findings) of heuristic evaluations. As it was validated in (Jimenez, Rusu, Rusu, Roncagliolo, & Inostroza, 2012), “depending on evaluators’ experience they could need more specific heuristics or at least a detailed specification of generic ones”. Since it was proposed by Nielsen (1990-a), the set of ten usability heuristics has been probably the most accepted and applied usability heuristics. However, these heuristics are considered too general to be applied to all software domains. Other proposals of heuristics are continuously developed for covering specific features of different application domains.

1.3 Literature Review

According to Okoli (2010), Systematic Literature Review (SLR) is “*a systematic, explicit, comprehensive and reproducible method for identifying, evaluating, and synthesizing the existing body of completed*

Specific-Domain Usability Heuristics: Are they really necessary?

recorded work produced by researchers, scholars, and practitioners". The main goal for conducting a systematic literature review is to capture all available information regarding a particular area, or topic of interest. This systematic process is not a merely collection of information. The aim should include the analytical criticism about the topic and how it has changed over the course of its development or research.

There are several works and guides about how to perform a systematic literature review in the information systems field (Keele, 2007; Okoli & Schabram, 2010; Reis & Prates, 2011). The three phases (planning, conducting and reporting) proposed in the guidelines of Kitchenham and Charters (Keele, 2007) are commonly used in several research works. However, in this study, the guide proposed by Okoli (2010) was followed. Okoli's guide is composed by eight steps which are summarized as follow:

1. Clearly identify the purpose and goals of the review.
2. Establish the protocol and train participants (reviewers) for following a detailed procedure in the execution of the review.
3. Establish explicit details of the literature search for ensuring the comprehensiveness of the search.
4. Establish criteria for including/excluding studies from the review.
5. Establish criteria for excluding studies due their insufficient quality to be included in the review synthesis.
6. Systematically extract the relevant information from each selected study.
7. Analyze the extracted information using suitable techniques (quantitative and/or qualitative).
8. Report in sufficient detail the obtained results.

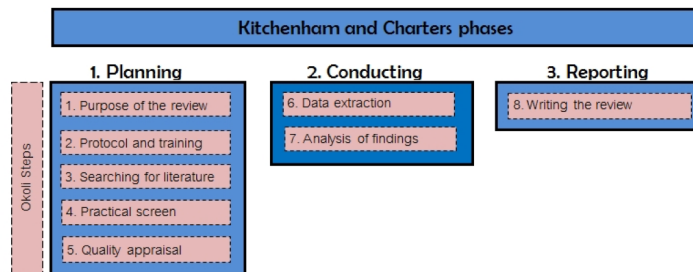


Figure 1. Mapping between Kitchenham phases and Okoli steps for systematic review

2. Execution of the Review

2.1 Purpose and research questions

The use of an adequate set of usability heuristics can improve the performance (usability findings) of usability evaluations. Traditional Nielsen's usability heuristics are too general for covering all specific software applications. This fact has motivated the development of new sets of heuristics, as tools for evaluating the usability of specific software systems.

The main purpose of this systematic review was to categorize and summarize the current work about the development and use of usability heuristics, and also identify gaps and useful knowledge about this topic of research. In order to achieve this purpose, a set of four research questions was established. Table 1 presents the research questions and the motivation of each one of them.

Table 1. Research questions

Id	Research Question	Motivation
RQ1	What usability heuristics are being used for evaluating software systems?	To identify the existence and use of usability heuristics
RQ2	What do authors think about the utility of usability heuristics?	To summarize the comments about the use of usability heuristics
RQ3	What kind of methodological processes exist for developing usability heuristics?	To detect the origin of usability heuristics
RQ4	What kind of validation is used for analyze the performance of usability heuristics?	To identify how usability heuristics have been validated

2.2 Protocol of the review

Data source and search strategy

In order to clarify the limits of the review and reduce the research bias, the procedure to be followed was defined in basis of the subsequent parameters:

- *Primary research source*: The ScienceDirect database was selected as primary research source.
- *Date range*: Only articles published in the period from 2008 and the first trimester of 2015 were collected. Although usability heuristics have been used

Specific-Domain Usability Heuristics: Are they really necessary?

for more than 20 years, this study considers only almost 8 years because, even the pioneer proposals of heuristics have presented continuous updates. Then, we think that a range of almost 8 years would provide us the enough knowledge to be analyzed and synthetized.

- *Language*: English, Spanish, French, Italian and Portuguese articles were considered, due the familiarity with the reviewers' language.
- *Registration search form*: A standard form for registering the performed searches was created. The form included the following fields: Search code, database, journal, keywords and search date. Table 2 presents a description of the fields contained in the registration search form.
- *References Manager*: In order to arrange the selected articles, the Mendeley references manager was used (www.Mendeley.com). Through this system it was possible to organize the articles and also add information regarding the analysis.

Table 2. Registration search form

Field	Description
Search Code	Identifier assigned to each conducted search. It was established as an auto-incremental field in order to keep a record of the quantity of conducted searches.
Journal	The specific journal defined for the search.
Keywords	The keywords used in the search string.
Search date	The date in what the search was conducted.

Primary studies selection strategy

In order to select the articles that should be included into the review, an *Approval/Reject* process was clearly established. Figure 2 depicts the flowchart followed with each article during this process. All obtained results from the automatic search were manually managed at this stage. The manual process involved reading each article (title and abstract) and including/excluding it in basis of some established criteria. If both, title and abstract had no relevant information for the literature review purpose then, the article should be excluded from the review. Table 3 shows the inclusion and exclusion criteria took into account for approving or rejecting every article. Boolean "OR" was used in order to consider alternative terms related with the topic of study.

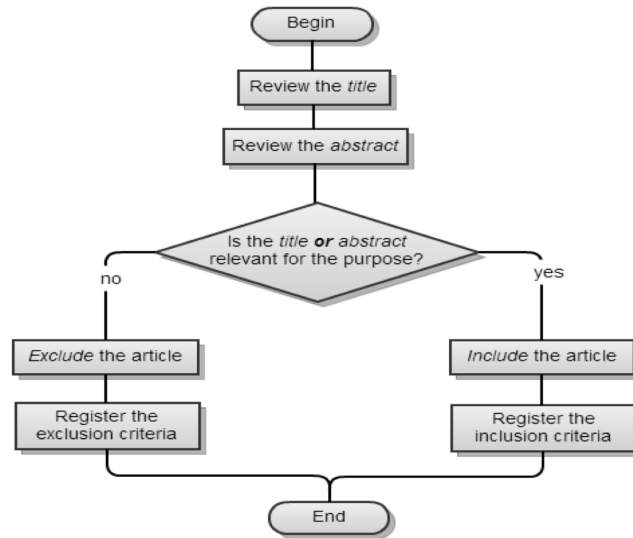


Figure 2. Approval/Reject process

Table 3. Inclusion / exclusion criteria

Inclusion	The article contains terms related to the search string (“usability heuristics”) such as: usability principles, guidelines and checklists.	OR
	The article mentions the use of usability evaluation methods.	OR
	The article is focused on usability evaluations of software systems.	OR
Exclusion	The article has poor and/or irrelevant information to be analyzed regarding the purpose of the review.	

As a tool for supporting the approval/reject process, a standard form was created. The form was used for registering all found articles and also the reasons why they were included or excluded from the review. This information could allow reviewers to reconsider the decision about inclusion/exclusion of a particular article. Table 4 presents a description of the fields contained in the Approval/Reject form.

Specific-Domain Usability Heuristics: Are they really necessary?

Table 4. Approval/Reject Form

Field	Description
Code	Unique identifier of the article
Review date	Date in which the article was reviewed
Approved	Indicator of approval (Yes/No)
Approval/Reject criteria	Summary of reasons to approval or reject

Quality assessing and data extraction strategy

Once selected the set of primary studies to be included in the review, the quality analysis process was performed. This study considers both, quantitative and qualitative analysis of collected information. The primary studies were submitted to a screening process in order to assess their quality to be included or excluded from the analysis. More detailed inclusion/exclusion criteria were established for assessing articles and extracting specific relevant information (data). If an article had no relevant information to be extracted, then it could be excluded from the analysis.

Taking into account the purpose of the review, a set of seven criteria for data extraction was established. Table 5 presents the criteria and its justification, categorized according to the research questions defined in Table 1.

Table 5. Data extraction criteria

RQ	Id	Criterion	Justification
RQ 1	C1	The article employs a set of usability heuristics.	Identify the set of heuristics used in the evaluation of software systems.
RQ 2	C2	The article mentions advantages/disadvantages in the use of usability heuristics.	Collect authors opinion about the use of usability heuristics.
RQ 3	C3	The set of usability heuristics is an original proposal or the adaptation of another set.	Identify new heuristics and how they were obtained.
	C4	The article describes a methodology or explicit process for developing of usability heuristics.	Identify methodologies or stages for developing usability heuristics.
RQ 4	C5	The article is focused on the evaluation of a software system.	Identify the domain of applications which are been evaluated.
	C6	The article mentions the use of usability heuristics for evaluating a software system.	Identify the domain in which the development or use of usability heuristics is predominating.
	C7	The usability heuristics have been used by others than those researchers who developed them.	Detect internal/external validation of heuristics.

2.3 Literature Review Conduction

Once defined the protocol of the review, the search, selection and analysis processes were conducted. The following sections summarize the results obtained in each stage.

Automatic search and primary studies selection process.

In order to identify potential primary studies to be included in the review, a web search was performed as an initial step. The ScienceDirect catalog was used as primary data source. The search string **usability heuristics** was used and the results were limited taking into account previously established search strategy. As result of this stage, 124 primary studies were obtained and all of them were submitted to the Approval/Reject. A total of 59 articles were excluded from the review based on the inclusion/exclusion criteria presented in Table 3 and only 65 articles were finally selected to be analyzed according to data extraction criteria presented in Table 5.

Figure 3 depicts the distribution of analyzed articles according to the year in which they were published. It is possible to observe that most articles were published at 2011 (11 articles). There is an apparent trend to increase the number of related articles from 2008 to 2011. However, the quantity of articles published from 2011 to 2015 seems to be decreasing. It is too early to conclude about the reason of this decreasing. The results of the rest of 2015 and publications from other catalogs should be analyzed in order to get more information about this fact.

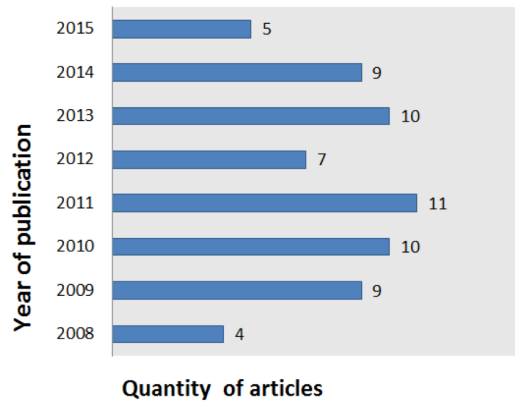


Figure 3. Quantity of articles by year of publication.

Data extraction and quality assessing process

The 65 selected articles were evaluated in order to assess their quality and also extract significant information for the purpose of the review. Each article was evaluated taking into account the seven criteria for data extraction defined in Table 5. If an article had no relevant information regarding any of the seven criteria, then it was excluded from the analysis. During this process, a total of 8 articles (Bardzell, 2011; Broström, Bengtsson, & Axelsson, 2011; Fernandez, Insfran, & Abrahão, 2011; González, Lorés, & Granollers, 2008; Hollender, Hofmann, Deneke, & Schmitz, 2010; Ling & Salvendy, 2009; Youngblood & Mackiewicz, 2012) were excluded. They had no significant information to be extracted in order to fulfill the purpose of the review. Thus, the 57 remaining articles were finally analyzed taking into account the seven established criteria.

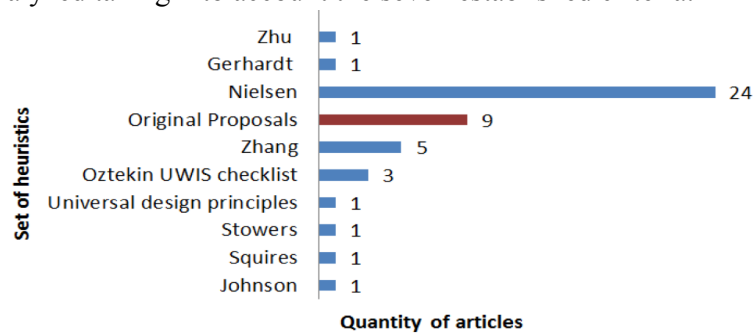


Figure 4. Sets of heuristics mentioned in the analyzed articles.

Table 6 presents the list of 57 analyzed articles with the bibliographic references. A detail of the extracted information can be consulted in the previous version of this paper (Jimenez, Lozada & Rosas, 2016) however, a summary of the most relevant information is presented below:

- Mentioned or used heuristics:* 50 out of 57 analyzed articles mention the use of a specific set of heuristics. In 7 articles (A3, A6, A7, A12, A19, A21 and A22) no set of heuristics is mentioned or used. As it can see, in Figure 4, Nielsen's heuristics are the most commonly used set and Zhang's heuristics are growing in use. It is remarkable to mention that in 9 articles (A10, A14, A16, A20, A25, A30, A34, A42 and A49), authors propose their own set of specific-domain usability heuristics.

- *Software evaluation and domain application:* Most of the articles mentions the evaluation a system belonging specific domain. Only 12 out of 57 articles were not focused on the evaluation of a software system. Medical and educational domains group the majority of the analyzed articles (See Figure 5).

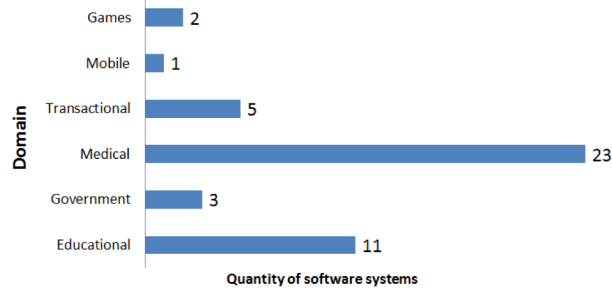


Figure 5. Quantity of evaluated systems by application domain.

- *How heuristics were created:* Most of the mentioned sets of heuristics were developed by combination or adaptation of other different existent sets or even by the addition of features regarding the specific application domain. The analysis of the articles showed that no methodology or explicit process was used for developing the sets of heuristics.
- *External validation of heuristics:* Information about who used the sets of heuristics was collected in order to identify if the proposals of heuristics are validated by others than the authors. The results showed that only Nielsen's and Zhang's heuristics were used by other different researchers. Zhang's heuristics were mentioned in 6 researches (A2, A17, A26, A28, A29 and A46). Oztekin's heuristics were mentioned in 3 researches (A10, A18 and A24) but, all are works of the same author who developed the set of heuristics. Except Nielsen's and Zhang's heuristics, remaining sets have showed limited external validation. In this review, it was not possible to detect evidence in which the other sets of heuristics had been applied.
- *Advantages and disadvantages of heuristics:* Authors opined about the use of the heuristics. Disadvantages were often regarding Nielsen's heuristics because they are too general and not sufficiently

comprehensive. Some authors mentioned that a better performance in the use of Nielsen's heuristics could depend on the evaluator's expertise. Some authors stated that new or adapted specific-domain heuristics have major potential for discovering more specific usability problems because incorporate features of each application domain.

2.4. Synthesis of the Study

RQ1: What usability heuristics are being used for evaluating software systems?

Nielsen's usability heuristics has been widely used in several research works. Nowadays, there are efforts for developing and providing new sets of specific-domain heuristics. So far, Nielsen's heuristics are used as basis for developing or adapting new sets. There are few researches that mention the use of Nielsen's heuristics without adaptation. Due to the lack of specificity of Nielsen's heuristics, some works have included their own proposals that pretend to solve the problem of generality by adding parameters for covering features of the specific applications domains.

Even though many sets of heuristics were identified in this study, other existent ones were not included due to the established limits of the review. Over the last few years, the usability research group (UseCV) has contributed with other domain-specific usability heuristics: Grid Computing (Rusu, Roncagliolo, et al., 2011), Virtual Worlds (Rusu, Muñoz, et al., 2011), Interactive Television (Solano et al., 2011) and Mobile Devices (Inostroza, Rusu, Roncagliolo, Jimenez, & Rusu, 2012) have been developed and tested. The results have been already published, but the catalog that contains them was not included in the protocol of the review.

RQ1: What usability heuristics are being used for evaluating software systems?

The evaluation of specific software systems requires heuristics developed taking into account specific-domain features. The review allowed concluding that specific-domain usability heuristics could provide better results in the evaluation of specific-domain application systems.

Table 6. List of selected articles and Bibliographic Reference

ID	Bibliographic Reference	ID	Bibliographic Reference
A1	(Edwards, Moloney, Jacko, & Sainfort, 2008)	A29	(Chan et al., 2012)
A2	(Martin, Norris, Murphy, & Crowe, 2008)	A30	(Nykänen, Kaipio, & Kuusisto, 2012)
A3	(M. Virvou & Katsionis, 2008; M. a. K. G. Virvou, 2008)	A31	(Khajouei, Peute, Hasman, & Jaspers, 2011)
A4	(Afacan & Erbug, 2009)	A32	(Tambascia, Menezes, Kutiishi, & Barbosa, 2012)
A5	(Baker, 2009)	A33	(Yovcheva, van Elzakker, & Köbben, 2012)
A6	(Conrad et al., 2009)	A34	(M. Brown, Sharples, & Harding, 2013)
A7	(Jaspers, 2009)	A35	(Castilla et al., 2013)
A8	(Kılıç Delice & Güngör, 2009)	A36	(Cobos et al., 2013)
A9	(Muñoz-Arteaga, 2009)	A37	(Erdem, Pala, & Baş, 2013)
A10	(Oztekin, Nikov, & Zaim, 2009)	A38	(Preece et al., 2013)
A11	(Peleg, Shachak, Wang, & Karnieli, 2009)	A39	(Raji, Mahmud, & Abubakr, 2013)
A12	(Boring, Hendrickson, Forester, Tran, & Lois, 2010)	A40	(Rashid, Soo, Sivaji, Naeni, & Bahri, 2013)
A13	(Choi & Bakken, 2010)	A41	(Rogers, Sockolow, Bowles, Hand, & George, 2013)
A14	(Doherty, Coyle, & Matthews, 2010)	A42	(Tan, Goh, Ang, & Huan, 2013)
A15	(Gartner, 2010; Gärtner, Seidel, Froschauer, & Berger, 2010)	A43	(Walsh, Carroll, & Sleator, 2013)
A16	(Horsky et al., 2010)	A44	(Bauleo et al., 2014)
A17	(Khajouei, Peek, Wierenga, Kersten, & Jaspers, 2010)	A45	(Cravo et al., 2014)
A18	(Oztekin, Kong, & Uysal, 2010)	A46	(Devine et al., 2014)
A19	(Welle Donker-Kuijter, de Jong, & Lentz, 2010)	A47	(Flewellling, Easty, Vicente, & Cafazzo, 2014)
A20	(D. J. Brown et al., 2011)	A48	(Huang & Benyoucef, 2014)
A21	(Granić, Mitrović, & Marangunić, 2011)	A49	(Lacerda, von Wangenheim, von Wangenheim, & Giuliano, 2014)
A22	(Khajouei et al., 2010)	A50	(Melo, Abreu, & Silva, 2014)
A23	(Lanzilotti, Ardito, Costabile, & De Angeli, 2011; Okoli & Schabram, 2010)	A51	(Pinho, Oliveira, Oliveira, Dinis, & Marques, 2014)
A24	(Oztekin, 2011)	A52	(Van Rosmalen, Boon, Bitter-Rijkema, Sie, & Sloep, 2014)
A25	(Taylor, Sullivan, Mullen, & Johnson, 2011)	A53	(Chang et al., 2015)
A26	(Zhang & Walji, 2011)	A54	(Kim & Lee, 2015)
A27	(Andrade-Aréchiga, López, & López-Morteo, 2012)	A55	(Lilholt, Jensen, & Hejlesen, 2015)
A28	(Bakhshi-Raiez et al., 2012)	A56	(Osman, Yahaya, & Ahmad, 2015)
		A57	(Tawfik & Anya, 2015)

RQ3: What kind of methodological processes exist for developing usability heuristics?

There was no evidence of a methodology or explicit process for developing usability heuristics. The identified sets of heuristics were obtained through the adaptation and/or combination of other existing ones. It is important to mention that in some cases, not only existing usability heuristics were considered, but also parameters of software quality, experts' recommendations and features of specific-domain applications.

RQ4: What kind of validation is used to analyze the performance of usability heuristics?

In most of the analyzed articles, usability heuristics were used for assessing the usability of software systems. Medical and educational applications were the most commonly kind of evaluated systems. Through this review, the use of some different sets of usability heuristics was identified. Some of these sets were just applied only by the researchers who developed them. Other sets of heuristics were applied by original authors but, also by other. The use of Zhang's heuristics for medical applications is increasing nowadays.

3. Conclusions and Future Work

Through this work, it was possible to evidence that the need for providing specific-domain usability heuristics is a reality. Through a literature review, a preliminary investigation regarding the use, existence, and progress in the development of usability heuristics was conducted. The obtained results pointed to conclude that the widely-accepted Nielsen's heuristics present some deficiencies for evaluating specific domain software systems. Nowadays, it is common to find new specific heuristics obtained through the adaptation of Nielsen's ones considering specific features of every domain of application.

Even though the generality of Nielsen's heuristics was stated as a main disadvantage, at the same time the most of the efforts for providing new sets

of heuristics often use the ten usability heuristics as a basis. The main reason for incorporating specific-domain features is focused on to avoid that usability problems are related to the wrong category of problems, or in worst of the cases, they are not been identified.

Although, the need for specific-domain usability heuristics is a reality, so far, there is no evidence about explicit process or methodology for guiding their development and validation. The results of this work were taken as the basis for creating and validating a proposal of methodology which has been named PROMETHEUS (PROcedural METHodology for developing HEuristics of USability); but, it will be presented in a future work. PROMETHEUS includes a set of stages that clearly define several activities for developing and describing specific-domain usability heuristics.

Finally, this study included a wide period time for collecting primary studies. However, we are conscious that the future work points to extend the study by incorporating other search sources such as IEEE or ACM catalogs. Surely, additional relevant information regarding the development of usability heuristics may be found in conferences and workshops catalogues. It is possible that researchers prefer spreading their results in another kind of scientific publications, due to the fast feedback in the revision process.

In a recent related study (Hermawati & Lawson 2016) performed a bibliographic review of 70 articles related to the development of specific domain heuristics, identifying two large problems: (1) significant deficiencies in efforts to validate new heuristics, and (2) lack of rigor, robustness and standardization in the effectiveness analysis of domain heuristics. The study performed by Hermawati & Lawson, included articles of several databases (not only Science Direct) and the results could help to support the findings presented in our study regarding the need of specific usability heuristics and also the lack of a formal methodology to develop those heuristics.

References

- Afacan, Y., & Erbug, C. (2009). An interdisciplinary heuristic evaluation method for universal building design. *Applied Ergonomics*, 40(4), 731-744.
- Andrade-Aréchiga, M., López, G., & López-Morteo, G. (2012). Assessing effectiveness of learning units under the teaching unit model in an undergraduate mathematics course. *Computers & Education*, 59(2), 594-606.
- Baker, D. L. (2009). Advancing E-Government performance in the United States through enhanced usability benchmarks. *Government Information Quarterly*, 26(1), 82-88.

Specific-Domain Usability Heuristics: Are they really necessary?

- Bakhshi-Raiey, F., de Keizer, N. F., Cornet, R., Dorrepaal, M., Dongelmans, D., & Jaspers, M. W. M. (2012). A usability evaluation of a SNOMED CT based compositional interface terminology for intensive care. *International Journal of Medical Informatics*, 81(5), 351-362.
- Bardzell, J. (2011). Interaction criticism: An introduction to the practice. *Interacting with Computers*, 23(6), 604-621.
- Bauleo, E., Carnevale, S., Catarci, T., Kimani, S., Leva, M., & Mecella, M. (2014). Design, realization and user evaluation of the SmartVortex Visual Query System for accessing data streams in industrial engineering applications. *Journal of Visual Languages & Computing*, 25(5), 577-601.
- Boring, R. L., Hendrickson, S. M. L., Forester, J. A., Tran, T. Q., & Lois, E. (2010). Issues in benchmarking human reliability analysis methods: A literature review. *Reliability Engineering & System Safety*, 95(6), 591-605.
- Broström, R., Bengtsson, P., & Axelsson, J. (2011). Correlation between safety assessments in the driver-car interaction design process. *Applied Ergonomics*, 42(4), 575-582.
- Brown, D. J., McHugh, D., Standen, P., Evett, L., Shopland, N., & Battersby, S. (2011). Designing location-based learning experiences for people with intellectual disabilities and additional sensory impairments. *Computers & Education*, 56(1), 11-20.
- Brown, M., Sharples, S., & Harding, J. (2013). Introducing PEGI: A usability process for the practical evaluation of Geographic Information. *International Journal of Human-Computer Studies*, 71(6), 668-678.
- Castilla, D., Garcia-Palacios, A., Bretón-López, J., Miralles, I., Baños, R. M., Etchemendy, E., . . . Botella, C. (2013). Process of design and usability evaluation of a telepsychology web and virtual reality system for the elderly: Butler. *International Journal of Human-Computer Studies*, 71(3), 350-362.
- Chan, A. J., Islam, M. K., Rosewall, T., Jaffray, D. A., Easty, A. C., & Cafazzo, J. A. (2012). Applying usability heuristics to radiotherapy systems. *Radiotherapy and Oncology*, 102(1), 142-147.
- Chang, C.-W., Ma, T.-Y., Choi, M.-S., Hsu, Y.-Y., Tsai, Y.-J., & Hou, T.-W. (2015). Electronic Personal Maternity Records: Both Web and Smartphone Services. *Computer Methods and Programs in Biomedicine*.
- Choi, J., & Bakken, S. (2010). Web-based education for low-literate parents in Neonatal Intensive Care Unit: Development of a website and heuristic evaluation and usability testing. *International Journal of Medical Informatics*, 79(8), 565-575.
- Cobos, C., Rodriguez, O., Rivera, J., Betancourt, J., Mendoza, M., León, E., & Herrera-Viedma, E. (2013). A hybrid system of pedagogical pattern recommendations based on singular value decomposition and variable data attributes. *Information Processing & Management*, 49(3), 607-625.
- Conrad, F. G., Bederson, B. B., Lewis, B., Peytcheva, E., Traugott, M. W., Hanmer, M. J., . . . Niemi, R. G. (2009). Electronic voting eliminates hanging chads but introduces new usability challenges. *International Journal of Human-Computer Studies*, 67(1), 111-124.

- Cravo, J., Almeida, F., Abreu, P. H., Reis, L. P., Lau, N., & Mota, L. (2014). Strategy planner: Graphical definition of soccer set-plays. *Data & Knowledge Engineering, 94, Part A(0)*, 110-131.
- Devine, E. B., Lee, C.-J., Overby, C. L., Abernethy, N., McCune, J., Smith, J. W., & Tarczy-Hornoch, P. (2014). Usability evaluation of pharmacogenomics clinical decision support aids and clinical knowledge resources in a computerized provider order entry system: A mixed methods approach. *International Journal of Medical Informatics, 83(7)*, 473-483.
- Doherty, G., Coyle, D., & Matthews, M. (2010). Design and evaluation guidelines for mental health technologies. *Interacting with Computers, 22(4)*, 243-252.
- Edwards, P. J., Moloney, K. P., Jacko, J. A., & Sainfort, F. (2008). Evaluating usability of a commercial electronic health record: A case study. *International Journal of Human-Computer Studies, 66(10)*, 718-728.
- Erdem, M., Pala, F. K., & Baş, T. (2013). A Usability Study of an Online Instructional Multi-media Discussion Environment. *Procedia - Social and Behavioral Sciences, 83(0)*, 786-792.
- Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. *Information and Software Technology, 53(8)*, 789-817.
- Flewwelling, C. J., Easty, A. C., Vicente, K. J., & Cafazzo, J. A. (2014). The use of fault reporting of medical equipment to identify latent design flaws. *Journal of Biomedical Informatics, 51(0)*, 80-85.
- Gartner, M. a. S. I. a. F. J. a. B. H. (2010). The formation of virtual organizations by means of electronic institutions in a 3D e-Tourism environment. *Information Sciences, 180(17)*, 3157--3169. doi: 10.1016/j.ins.2010.03.010
- Gerhardt-Powals, J. (1996). Cognitive engineering principles for enhancing human-computer performance. *Int. J. Hum.-Comput. Interact., 8(2)*, 189-211. doi: 10.1080/10447319609526147
- González, M. P., Lorés, J., & Granollers, A. (2008). Enhancing usability testing through datamining techniques: A novel approach to detecting usability problem patterns for a context of use. *Information and Software Technology, 50(6)*, 547-568.
- Gould, J., & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM, 28(3)*, 300--311.
- Granić, A., Mitrović, I., & Marangunić, N. (2011). Exploring the usability of web portals: A Croatian case study. *International Journal of Information Management, 31(4)*, 339-349.
- Gärtner, M., Seidel, I., Froschauer, J., & Berger, H. (2010). The formation of virtual organizations by means of electronic institutions in a 3D e-Tourism environment. *Information Sciences, 180(17)*, 3157-3169.
- Hermawati, S., & Lawson, G. (2016). Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus?. *Applied ergonomics, 56*, 34-51.
- Hollender, N., Hofmann, C., Deneke, M., & Schmitz, B. (2010). Integrating cognitive load theory and concepts of human-computer interaction. *Computers in Human Behavior, 26(6)*, 1278-1288.

Specific-Domain Usability Heuristics: Are they really necessary?

- Holzinger, A. (2005). Usability engineering methods for software developers. *Communications of the ACM*, 48(1), 71-74.
- HONcode: Principles - Quality and trustworthy health information. (1997). from
- Horsky, J., McColgan, K., Pang, J. E., Melnikas, A. J., Linder, J. A., Schnipper, J. L., & Middleton, B. (2010). Complementary methods of system usability evaluation: Surveys and observations during software design and development cycles. *Journal of Biomedical Informatics*, 43(5), 782-790.
- Huang, Z., & Benyoucef, M. (2014). Usability and credibility of e-government websites. *Government Information Quarterly*, 31(4), 584-595.
- Inostroza, R., Rusu, C., Roncagliolo, S., Jimenez, C., & Rusu, V. (2012). *Usability heuristics for touchscreen-based mobile devices*. Paper presented at the Information Technology: New Generations (ITNG), 2012 Ninth International Conference on.
- Jaspers, M. W. M. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. *International Journal of Medical Informatics*, 78(5), 340-353.
- Jimenez, C., Lozada, P., Rosas, P. (2016). *Usability Heuristics: A Systematic Literature Review*. Paper presented at the Proceedings of the 11 Congreso Colombiano de Computación. IEEE Xplorer Digital Library. *IN-PRESS*
- Jimenez, C., Rusu, C., Rusu, V., Roncagliolo, S., & Inostroza, R. (2012). *Formal specification of usability heuristics: how convenient it is?* Paper presented at the Proceedings of the 2nd international workshop on Evidential assessment of software technologies.
- Johnson, M. L., & Zurko, M. E. (2007). *Security user studies and standards: Creating best practices*. *Proceedings of Workshop on Security User Studies*, San Jose, California.
- Keele, S. (2007). Guidelines for performing systematic literature reviews in software engineering: Technical report, EBSE Technical Report EBSE-2007-01.
- Khajouei, R., Peek, N., Wierenga, P. C., Kersten, M. J., & Jaspers, M. W. M. (2010). Effect of predefined order sets and usability problems on efficiency of computerized medication ordering. *International Journal of Medical Informatics*, 79(10), 690-698.
- Khajouei, R., Peute, L. W. P., Hasman, A., & Jaspers, M. W. M. (2011). Classification and prioritization of usability problems using an augmented classification scheme. *Journal of Biomedical Informatics*, 44(6), 948-957.
- Kılıç Delice, E., & Güngör, Z. (2009). The usability analysis with heuristic evaluation and analytic hierarchy process. *International Journal of Industrial Ergonomics*, 39(6), 934-939.
- Kim, Y.-H., & Lee, J.-H. (2015). Game interface enhancement under smartphone platform focused on touchscreen interaction. *Computers & Industrial Engineering*, 80(0), 45-61.
- Lacerda, T. C., von Wangenheim, C. G., von Wangenheim, A., & Giuliano, I. (2014). Does the use of structured reporting improve usability? A comparative evaluation of the usability of two approaches for findings reporting in a large-scale telecardiology context. *Journal of Biomedical Informatics*, 52(0), 222-230.
- Lanzilotti, R., Ardito, C., Costabile, M. F., & De Angeli, A. (2011). Do patterns help novice evaluators? A comparative study. *International Journal of Human-Computer Studies*, 69(1-2), 52-69.

- Lilholt, P. H., Jensen, M. H., & Hejlesen, O. K. (2015). Heuristic evaluation of a telehealth system from the Danish TeleCare North Trial. *International Journal of Medical Informatics, 84*(5), 319-326.
- Ling, C., & Salvendy, G. (2009). Effect of evaluators' cognitive style on heuristic evaluation: Field dependent and field independent evaluators. *International Journal of Human-Computer Studies, 67*(4), 382-393.
- Lorés, J. a. G. T. a. L. S. (2002). Introducción a la Interacción Persona Ordenador.
- Martin, J. L., Norris, B. J., Murphy, E., & Crowe, J. A. (2008). Medical device development: The challenge for ergonomics. *Applied Ergonomics, 39*(3), 271-283.
- Melo, L., Abreu, J., & Silva, T. (2014). Mobile Applications for Natural Parks: Guidelines Study for the Development of a Mobile Device Application. *Procedia Technology, 16*(0), 263-268.
- Muñoz-Arteaga, J. a. G. R. M. a. M. M. V. a. V. J. a. Á.-R. F. (2009). A methodology for designing information security feedback based on User Interface Patterns. *Advances in Engineering Software, 40*(12), 1231--1241. doi: 10.1016/j.advengsoft.2009.01.024
- Nielsen, J. (1990-a). 10 Heuristics for User Interface Design: Article by Jakob Nielsen.
- Nielsen, J. (1993). *Usability Engineering* (Vol. 44): Morgan Kaufmann.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Chi '90, 17*(1), 249--256. doi: 10.1145/97243.97281
- Nykänen, P., Kaipio, J., & Kuusisto, A. (2012). Evaluation of the national nursing model and four nursing documentation systems in Finland – Lessons learned and directions for the future. *International Journal of Medical Informatics, 81*(8), 507-520.
- Okoli, C., & Kira, S. (2010). Working Papers on Information Systems A Guide to Conducting a Systematic Literature Review of Information Systems Research. *10*(2010).
- Okoli, C., & Schabram, K. (2010). Working Papers on Information Systems A Guide to Conducting a Systematic Literature Review of Information Systems Research. *10*(2010).
- Osman, A., Yahaya, W. A. J. W., & Ahmad, A. C. (2015). Educational Multimedia App for Dyslexia Literacy Intervention: A Preliminary Evaluation. *Procedia - Social and Behavioral Sciences, 176*(0), 405-411.
- Oztekin, A. (2011). A decision support system for usability evaluation of web-based information systems. *Expert Systems with Applications, 38*(3), 2110-2118.
- Oztekin, A., Kong, Z. J., & Uysal, O. (2010). UseLearn: A novel checklist and usability evaluation method for eLearning systems by criticality metric analysis. *International Journal of Industrial Ergonomics, 40*(4), 455-469.
- Oztekin, A., Nikov, A., & Zaim, S. (2009). UWIS: An assessment methodology for usability of web-based information systems. *Journal of Systems and Software, 82*(12), 2038-2050.
- Peleg, M., Shachak, A., Wang, D., & Karnieli, E. (2009). Using multi-perspective methodologies to study users' interactions with the prototype front end of a guideline-based decision support system for diabetic foot care. *International Journal of Medical Informatics, 78*(7), 482-493.

Specific-Domain Usability Heuristics: Are they really necessary?

- Pinho, C., Oliveira, A., Oliveira, D., Dinis, J., & Marques, A. (2014). RIBS@UA: Interface to collect and store respiratory data, a preliminary study. *Computers in Biology and Medicine*, 47(0), 44-57.
- Preece, M. H. W., Hill, A., Horswill, M. S., Karamatic, R., Hewett, D. G., & Watson, M. O. (2013). Applying heuristic evaluation to observation chart design to improve the detection of patient deterioration. *Applied Ergonomics*, 44(4), 544-556.
- Raji, S. O., Mahmud, M., & Abubakr, A. (2013). Evaluation of University Teaching Hospital Websites in Nigeria. *Procedia Technology*, 9(0), 1058-1064.
- Rashid, S., Soo, S. T., Sivaji, A., Naeni, H. S., & Bahri, S. (2013). Preliminary Usability Testing with Eye Tracking and FCAT Analysis on Occupational Safety and Health Websites. *Procedia - Social and Behavioral Sciences*, 97(0), 737-744.
- Reis, D. S., & Prates, R. O. (2011). *Applicability of the semiotic inspection method: a systematic literature review*. Paper presented at the Proceedings of the 10th Brazilian Symposium on on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction.
- Rogers, M. L., Sockolow, P. S., Bowles, K. H., Hand, K. E., & George, J. (2013). Use of a human factors approach to uncover informatics needs of nurses in documentation of care. *International Journal of Medical Informatics*, 82(11), 1068-1074.
- Rusu, C., Muñoz, R., Roncagliolo, S., Rudloff, S., Rusu, V., & Figueroa, A. (2011). *Usability heuristics for virtual worlds*. Paper presented at the AFIN 2011, The Third International Conference on Advances in Future Internet.
- Rusu, C., Roncagliolo, S., Tapia, G., Hayvar, D., Rusu, V., & Gorgan, D. (2011). *Usability heuristics for grid computing applications*. Paper presented at the ACHI 2011, The Fourth International Conference on Advances in Computer-Human Interactions.
- Shneiderman, S. B., & Plaisant, C. (2005). Designing the user interface 4 th edition. *ed: Pearson Addison Wesley, USA*.
- Solano, A., Rusu, C., Collazos, C., Roncagliolo, S., Arciniegas, J. L., & Rusu, V. (2011). *Usability Heuristics for Interactive Digital Television*. Paper presented at the AFIN 2011, The Third International Conference on Advances in Future Internet.
- Squires, D. a. P. J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with computers*, 11(5), 467-483.
- Stowers, G. (2002). The state of federal websites: The pursuit of excellence E-government
- Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.*, 3(3), 3-3.
- Tambascia, C. A., Menezes, E. M., Kutiishi, S. M., & Barbosa, R. C. (2012). Usability Evaluation of Electronic Government Services for Interactive TV. *Procedia Computer Science*, 14(0), 301-310.
- Tan, J. L., Goh, D. H.-L., Ang, R. P., & Huan, V. S. (2013). Participatory evaluation of an educational game for social skills acquisition. *Computers & Education*, 64(0), 70-80.
- Tawfik, H., & Anya, O. (2015). Evaluating practice-centered awareness in cross-boundary telehealth decision support systems. *Telematics and Informatics*, 32(3), 486-503.

- Taylor, H. A., Sullivan, D., Mullen, C., & Johnson, C. M. (2011). Implementation of a user-centered framework in the development of a web-based health information database and call center. *Journal of Biomedical Informatics*, 44(5), 897-908.
- The Center for Universal Design - Universal Design Principles. (1997). from http://www.ncsu.edu/ncsu/design/cud/about_ud/udprinciplestext.htm
- Tognazzini, B. (2014). First Principles of Interaction Design (Revised & Expanded) | askTog. from <http://asktog.com/atc/principles-of-interaction-design/>
- Tufte, E. (2001). The visual display of quantitative information. *Graphics Press, Cheshire, USA*, 4(5), 6.
- Van Rosmalen, P., Boon, J., Bitter-Rijkema, M., Sie, R., & Sloep, P. (2014). Supporting co-creation with software, the idSpace platform. *Computers in Human Behavior*, 37(0), 413-422.
- Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education*, 50(1), 154-178.
- Virvou, M. a. K. G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education*, 50(1), 154--178. doi: 10.1016/j.compedu.2006.04.004
- W3C. (1999). Web Content Accessibility Guidelines 1.0.
- Walsh, P., Carroll, J., & Sleator, R. D. (2013). Accelerating in silico research with workflows: A lesson in Simplicity. *Computers in Biology and Medicine*, 43(12), 2028-2035.
- Welle Donker-Kuijter, M., de Jong, M., & Lentz, L. (2010). Usable guidelines for usable websites? An analysis of five e-government heuristics. *Government Information Quarterly*, 27(3), 254-263.
- Winker, M. A. a. F. A. a. C.-L. B. a. W. J. a. A. K. a. K. R. L. a. D. C. D. a. M. R. A. (2000). Guidelines for medical and health information sites on the Internet. *JAMA: the journal of the American Medical Association*, 283(12), 1600--1606.
- Youngblood, N. E., & Mackiewicz, J. (2012). A usability analysis of municipal government website home pages in Alabama. *Government Information Quarterly*, 29(4), 582-588.
- Yovcheva, Z., van Elzakker, C. P. J. M., & Köbben, B. (2012). User requirements for geo-collaborative work with spatio-temporal data in a web-based virtual globe environment. *Applied Ergonomics*, 44(6), 929-939.
- Zhang, J., R., J. T., L., P. V., L., P. D., & Tate, K. (2003). Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics*, 36(1-2), 23--30.
- Zhang, J., Todd, J., Vimla, P., Danielle, P., & Tate, K. (2003). Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics*, 36(1-2), 23--30.
- Zhang, J., & Walji, M. F. (2011). TURF: Toward a unified framework of EHR usability. *Journal of Biomedical Informatics*, 44(6), 1056-1067.
- Zhu, W., Vu, K.-P. L., & Proctor, R. W. (2005). Evaluating web usability. *Handbook of human factors in web design*, 321-337.