International Journal of Engineering & Technology, 7 (2.28) (2018) 10-19

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



A Formal Protocol to Conduct Usability Heuristic Evaluations in the Context of the Software Development Process

Freddy Paz*[1], Freddy A. Paz[2], José Antonio Pow-Sang[1], César Collazos[3]

[1]Pontificia Universidad Católica del Perú, San Miguel, Lima 32, Lima, Peru
[2]Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Peru
[3]Universidad del Cauca, Popayán, Colombia
**Email: fpaz@pucp.pe*

Abstract

Heuristic evaluation is one of the most used techniques to evaluate the level of usability of a software product. In this research, we performed a comprehensive analysis of the recent studies which report the use of this method in the context of a software development process. The purpose was to identify the specific way in which each author performs this usability evaluation method, in order to propose a formal protocol. After an indeed examination of these studies, we have determined there are several differences in the way this technique is conducted according to the literature. There is no agreement about the number of inspectors that should participate, the usability principles that should be used, the profile of the specialists who must be part of the assessment team, or the evaluation process that should be followed. This work highlights the available settings and a detailed procedure to perform a heuristic evaluation in the domain of software products.

Keywords: *assessment protocol; heuristic evaluation; Human-Computer Interaction; usability evaluation methods; usability principles.*

1. Introduction

Usability is considered one of the most important aspects of software quality nowadays [13]. If a computer system is hard to use or provides an interface that is difficult to understand, the application leans towards failure [15]. As a result of this fact, several techniques have emerged to provide software developers with specific mechanisms to measure the usability of computer systems. The heuristic evaluation method is widely used in this field because of the advantages it offers. According to Nielsen [16], this technique allows identifying about 75% of the usability problems in the graphical user interfaces, demanding the participation of only five specialists. In contrast to usability tests that involve a large number of end users, the heuristic evaluation method represents a useful alternative regarding cost-benefit ratio [18].

Although the heuristic evaluation method is frequently employed in the software development process to ensure the system reaches acceptable levels of usability [20], there is no agreement on various aspects of the execution of this technique. In this research, we have analyzed all the recent studies that report the use of the heuristic evaluation method in the software domain. After a comprehensive analysis of all the case studies, we have determined the differences in the settings and procedure that the authors establish to perform this technique. The discrepancies are noticeable in five aspects: (1) the number of specialists that the authors state is appropriate for the evaluation, (2) the profile that the specialists should meet to participate as reviewers in the execution of this method, (3) the set of usability principles that must be used as guidelines to identify design problems, (4) the phase of the software development process in which the method should be used, and (5) the procedure must be followed to conduct an appropriate heuristic evaluation.

This work details the diversity of considerations are supported by the literature to carry out heuristic evaluations in the domain of software applications. Furthermore, an assessment protocol was developed with the purpose to provide specialists with a structured and complete guide to properly execute this evaluation method. This new assessment proposal collects the best practices of all the studies that were analyzed, and it describes step by step the set of activities are required to reach a successful usability assessment. The new approach that is established in this work is focused on the evaluation of software products. However, future experimental case studies can demonstrate its validity in other scenarios.

This paper is structured as follows. In section 2, we describe the main concepts of the field of Human-Computer Interaction that have allowed the development of the present research. In section 3, we explain the methodology was used to propose our new approach. In Section 4, we discuss the results of our research. Finally, the conclusions and future works are established in Section 5.

2. Background

2.1. Usability

Usability is a quality attribute related to the ease of use of a product. This concept is not exclusive of software applications, but it can as well be applied to hardware products. The International Organization for Standardization (ISO) offers two definitions, one oriented for each context. ISO 9241-11 provides a wide general meaning of usability that is associated with the interaction between humans and machines. In this particular standard, usability

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

is described as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [10]. However, in the norm ISO 9126-1, this concept is framed into the field of the software applications. Usability is defined as "the capability of a software product to be understood, learned, used and attractive to users, when it is used under specified conditions" [11]. Although each definition is established in a particular area, both concepts highlight the importance of an intuitive interface that allows users, through the use of the technological product, to reach their purpose on the edge of a pleasant experience.

Another recognized definition is proposed by Nielsen [15], who states usability is related to user satisfaction that arises as a result of the use of a computer system. For this author, the usability is how comfortable and pleasant the features of a software product are to use. In this theory, the concept of usability is defined by five attributes: learnability, efficiency, memorability, errors, and satisfaction. If all these conditions are met, users will return to use the software again, and there would be a high probability for the system to become a successful product.

2.2. Usability evaluation methods

Usability has currently become an important aspect to be considered during all the phases of the software development process. The importance of designing usable interfaces has led to the emergence of several techniques to evaluate the usability of computer systems. These methods provide specialists with a set of mechanisms to determine, in a systematic way, if the graphical user interface of a specific software meets an appropriate degree of usability.

A usability evaluation method is defined in the literature as "a procedure which is composed of a set of well-defined activities for collecting usage data related to end-user interaction with a software product" [4]. These evaluation techniques allow specialists to analyze how specific properties of a particular software system, help to achieve the desired degree of usability in the interface.

There are several taxonomies to classify the usability evaluation methods. Nielsen [16], Whitefield et al. [23] and Holzinger [6] establish these techniques can be categorized according to the type of participant is considered to perform the assessment. In this model, authors propose two categories:

- *Inspection methods* which involve the participation of specialists in the field of Human-Computer Interaction (HCI). Some examples of this type of methods are the heuristic evaluation, the cognitive walkthrough, and the review of checklists.
- *Testing methods* which involve the participation of representative users of the software system to be evaluated. In this category are considered the user testing, the thinking aloud tests, and the co-discovery.

Other classification proposals, such as Ivory et al. [12] and Fernandez [4], include three more categories than the traditional approaches. This taxonomy is established by the way the information is collected, and the type of results are obtained:

- *Inquiry methods* which allow obtaining feedback from participants as well as their preferences and feelings about the interface of a software product. Examples of this category are the interview, the questionnaire, and the focus group.
- *Analytical modeling methods* that establish an engineering approach to obtain quantitative results about the usability of a software product. Some instances of this category are the software metrics, the GOMS analysis, and the task analysis.
- *Simulation methods* which provide algorithms to simulate the user behaviour. The place/transition net models, and the automated tools to evaluate the usability are considered in this group.

In this work, we focus on studying the heuristic evaluation, which is classified as a usability inspection method. All studies which describe the use of this method were analyzed to identify the differences in the assessment process as well as in the initial considerations.

2.3. Heuristic evaluation

The heuristic evaluation is an inspection method whose use is one of the most reported in the literature [19]. This technique involves having a group of usability specialists who determines whether each element of the software interface follows established usability principles called "heuristics" [16]. Nielsen proposed this type of evaluation as an alternative to user tests [15]. In contrast to those methods in which the participation of a high number of end-users is required, the heuristic evaluation allows identifying 75% of the total number of usability problems in the interface, with the involvement of only five specialists. According to Nielsen, the appropriate number of evaluators is from three to five usability experts to maintain the cost-benefit relation of the software project [15]. The participation of more than five professionals would represent an additional cost for the project with only minimal benefits, given that most of the problems would be over-detected.

The evaluation process establishes that each evaluator should examine all the system interfaces and judge its compliance with recognized usability principles. The set of principles that are commonly used to perform this method are the ten heuristics for user interface design proposed by Nielsen [14]. However, these guidelines are general, and fail to cover specific usability aspects of the software domain [17]. There is enough evidence in the literature proving that these conventional heuristics are no longer suitable for the new categories of software applications that are emerging nowadays [8] [3] [2], such as, videogames, virtual worlds, applications for mobile devices, augmented reality applications, social networks, and software for research in basic and applied sciences. For this reason, new heuristics have been developed by different authors to obtain more accurate and effective results when a heuristic evaluation is performed to a non-traditional type of software.

In the present research, we discuss the categories of software that are currently assessed by the heuristic evaluation method, the usability principles are used to perform this evaluation, the number of evaluators that are considered to identify problems, and the type of specialists are recruited for these assessments. Finally, we establish an evaluation protocol to carry out the heuristic evaluation method based on the analysis of the procedures described by the authors who report the use of this method in the context of the software development process.

3. Development of the protocol

In a previous work [19], a systematic mapping study was conducted to identify all cases studies that inform about the use of a usability evaluation method in the software domain. The purpose of this revision was to determine the most reported techniques in the last three years (from 2012 to 2015) to evaluate the usability of software applications. The automated search was performed once more to update our systematic study until October 9th, 2016 with relevant studies. From these results (a total of 1615 papers), only those, which specified the employment of the heuristic evaluation, were submitted to a comprehensive analysis. Our proposed protocol is based on the examination of 71 studies, which are presented in Table 1. Each study is given an identifier that will be used during all document to reference the paper.

Table 1: List of relevant papers

ID	Paper
P01	Alexander GL, Wakefield BJ, Anbari AB, Lyons V, Prentice D, Shepherd M, Strecker EB & Weston MJ, "A usability evaluation exploring the design of American Nurses Association State web sites", CIN: Computers, Informatics, Nursing, vol. 32, no. 8, 2014, pp. 378-387.
P02	Alomar N, Almobarak N, Alkoblan S, Alhozaimy S & Alharbi S, "Usability engineering of agile software project management tools", Proceedings of the 18th International Conference on Human-Computer Interaction (HCI International 2016), 2016, pp. 197-208.
P03	Alroobaea R & Mayhew PJ, "How many participants are really enough for usability studies?", Proceedings of the 2014 Science and Information Conference (SAI 2014), 2014, pp. 48-56.
P04	Altin Gumussoy C, "Usability guideline for banking software design", Computers in Human Behavior, vol. 62, no. C, 2016, pp. 277-285.
P05	Allende-Hernández O & Caballero-Morales S-O, "MODELI: An emotion-based software engineering methodology for the development of digital learning objects for the preservation of the Mixtec language", Sustainability, vol. 7, no. 7, 2015, pp. 9344-9394.
P06	Anwar S, Motla YH, Siddiq Y, Asghar S, Hassan MS & Khan ZI, "User-centered design practices in scrum development process: A distinctive advantage?", Proceedings of the 2014 IEEE 17th International Multi-Topic Conference (INMIC 2014), 2014, pp. 161-166.
P07	Arshad F, Nnamoko N, Wilson J, Roy B & Taylor M, "Improving healthcare system usability without real users: A semi-parallel design approach", International Journal of Healthcare Information Systems and Informatics (IJHISI), vol. 10, no. 1, 2015, pp. 67-81.
P08	Bandi A & Heeler P, "Usability testing: A software engineering perspective", Proceedings of the IEEE International Conference on Human Computer Interactions (ICHCI'13), 2013, pp. 1-8.
P09	Bonifácio B, Fernandes P, Santos F, d. Oliveira FHABF & Conte T, "Usability of mobile web applications: Evaluating a new approach inspection through experimental studies", Proceedings of the XV Ibero-American Conference on Software Engineering (CIbSE 2012), 2012.
P10	Bošković D & Borovina N, "Heuristic evaluation in the human-computer interaction course", Proocedings of the 36th International Convention on Information & Communication Technology Electronics & Microelectronics (MIPRO 2013), 2013, pp. 685-688.
P11	Bozyer Z & Durdu PO, "Heuristic evaluation of a MMORPG: Guild Wars 2", Proceedings of the 16th International Conference on Human-Computer Interaction (HCI International 2014), 2014, pp. 9-14.
P12	Brayshaw M, Gordon N, Nganji J, Wen L & Butterfield A, "Investigating heuristic evaluation as a methodology for evaluating pedagogical software: An analysis employing three case studies", Proceedings of the 16th International Conference on Human-Computer Interaction (HCI International 2014), 2014, pp. 25-35.
P13	Cassino R, Tucci M, Vitiello G & Francese R, "Empirical validation of an automatic usability evaluation method", Journal of Visual Languages & Computing, vol. 28, 2015, pp. 1-22.
P14	Cavichi de Freitas R, Rodrigues LA & Marques da Cunha A, "AGILUS: A method for integrating usability evaluations on agile software development", Proceedings of the 18th International Conference on Human-Computer Interaction (HCI International 2016), 2016, pp. 542-552.
P15	Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J, Fisk AD, Lee CC & Nair SN, "The personalized reminder information and social management system (PRISM) trial: Rationale, methods and baseline characteristics", Contemporary Clinical Trials, vol. 40, 2015, pp. 35-46.
P16	Chuan NK, Sivaji A & Ahmad WFW, "Proposed usability heuristics for testing gestural interaction", Proceedings of the 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology (ICAIET 2014), 2014, pp. 233-238.
P17	de Carvalho JMI, da Silva TS & Silveira MS, "Agile and UCD integration based on pre-development usability evaluations: An experience report", Proceedings of the 18th International Conference on Human-Computer Interaction (HCI International 2016), 2016, pp. 586-597.
P18	Desurvire H & El-Nasr MS, "Methods for game user research: Studying player behavior to enhance game design", IEEE Computer Graphics and Applications, vol. 33, no. 4, 2013, pp. 82-87.
P19	Dirin A & Nieminen M, "Relevance of UCD education to software development – recommendation for curriculum design", Proceedings of the 8th International Conference on Computer Supported Education (CSEDU 2016), 2016, pp. 112-120.
P20	Duke JD, Morea J, Mamlin B, Martin DK, Simonaitis L, Takesue BY, Dixon BE & Dexter PR, "Regenstrief Institute's Medical Gopher: A next-generation homegrown electronic medical record system", International Journal of Medical Informatics, vol. 83, no. 3, 2014, pp. 170-179.
P21	Faulring A, Myers BA, Oren Y & Rotenberg K, "A case study of using HCI methods to improve tools for programmers", Proceedings of the 5th International Workshop on Co-operative and Human Aspects of Software Engineering (CHASE 2012), 2012, pp. 37-39.
P22	Feldmann F & Filippis MD, "Usability testing for German railway HMIS: A chance to optimize software engineering", Proceedings of the 3rd International Conference on Rail Human Factors, 2012, pp. 273-282.
P23	Fernandez A, Abrahão S, Insfran E & Matera M, "Further analysis on the validation of a usability inspection method for model-driven web development", Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2012), 2012, pp. 153-156.
P24	Fernandez A, Abrahão S, Insfran E & Matera M, "Usability inspection in model-driven web development: Empirical validation in WebML", Proceedings of the 16th International Conference on Model Driven Engineering Languages and Systems (MODELS 2013), 2013, pp. 740-756.
P25	González CS, Toledo P, Muñoz V, Noda MA, Bruno A & Moreno L, "Inclusive educational software design with agile approach", Proceedings of the 1st International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM 2013), 2013, pp. 149-155.
P26	Hasan L, Morris A & Probets S, "E-commerce websites for developing countries - a usability evaluation framework", Online Information Review, vol. 37, no. 2, 2013, pp. 231-251.
P27	Hurtado N, Ruiz M, Orta E & Torres J, "Using simulation to aid decision making in managing the usability evaluation process", Information and Software Technology, vol. 57, no. 1, 2015, pp. 509-526.
P28	Hvannberg ET, "Identifying and explicating knowledge on method transfer: A sectoral system of innovation approach", Universal Access in the Information Society, vol. 14, no. 2, 2015, pp. 187-202.
P29	Inostroza R, Rusu C, Roncagliolo S, Rusu V & Collazos CA, "Developing SMASH: A set of SMArtphone's uSability Heuristics", Computer Standards & Interfaces, vol. 43, 2016, pp. 40-52.
P30	Jiménez C, Rusu C, Roncagliolo S, Inostroza R & Rusu V, "Evaluating a Methodology to Establish Usability Heuristics", Proceedings of the 31st International Conference of the Chilean Computer Science Society (SCCC 2012), 2012, pp. 51-59.
P31	Jiménez C, Rusu C, Rusu V, Roncagliolo S & Inostroza R, "Formal specification of usability heuristics: how convenient it is?", Proceedings of the 2nd International Workshop on Evidential Assessment of Software Technologies (EAST'12), 2012, pp. 55-60.
P32	Kaufmann D, Parry D, Carlsen V, Carter P, Parry E & Westbrook L, "Opportunistically discovering usability requirements for a clinical handover system", Studies in Health Technology and Informatics, vol. 192, 2013, pp. 157-161.
P33	Konstantinidis G, Anastassopoulos GC, Karakos AS, Anagnostou E & Danielides V, "A user-centered, object-oriented methodology for developing health information systems: A clinical information system (CIS) example", Journal of Medical Systems, vol. 36, no. 2, 2012, pp. 437-450.
P34	Lasierra N, Kushniruk A, Alesanco A, Borycki E & García J, "A methodological approach for designing a usable ontology-based GUI in healthcare", Studies in Health Technology and Informatics, vol. 192, no. 1-2, 2013, pp. 1040-1040.
P35	Losada B, Urretavizcaya M, López-Gil J-M & Fernández-Castro I, "Applying usability engineering in InterMod agile development methodology. A case study in a mobile application", Journal of Universal Computer Science, vol. 19, no. 8, 2013, pp. 1046-1065.
P36	Luna H, Mendoza R, Vargas M, Muñoz J, Alvarez FJ & Rodriguez LC, "Using design patterns as usability heuristics for mobile groupware systems", IEEE Latin America Transactions, vol. 13, no. 12, 2015, pp. 4004-4010.

P37	Medina-Flores R & Morales-Gamboa R, "Usability evaluation by experts of a learning management system", IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, vol. 10, no. 4, 2015, pp. 197-203.
P38	Mi N, Cavuoto LA, Benson K, Smith-Jackson T & Nussbaum MA, "A heuristic checklist for an accessible smartphone interface design", Universal Access in the Information Society, vol. 13, no. 4, 2014, pp. 351-365.
P39	Monkman H & Kushniruk A, "A health literacy and usability heuristic evaluation of a mobile consumer health application", Studies in Health Technology and Informatics, vol. 192, 2013, pp. 724-728.
P40	Myers BA, Ko AJ, LaToza TD & Yoon Y, "Programmers are users too: Human-centered methods for improving programming tools", Computer, vol. 49, no. 7, 2016, pp. 44-52.
P41	Nascimento I, Silva W, Gadelha B & Conte T, "Userbility: A technique for the evaluation of user experience and usability on mobile applications", Proceedings of the 18th International Conference on Human-Computer Interaction (HCI International 2016), 2016, pp. 372-383.
P42	Orlandini G, Castadelli GA & Presumido Braccialli LM, "Ergonomics and usability in sound dimension: Evaluation of a haptic and acoustic interface application for mobile devices", Proceedings of the 16th International Conference on Human-Computer Interaction (HCI International 2014), 2014, pp. 193-202.
P43	Paz F, Paz FA, Villanueva D & Pow-Sang JA, "Heuristic evaluation as a complement to usability testing: A case study in web domain", Proceedings of the 12th International Conference on Information Technology - New Generations (ITNG 2015), 2015, pp. 546-551.
P44	Paz F, Villanueva D, Rusu C, Roncagliolo S & Pow-Sang JA, "Experimental evaluation of usability heuristics", Proceedings of the 10th International Conference on Information Technology: New Generations (ITNG 2013), 2013, pp. 119-126.
P45	Pivetta EM, Saito DS, da Silva Flor C, Ulbricht VR & Vanzin T, "Automated accessibility evaluation software for authenticated environments", Proceedings of the 16th International Conference on Human-Computer Interaction (HCI International 2014), 2014, pp. 77-88.
P46	Quiñones D, Rusu C & Roncagliolo S, "Redefining usability heuristics for transactional web applications", Proceedings of the 11th International Conference on Information Technology: New Generations (ITNG 2014), 2014, pp. 260-265.
P47	Rajanen M & Livari N, "Examining usability work and culture in OSS", Proceedings of the 11th IFIP WG 2.13 International Conference on Open Source Systems (OSS 2015), 2015, pp. 58-67.
P48	Rajanen M, Iivari N & Keskitalo E, "Introducing usability activities into open source software development projects: A participative approach", Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordiCHI'12), 2012, pp. 683-692.
P49	Rivero L, Vincenzi A, Maldonado JC & Conte T, "Evaluating software engineers' acceptance of a technique and tool for web usability inspection", Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering (SEKE 2015), 2015, pp. 140-145.
P50	Roisko V, Kämppi P & Luojus S, "Touch screen based TETRA vehicle radio: Preliminary results of multi-methodology usability testing prototype", Proceedings of the 2013 International Conference on Connected Vehicles and Expo (ICCVE), 2013, pp. 941-952.
P51	Rusu C, Roncagliolo S, Figueroa A, Rusu V & Gorgan D, "Evaluating the usability and the communicability of grid computing applications", Proceedings of the 5th International Conference on Advances in Computer-Human Interactions (ACHI 2012), 2012, pp. 204-207.
P52	Saini A, Nanchen B & Evequoz F, "Putting the customer back in the center of SOA with service design and user-centered design", Proceedings of the European Conference on Service-Oriented and Cloud Computing (ESOCC 2013), 2013, pp. 94-103.
P53	Sanches LMP, Harris MR, Abbott PA, Novaes MA & Lopes MHBM, "Collaborative software development for a brazilian telehealth program", Proceedings of the 12th International Congress on Nursing Informatics: East Meets West eSMART+ (NI 2014), 2014, pp. 211-218.
P54	Santana DS, Santos CHdS & Figueroa HEH, "Human-computer interface techniques to design and evaluate an electromagnetic simulator", IEEE Latin America Transactions, vol. 12, no. 4, 2014, pp. 725-732.
P55	Santos CHS & Figueroa HEH, "Free visual FDTD 2D simulator to support the telecommunication teaching-learning process", IEEE Latin America Transactions, vol. 13, no. 3, 2015, pp. 818-824.
P56	Santos R, Werner C, Costa H, Abílio R & Borges H, "Managing reusable learning objects and experience reports in EduSE portal", Proceedings of the IEEE 13th International Conference on Information Reuse and Integration (IRI 2012), 2012, pp. 631-638.
P57	Sedlmair M, Frank A, Munzner T & Butz A, "RelEx: Visualization for actively changing overlay network specifications", IEEE Transactions on Visualization and Computer Graphics, vol. 18, no.12, 2012, pp. 2729-2738.
P58	Shoukry L, Sturm C & Galal-Edeen GH, "Pre-MEGa: A proposed framework for the design and evaluation of preschoolers' mobile educational games", Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering, Lecture Notes in Electrical Engineering, vol. 313, 2015, pp. 385-390.
P59	Sim G, Read JC, Gregory P & Xu D, "From England to Uganda: Children designing and evaluating serious games", Human–Computer Interaction, vol. 30, no. 3-4, 2015, pp. 263-293.
P60	Solano A., Cerón JC, Collazos CA, Fardoun HM, & Arciniegas JL, "ECUSI: A software tool that supports the usability collaborative evaluation of interactive systems", Proceedings of the XVI International Conference on Human Computer Interaction (Interacción '15), 2015, article No. 42.
P61	Soomro S, Ahmad WFW & Sulaiman S, "Evaluation of mobile games with playability heuristic evaluation system", Proceedings of the 2014 International Conference on Computer and Information Sciences (ICCOINS 2014), 2014.
P62	Stråt B, Johansson M & Warpefelt H, "Evaluating game heuristics for measuring player experience", Proocedings of the 14th International Conference on Intelligent Games and Simulation (GAME-ON 2013), 2013.
P63	Stråt B, Rutz F & Johansson M, "Does game quality reflect heuristic evaluation?: Heuristic evaluation of games in different quality strata", International Journal of Gaming and Computer-Mediated Simulations, vol. 6, no. 4, 2014, pp. 45-58.
P64	Valentim NMC & Conte T, "Improving a usability inspection technique based on quantitative and qualitative analysis", Proceedings of the 28th Brazilian Symposium on Software Engineering (SBES 2014), 2014, pp. 171-180.
P65	Vaz VT, Travassos GH & Conte T, "Empirical assessment of WDP tool: A tool to support web usability inspections", Proceedings of 2012 XXXVIII Conferencia Latinoamericana en Informática (CLEI 2012), 2012, pp. 1-9.
P66	Vélez O, Okyere PB, Kanter AS & Bakken S, "A usability study of a mobile health application for rural Ghanaian midwives", Journal of Midwifery & Women's Health, vol. 59, no. 2, 2014, pp. 184-191.
P67	Vogel B, Kurti A, Milrad M & Mikkonen T, "Architectural concepts: Evolution of a software system across design and implementation stages in dynamically changing environments", Proceedings of the IEEE 36th Annual Computer Software and Applications Conference Workshops (COMPSACW 2012), 2012, pp. 144-151.
P68	Yadav K, Chamberlain JM, Lewis VR, Abts N, Chawla S, Hernandez A, Johnson J, Tuveson G & Burd RS, "Designing real-time decision support for trauma resuscitations", Academic Emergency Medicine, vol. 22, no. 9, 2015, pp. 1076-1084.
P69	Yuan MJ, Finley GM, Long J, Mills C & Johnson RK, "Evaluation of user interface and workflow design of a bedside nursing clinical decision support system", Journal of Medical Internet Research, vol. 15, no. 1, 2013, pp. e4.
P70	Zargaran E, Schuurman N, Nicol AJ, Matzopoulos R, Cinnamon J, Taulu T, Ricker B, Garbutt Brown DR, Navsaria P & Hameed SM, "The electronic trauma health record: design and usability of a novel tablet-based tool for trauma care and injury surveillance in low resource settings", Journal of the American College of Surgeons, vol. 218, no. 1, 2014, pp. 41-50.
P71	Zaharias P & Koutsabasis P, "Heuristic evaluation of e-learning courses: A comparative analysis of two e-learning heuristic sets", Campus-Wide Information Systems, vol. 29, no. 1, 2012, pp. 45-60.

4. Analysis of relevant papers

After a complete revision of 1615 papers, we examined only those studies which report the use of the heuristic evaluation method in the software domain. For each relevant paper that was identified by the systematic literature review, we analyzed the following criteria:

- ***Software domain:*** the category of software product that was assessed by the heuristic evaluation method.
- ***Usability principles:*** the set of heuristics that was used to identify usability problems in the graphical user interface.
- ***Evaluators' profile:*** the experience and knowledge of each reviewer who was involved in the evaluation process.
- ***Number of evaluators:*** the number of participants that were involved in the evaluation process.

4.1. Software domain

The first goal of this research was to identify the types of software products that are assessed by the heuristic evaluation method. In Table 2, we present the papers which describe the use of this evaluation technique in a particular software category. According to these results, the heuristic evaluation is more used to evaluate software products that are related to the educational domain. Some examples of this category are learning management systems, mobile learning applications, software tools that are used in classroom as support for the teaching of a topic, serious games whose main purpose is the education, educational software for people with special abilities, e-learning platforms, content management systems, virtual learning environments socio-pedagogical monitoring systems, and tutor systems.

The results establish, furthermore, that the heuristic evaluation is also frequently used to measure the usability of software products related to the field of health informatics, e-commerce, and gaming. In some cases, the method is performed with the use of the conventional heuristics proposed by Nielsen [14]. In other cases, new specific proposals of heuristic principles are employed to assess usability aspects of the software domain that are not considered by the traditional approach. This matter is discussed in Section 5.2.

We highlight the fact that some papers can appear in more than one domain. This classification is due to two situations: (1) the software product that was evaluated can be categorized in two or more domains, for instance, serious games can be considered in the domain of gaming, however, they have a solid component of education, or (2) the study report the heuristic evaluation of two or more different types of software products.

Table 2: Number of studies that report the use of the heuristic evaluation method in a particular category of software application

Domain	Number of studies	Papers
Education	17	[P01] [P03] [P05] [P06] [P12] [P14] [P15] [P19] [P25] [P28] [P37] [P39] [P55] [P56] [P58] [P67] [P71]
Health Informatics	10	[P07] [P20] [P32] [P33] [P34] [P53] [P66] [P68] [P69] [P70]
E-Commerce	10	[P20] [P23] [P24] [P26] [P31] [P41] [P43] [P44] [P46] [P52]
Gaming	7	[P11] [P18] [P48] [P59] [P61] [P62] [P63]
Personal Information Manager	5	[P14] [P23] [P24] [P29] [P41]
Business Management	3	[P08] [P40] [P64]
Software Development Tool	3	[P21] [P45] [P60]
Social Network	3	[P03] [P09] [P15]
Support Software for Industrial Processes	2	[P49] [P57]
Accessible Applications	2	[P38] [P42]
Project Management Tool	2	[P02] [P17]
Software for Physics	2	[P54] [P55]
Software for Transportation Sector	1	[P50]
Banking Software	1	[P04]
Geographic Information System	1	[P51]
Software for Users' Communication	1	[P09]
Financial Software	1	[P65]
Web Search Engine	1	[P13]
Personal Guide System	1	[P35]
TOTAL	**73**	

4.2. Usability principles

The heuristic evaluation method establishes the use of a set of principles to verify that the system interface is usable. The most recognized heuristics in the literature are those who were designed by Nielsen [14], the same author of the technique. The ten usability heuristics for user interface design are widely known and used in the execution of this method. However, in the last years, new usability principles have been developed to evaluate software applications of specific domains [5]. This fact is because the traditional heuristics fail to cover those features that the new types of software products include in their interface design [22]. When the conventional approach of Nielsen is used to evaluate the usability of emerging categories of software, the results are inaccurate [21]. Depending on the kind of software, the features change and in the same way, the usability aspects to evaluate. In this research, we analyzed if the authors still continue using the ten principles proposed by Nielsen, or if new approaches are considered for the evaluation process. In Table 3, we present the number of studies which report the use of both, Nielsen's usability heuristics and different proposals. Although most studies consider that the general guidelines of Nielsen are enough to perform a heuristic evaluation, a significant number of papers (29 out of 71) describe the use of new heuristics that were designed to evaluate new software domains. In Table 4, we describe the proposal that was used in those works which report the use of a different set of principles to the traditional ones.

Table 3: Number of studies that report the use of a specific type of heuristics

Set of heuristics	Number of studies	Papers
Nielsen's usability heuristics	33	[P01] [P02] [P03] [P04] [P05] [P07] [P08] [P10] [P15] [P17] [P19] [P21] [P23] [P24] [P25] [P31] [P32] [P33] [P40] [P41] [P42] [P43] [P44] [P49] [P50] [P52] [P54] [P55] [P56] [P60] [P64] [P68] [P69]
Other proposals of usability heuristics	29	[P09] [P11] [P12] [P13] [P16] [P18] [P22] [P26] [P28] [P29] [P30] [P35] [P36] [P37] [P38] [P39] [P45] [P46] [P48] [P51] [P58] [P59] [P61] [P62] [P63] [P65] [P66] [P70] [P71]
Not mentioned	9	[P06] [P14] [P20] [P27] [P34] [P47] [P53] [P57] [P67]
TOTAL	**71**	

Table 4: New proposals of usability heuristics that were considered in each study

ID	Set of usability heuristic that was used
P09	Heuristics proposed by Bertini et al. to evaluate the usability of applications for mobile devices. • Bertini E, Gabrielli S & Kimani S, "Appropriating and assessing heuristics for mobile computing", Proceedings of the 8th International Working Conference on Advanced Visual Interfaces (AVI 2006), 2006, pp. 119-126.
P11	The authors of this paper determined a set of heuristics to assess the usability of a massively multiplayer online role-playing game. This proposal is based on the study conducted by Song et al. • S. Song, J. Lee, and I. Hwang, "A new framework of usability evaluation for massively multi-player online game: case study of "World of Warcraft" game", Proceedings of the 12th International Conference on Human-Computer Interaction (HCI International 2007), 2007, pp. 341-350.
P12	A set of heuristic principles proposed by Squires & Preece to evaluate the usability of pedagogical software. This evaluation approach for eLearning systems is based on the Nielsen's traditional guidelines. • Squires D & Preece J, "Predicting quality in educational software: Evaluating for learning, usability and the synergy between them", Interacting with Computers, vol. 11, no. 5, 1999, pp. 467-483.
P13	The authors of this paper derived a list of fifteen usability properties that should be examined during the evaluation of an interactive software application. This proposal is entirely detailed in the same paper [P13]. This new approach is based on Nielsen's heuristics and the work performed by Hvannberg. • Hvannberg ET, Law EL & Lárusdóttir MK, "Heuristic evaluation: Comparing ways of finding and reporting usability problems", Interacting with Computers, vol. 19, no. 2, 2007, pp. 225.240.
P16	The authors of this study proposed new usability heuristics to evaluate gestural interactions of software applications for touch screen devices. The list of principles is presented in the paper [P16]. This proposal is based on the researches performed by Ryu et al. and Norman et al. • Ryu T, Lee J, Yun MH & Lim JH, "Conditions of applications, situations and functions applicable to gesture interface", Proceedings of the 15th International Conference on Human-Computer Interaction (HCI International 2013), 2013, pp. 368-377. • Norman D & Nielsen J, "Gestural interfaces: a step backward in usability", Interactions, vol. 17, no. 5, 2010, pp. 46-49.
P18	A set of principles proposed by Desurvire & Wiberg to evaluate the usability of videogames. The authors make reference to two studies: • Desurvire H & Wiberg C, "Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration", Proceedings of the 13th International Conference on Human-Computer Interaction (HCI International 2009), 2009, pp. 557-566. • Desurvire H & Wiberg C, "Master of the game: Assessing approachability in future game design", Proceedings of the 26th CHI Conference on Human Factors in Computing Systems (CHI '08), 2008, pp. 3177-3182.
P22	The authors developed a set of heuristics on the basis of the standard ISO 9241-110. However, these new principles are not described in the paper. • Ergonomics of human-system interaction -- Part 110: Dialogue principles, ISO Std. 9241-11:2006, International Organization for Standardization, 2006.
P26	The authors devised a set of heuristics to evaluate e-commerce websites specifically. This new proposal was based on the Nielsen's traditional approach and the studies performed by Elliot et al. and Pearson et al. Nevertheless, the new set of principles are not detailed. • Elliot S, Mørup-Petersen A & Bjørn-Andersen N, "Towards a framework for evaluation of commercial web sites", Proceedings of the 13th International Bled Electronic Commerce Conference, 2000. • Pearson JM, Pearson A & Green D, "Determining the importance of key criteria in web usability", Management Research News, vol. 30, no. 11, 2007, pp. 816-828.
P28	Heuristics proposed by Sutcliffe & Gault to evaluate the usability of virtual reality software applications. • Sutcliffe A & Gault B, "Heuristic evaluation of virtual reality applications", Human Computer Interaction in Latin America, vol. 16, no. 4, 2004, pp. 831-849.
P29	The authors of this paper developed a new set of heuristics to evaluate the usability of software applications for touchscreen-based mobile devices. The list of principles is described entirely in the same paper [P29]. This new proposal is based on the Nielsen's traditional approach.
P30	The authors of this paper assessed four different categories of software applications. For each type, a particular set of usability heuristics was used. 1. Heuristics proposed by Roncagliolo et al. to evaluate the usability of grid computing-based applications. • Roncagliolo S, Rusu V, Rusu C, Tapia G, Hayvar D & Gorgan D, "Grid computing usability heuristics in practice", Proceedings of the 8th International Conference on Information Technology: New Generations (ITNG 2011), 2011, pp. 145-150. 2. Heuristics proposed by Solano et al. to evaluate the usability of smart TV applications. • Solano A, Rusu C, Collazos C, Roncagliolo S, Arciniegas JL, & Rusu V, "Usability heuristics for interactive digital television", Proceedings of the 3rd International Conference on Advances in Future Internet (AFIN 2011), 2011, pp. 60-63. 3. Heuristics proposed by Inostroza to evaluate the usability of applications for touchscreen mobile devices. • Inostroza R, Rusu C, Roncagliolo S, Jimenez C & Rusu V, "Usability heuristics for touchscreen-based mobile devices", Proceedings of the 9th International Conference on Information Technology: New Generations (ITNG 2012), 2012, pp. 662-667. 4. Heuristics proposed by Muñoz & Rusu to evaluate the usability of virtual worlds. • Muñoz R & Rusu C, "Virtual worlds: Real usability?", Proceedings of the 5th Latin American Conference on Human Computer Interaction (CLIHC 2011), 2011.
P35	Usability heuristics proposed by Bertini to evaluate mobile applications. • Bertini E, Catarci T, Dix A, Gabrielli S, Kimani S & Santucci G, "Appropiating heuristic evaluation for mobile computing", International Journal of Mobile Human Computer Interaction, vol. 1, no. 1, 2011, pp. 20-41.
P36	The authors of this study proposed a new set of heuristics to evaluate the usability of groupware software applications for mobile devices. This new proposal is completely detailed in the same paper [P36]. The new principles are based on the researches performed by Ellis et al. and Johansen. • Ellis CA, Gibbs SJ & Rein G, "Groupware: Some issues and experiences", Communications of the ACM, vol. 34, no. 1, 1991, pp. 39-58. • Johansen R, "Groupware: Future directions and wild cards", Journal of Organizational Computing, vol. 1, no. 2, 1991, pp. 219-227.
P37	The authors of this paper established a new set of criteria to evaluate the usability of learning management systems. The list is completely described in the same paper [P37]. This new proposal is based on an analysis of the characteristics of this type of software systems.
P38	A heuristic checklist proposed by the authors of the paper to evaluate the usability and accessibility of a software application for a smartphone [P38].
P39	The authors generated a new set of heuristics based on a guide for the development of health literacy online. The principles are described in the same study [P39]. • Health literacy online: A guide to writing and designing easy-to-use health web sites, U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion,

	2010.
P45	The authors developed a set of heuristics to evaluate the usability and accessibility of a software product. The principles are detailed in the same study [P45].
P46	A set of heuristics to evaluate the usability of transactional web applications. This new proposal was developed by Paz et al. • Paz F, Paz FA, Pow-Sang JA & Collantes L, "Usability heuristics for transactional web sites", Proceedings of the 11th International Conference on Information Technology: New Generations (ITNG 2014), 2014, pp. 627-628.
P48	Usability heuristics proposed by Pinelle et al. and Desurvire et al. to evaluate videogames. • Pinelle D, Wong N & Stach T, "Heuristic evaluation for games: Usability principles for video game design", Proceedings of the 26th CHI Conference on Human Factors in Computing Systems (CHI '08), 2008, pp. 1453-1462. • Desurvire H, Caplan M & Toth J, "Using heuristics to evaluate the playability of games", Proceedings of the 22th CHI Conference on Human Factors in Computing Systems (CHI '04), 2004, pp. 1509-1512.
P51	A new set of principles proposed by Roncagliolo et al. to evaluate the usability of grid computing-based applications. • Roncagliolo S, Rusu V, Rusu C, Tapia G, Hayvar D & Gorgan D, "Grid computing usability heuristics in practice", Proceedings of the 8th International Conference on Information Technology: New Generations (ITNG 2011), 2011, pp. 145-150.
P58	The authors developed a new set of heuristics to evaluate the usability of mobile educational games for preschoolers. The approach was based on a literature review of the available tools to assess videogames and educational software. The list of principles is presented in the same paper [P58].
P59	Two set of heuristics proposed by Malone and Alsumait et al. to evaluate the usability of videogames. • Malone TW, "What makes things fun to learn? Heuristics for designing instructional computer games", Proceedings of the 3rd ACM SIGSMALL Symposium and the 1st SIGPC Symposium on Small Systems, 1980, pp. 162-169. • Alsumait A & Al-Osaimi A, "Usability heuristics evaluation for child e-learning applications", Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services (iiWAS '09), 2009, pp. 425-430.
P61	Usability heuristics proposed by Korhonen & Koivisto to evaluate mobile games. • Korhonen H & Koivisto EMI, "Playability heuristics for mobile games", Proceedings of the 8th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '06), 2006, pp. 9-16.
P62	A list of game playability principles developed by the authors from a contemporary literature review [P62]. These principles were based on the approach proposed by Desurvire & Wiberg. • Desurvire H & Wiberg C, "Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration", Proceedings of the 13th International Conference on Human-Computer Interaction (HCI International 2009), 2009, pp. 557-566.
P63	A list of heuristics proposed by Stråat et al. to evaluate the usability of videogames. • Stråat B, Johansson M & Warpefelt H, "Evaluating game heuristics for measuring player experience", Proceedings of the 14th annual European Conference on Simulation and Artificial Intelligence in Computer Games (GAMEON 2013), 2013, pp. 15-19.
P65	The authors proposed a new set of usability aspects based on the combination of the Nielsen's traditional approach and the work performed by Zhang et al. The aspects are presented in the same paper [P65]. • Z. Zhang, V. Basili, and B. Shneiderman, "Perspective-based usability inspection: An empirical validation of efficacy", Empirical Software Engineering, vol. 4, no. 1, 1999, pp. 43-69.
P66	A new set of heuristics proposed by Bertini et al. to evaluate the usability of software applications for mobile devices. • Bertini E, Gabrielli S & Kimani S, "Appropriating and assessing heuristics for mobile computing", Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '06), 2006, pp. 119-126.
P70	The authors established a set of usability attributes that should be considered in the heuristic evaluation of an electronic health record. These characteristics are described in the same paper [P70]. This proposal is based on the Nielsen's traditional approach.
P71	Two heuristic sets proposed by Reeves et al. and Mehlenbacher et al. to evaluate e-learning systems. • Reeves TC, Benson L, Elliott D, Grant M, Holschuh D, Kim B, Kim H, Lauber E, & Loh CS, "Usability and instructional design heuristics for e-learning evaluation", Proceedings of the 14th World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2002, pp. 2-8. • Mehlenbacher B, Bennett L, Bird T., Ivey M, Lucas J, Morton J & Whitman L. "Usable e-learning: A conceptual model for evaluation and design", Proceedings of the 11th International Conference on Human-Computer Interaction (HCI International 2005), 2005, pp. 1-10.

4.3. Profile of the evaluators

According to the original definition of heuristic evaluation proposed by Nielsen [16], the assessment process must necessarily be carried out by usability specialists. However, given the complexity of having the participation of this type of professionals, many authors have opted to involve evaluators from a different profile. In more recent research, Nielsen establishes that the process is influenced by the skills of the reviewers [15]. The ideal evaluator to obtain accuracy results would be a double expert at usability and the domain of the application [1]. For instance, if the purpose of the evaluation is an e-learning platform, a double expert would be a professional with enough knowledge in usability and education at the same time [9]. However, the majority of the studies that were analyzed report the participation of software specialists, professionals who are familiar with the software development process but with little or no experience in the field of usability neither the domain. In order to compensate for this situation, the software professionals are trained in the main concepts of usability and the heuristic evaluation process. Finally, there is a small group of authors who have mixed the participation of types of evaluators. In these studies, half of the participants were usability specialists and the other half experts in the domain. Table 5 summarizes the studies that describe the participation of a particular profile of evaluators.

Table 5: Number of studies that report the involvement of a specific type of specialist

Profile of the evaluators	Number of studies	Papers
Usability specialists	15	[P02] [P04] [P08] [P13] [P22] [P30] [P34] [P35] [P43] [P47] [P51] [P52] [P53] [P57] [P60]
Software professionals	21	[P05] [P09] [P10] [P16] [P17] [P19] [P23] [P24] [P29] [P31] [P33] [P41] [P42] [P44] [P45] [P46] [P48] [P49] [P61] [P64] [P65]
Experts in the domain	9	[P01] [P11] [P26] [P36] [P37] [P38] [P56] [P66] [P70]
Usability Specialists + Experts in the domain	4	[P20] [P25] [P54] [P69]
Double experts	4	[P03] [P12] [P68] [P71]
Not mentioned	18	[P06] [P07] [P14] [P15] [P18] [P21] [P27] [P28] [P32] [P39] [P40] [P50] [P55] [P58] [P59] [P62] [P63] [P67]
TOTAL	**71**	

4.4. Number of evaluators

The original proposal of the heuristic evaluation establishes the participation of three to five usability specialists. As stated by Nielsen [15], this amount of reviewers is enough to identify 75% of the total number of usability problems that are present in an interface design. The benefits of involving more than five evaluators are minor because many of the additional usability issues to be identified would only be cosmetic (with slight relevance in the usability of the product). Furthermore, the hiring of another expert would suppose an extra cost to the project, unbalancing the cost-benefit ratio. On the other hand, the participation of only one or two specialists would not let the identification of the most severe usability problems.

The analysis of the studies establishes that most authors follow the suggestion of Nielsen of involving the participation of three to five evaluators. However, there is a small percentage of scholars who do not adhere to the standard and include in the evaluation a diverse number of inspectors. Table 6 summarizes the studies which describe the involvement of a specific number of reviewers in the evaluation process.

Table 6: Studies that report the participation of a specific number of evaluators

Number of evaluators	Number of studies	Papers
From 1 to 2	10	[P07] [P08] [P22] [P34] [P39] [P42] [P53] [P59] [P62] [P71]
From 3 to 5 (Nielsen's approach)	27	[P02] [P03] [P04] [P09] [P10] [P13] [P17] [P26] [P29] [P31] [P33] [P35] [P43] [P44] [P45] [P46] [P47] [P48] [P52] [P54] [P56] [P57] [P60] [P61] [P63] [P68] [P69]
From 6 to 10	10	[P01] [P11] [P12] [P24] [P25] [P37] [P38] [P41] [P64] [P65]
From 11 to 14	1	[P36]
From 15 to more	3	[P23] [P49] [P70]
Not mentioned	20	[P05] [P06] [P14] [P15] [P16] [P18] [P19] [P20] [P21] [P27] [P28] [P30] [P32] [P40] [P50] [P51] [P55] [P58] [P66] [P67]
TOTAL	**71**	

5. Usability heuristic evaluation protocol

After a comprehensive analysis of the evaluation processes that are reported in the 71 relevant studies, we have developed a protocol to perform heuristic evaluations in the context of software products. The authors do not agree on how to carry out the evaluation process. Some of them perform activities that are not reported by others. In this work, we provide a formal procedure to allow specialists from different fields to conduct heuristic evaluations in a systematic and structured way. Based on the review of all the studies, we have divided the process into five phases: (1) planning, (2) training, (3) evaluation, (4) discussion and (5) report. The structure of this protocol was developed according to the approach proposed by Hurtado et al. [7] for usability inspections. The sequence of our procedure is illustrated in Figure 1.

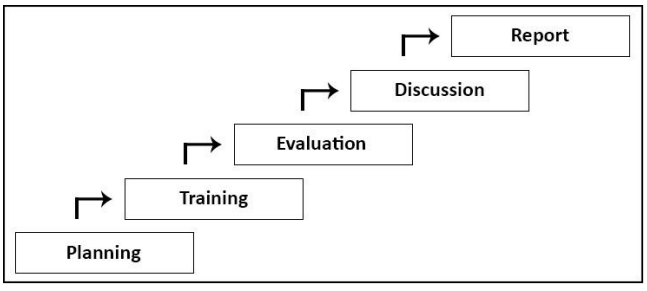


Fig. 1: New evaluation proposal to conduct heuristic evaluations

Each of the phases of this protocol is described as follows. This proposal is a general framework based on the assessment processes that are detailed in those studies which report the use of the heuristic evaluation as evaluation method in the context of the software development process.

- **Planning:** The first phase of this protocol involves the planning of the evaluation session, the selection of the appropriate evaluators, the choice of the heuristics that will be employed, the institution of the criteria that will be assumed to conduct the evaluation, and the materials that will be used. This phase is supported by the following papers that report the implementation of the mentioned activities: [P05] [P12] [P22] [P23] [P25] [P26] [P28] [P32] [P36] [P37] [P38] [P42] [P45] [P65] [P71].

1. The evaluation manager should establish the number of evaluators who will participate in the heuristic inspection. Although Nielsen recommends the participation of three to five reviewers, the analysis shows that many authors involve the participation of a varied number of evaluators that range between one to twenty-two specialists (Table 6).
2. The evaluation manager must determine the profile of the reviewers that will be part of the assessment team. According to the analysis, the participants can be usability specialists, experts in the domain, double experts or software professionals. It is also possible to have professionals with a different profile in the same evaluation group as the Table 5 describes.
3. The evaluation manager must prepare a general description of the software product that will be examined regarding usability. This information must be given to each of the evaluators that will form the inspection team.
4. Finally, the evaluation manager should recruit the inspection team according to the criteria that were previously established as the number and profile of the reviewers.

- **Training:** In this phase, those evaluators, with little or no experience in the conduction of heuristic evaluations, will receive training in the execution of the inspection method. The evaluation manager will explain the process, the goals, and the criteria that must be followed. These activities were proposed on the basis of the following papers: [P12] [P18] [P22] [P23] [P25] [P27] [P28] [P32] [P36] [P38] [P42] [P45] [P65] [P66] [P71].

5. If the evaluation team has little experience, each evaluator should receive training in the main concepts of usability and heuristic evaluation, as well as in the assessment process.
6. Each member of the evaluation team must be informed about the protocol which will be used to carry out the heuristic inspection.
7. All evaluators should collaboratively establish the set of heuristic principles that will be employed for the usability evaluation. This decision should be taken in accordance with the nature of the software product. In Table 3, we describe some new proposals that were used for specific domains. This aspect becomes highly relevant given that Nielsen's traditional guidelines fail to cover usability issues that new categories of software present.
8. (Optional) The evaluators can perform a free exploration of the system interface as a previous step to the usability evaluation. In this analysis, they can establish users' profiles, scenarios and in addition, a list of tasks that can be used to guide the inspection. The evaluators can assume these profiles during the examination of the interfaces, and perform the list of tasks as if they were real users of the system.

- **Evaluation:** In this phase, each of the specialists should examine individually the system interface. This activity is sup-

ported by the following papers: [P01] [P04] [P07] [P08] [P09] [P10] [P12] [P13] [P15] [P16] [P17] [P18] [P19] [P21] [P22] [P23] [P25] [P27] [P28] [P29] [P30] [P31] [P32] [P33] [P35] [P36] [P37] [P38] [P39] [P41] [P42] [P43] [P44] [P45] [P46] [P48] [P50] [P53] [P54] [P56] [P59] [P60] [P61] [P62] [P64] [P65] [P66] [P69] [P71].

9. Each evaluator should work individually to determine whether all graphical user interfaces are aligned to the usability heuristics. This evaluation process can be performed by a free examination of the system interface or by following pre-defined tasks. There are two ways to carry out the inspection. (1) The evaluators can identify usability problems heuristic by heuristic or (2) the evaluators can read all the heuristics and identify problems for any of them. A usability problem is an infringement of the heuristic principle. If the interface design does not follow the usability guidelines, these issues should be classified as identified problems. The evaluators can use templates to describe each problem.

- **Discussion:** Once each evaluator has identified individually a list of usability problems, the evaluation team should consolidate all the problems in a single list. For this purpose, a meeting should be organized in which the team should discus if each identified problem is indeed a usability problem, determine if the problems were identified by more than one evaluator, and determine if the problem as written describes the issue. These activities are based on the process established by the papers: [P02] [P03] [P04] [P07] [P10] [P12] [P13] [P23] [P25] [P26] [P27] [P28] [P31] [P33] [P35] [P36] [P37] [P38] [P39] [P42] [P43] [P44] [P45] [P46] [P52] [P54] [P56] [P61] [P64] [P65] [P66] [P71].

- **Report:** The evaluators should offer solutions to each usability problem that was identified. Finally, the results of this evaluation should be reported in a formal document. Both activities are supported by the following papers: [P10] [P26] [P27] [P38] [P42] [P44] [P45] [P46] [P48] [P52] [P54] [P66] [P69].

6. Conclusions and future works

The heuristic evaluation is currently a wide used method to evaluate the usability of the software applications. Although it is one of the most recognized techniques, there is not a formal protocol to carry out the evaluation process. This study evidences the different ways the heuristic evaluation is conducted. An analysis of the papers which report the use of this method in the software development process has allowed us to establish a proper and systematic procedure to execute the assessment.

This research is based on a previous work, in which a systematic literature review was performed. In the review, we identified the most recent studies that report the use of a method to measure the usability of software products in a context of development. This review was updated by running once more the search string in the scientific databases. The search retrieved 1615 studies. However, in this instance, only those papers which describe the use of the heuristic evaluation were selected for analysis. A total of 71 papers were chosen to determine the arrangements each author establish to execute this method.

The results establish that the heuristic evaluation method is used to assess software products from different domains. This technique is commonly employed to evaluate educational software, applications of medical informatics, e-commerce and gaming. In the same way, we have determined that many recommendations of the original author of the technique are followed. Most studies report the participation of three to five specialists, the use of the Nielsen's traditional heuristics and the involvement of usability specialists. However, a representative number of authors have proposed new settings to carry out the inspection. Based on the different ways to execute the evaluation process, we have developed a formal protocol. Nevertheless, it is still necessary to validate this evaluation procedure in several scenarios.

References

[1] Davids MR, Chikte U, Grimmer-Somers K, Halperin ML, "Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders", British Journal of Educational Technology, vol. 45, no. 2, 2014, pp. 367-381.

[2] Desurvire H & Wiberg C, "Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration", Proceedings of the 13th International Conference on Human-Computer Interaction (HCI International 2009), 2009, pp. 557-566.

[3] Díaz J, Rusu C & Collazos CA, "Experimental validation of a set of cultural-oriented usability heuristics: e-commerce websites evaluation", Computer Standards & Interfaces, vol. 50, 2017, pp. 160-178.

[4] Fernandez A, Insfran E & Abrahão S, "Usability evaluation methods for the web: A systematic mapping study", Information and Software Technology, vol. 53, no. 8, 2011, pp. 789-817.

[5] Hermawati S & Lawson G, "Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus?", Applied Ergonomics, vol. 56, 2016, pp. 34-51.

[6] Holzinger A, "Usability engineering methods for software developers", Communications of the ACM, vol. 48, no. 1, 2005, pp. 71-74.

[7] Hurtado N, Ruiz M, Orta E & Torres J, "Using simulation to aid decision making in managing the usability evaluation process", Information and Software Technology, vol. 57, no. 1, 2015, pp. 509-526.

[8] Inostroza R, Rusu C, Roncagliolo S, Rusu V & Collazos CA, "Developing SMASH: A set of SMArtphone's uSability Heuristics", Computer Standards & Interfaces, vol. 43, 2016, pp. 40-52.

[9] Isbandiputra IN, Santoso HB & Hasibuan ZA, "Usability evaluation of Indonesia open educational resources using multi-methods", Journal of Engineering and Applied Sciences, vol. 12, no. 12, 2017, pp. 3272-3276.

[10] ISO, Ergonomic requirements for office work with visual display terminals (vdts), -- Part 11: Guidance of usability, ISO Std. 9241-11:1998, International Organization for Standardization, 1998.

[11] ISO, Software engineering – Product quality – Part 1: Quality model, ISO Std. 9126-1:2001, International Organization for Standardization, 2001.

[12] Ivory MY & Hearst MA, "The state of the art in automating usability evaluation of user interfaces", ACM Computing Surveys, vol. 33, no. 4, 2001, pp. 470-516.

[13] Nagpal R, Mehrotra D & Bhatia PK, "Analytical modelling approach to measure the usability of website", International Journal of Software Engineering and Its Applications, vol. 10, no. 5, 2016, pp. 125-142.

[14] Nielsen J, 10 Usability Heuristics for User Interface Design, 1995, available online: https://www.nngroup.com/articles/ten-usability-heuristics/

[15] Nielsen J, Usability engineering, 1st ed., San Diego, CA, USA: Academic Press, 1993.

[16] Nielsen J, "Usability inspection methods", Proceedings of the Conference Companion on Human Factors in Computing Systems (CHI'94), 1994, pp. 413-414.

[17] Paz F, Paz FA & Pow-Sang JA, "Experimental case study of new usability heuristics", Proceedings of the 17th International Conference on Human-Computer Interaction (HCI International 2015), 2015, pp. 212-223.

[18] Paz F, Paz FA, Villanueva D & Pow-Sang JA, "Heuristic evaluation as a complement to usability testing: A case study in web domain", Proceedings of the 12th International Conference on Information Technology: New Generations (ITNG 2015), 2015, pp. 546-551.

[19] Paz F & Pow-Sang JA, "A systematic mapping review of usability evaluation methods for software development process", International Journal of Software Engineering and Its Applications, vol. 10, no. 1, 2016, pp. 165-178.

[20] Paz F & Pow-Sang JA, "Current trends in usability evaluation methods: A systematic review", Proceeding of the 7th International Conference on Advanced Software Engineering & Its Applications (ASEA 2014), 2014, pp. 11-15.

[21] Paz F, Villanueva D, Rusu C, Roncagliolo S & Pow-Sang JA, "Experimental evaluation of usability heuristics", Proceedings of the 10th International Conference on Information Technology: New Generations (ITNG 2013), 2013, pp. 119-126.

[22] Ten AC, Paz F & Pow-Sang JA, "A framework for the evaluation of user experience in information-driven websites", Journal of Engineering and Applied Sciences, vol. 12 (special issue 5), 2017, pp. 7093-7097.
[23] Whitefield A, Wilson F & Dowell J, "A framework for human factors evaluation", Behaviour & Information Technology, vol. 10, no. 1, 1991, pp. 65-79.