A Perspective-based Usability Inspection for ERP Systems

Joelma Choma[1], Diego H. Quintale [1], Luciana A. M. Zaina[1] and Daniela Beraldo[1, 2]
[1] *Department of Computing, UFSCar, Sorocaba, São Paulo, Brazil*
[2]*Indiana State University, Terre Haute, IN, U.S.A.*

Keywords: Usability Inspection, Heuristic Evaluation, ERP System.

Abstract: The inspection methods to evaluate the usability of ERP systems require more specific heuristics and most suitable criteria into this field. This article proposes a set of heuristics based on perspectives of presentation, and task support aiming to facilitate the inspection of usability in ERP systems especially for novice inspectors. An empirical study was conducted to verify the efficiency and effectiveness of inspections conducted with the proposed heuristics. The results indicate the efficiency and effectiveness to detect problems, mainly in medium-fidelity prototypes of ERP modules.

1 INTRODUCTION

The development of Enterprise Resource Planning systems (ERP) has always been directed to fit a very heterogeneous target market with the crucial goal of integration data for a variety of business processes. This effort on the requirement of data reliability has primarily focused the development of ERP systems on the functionality to properly meet complex business processes; therefore, little attention has been given to usability aspects.

ERP systems require an adequate structure to withstand an enormous amount of data and functionality. Uflacker & Busse (2007) point out that the user interfaces should be developed in order to match specific functional requirements, but at the same time being more adherent to the usability attributes for meeting the user needs, supporting their tasks, and contributing to the easiness of use.

A variety of studies addressing the human-computer interaction aspect in ERP systems have investigated the potential causes of usability problems, nevertheless few contributions concerning inspection methods for properly evaluation of such systems have been found in the literature (Scholtz, et al., 2013; Lucas, et al., 2013; and Lambeck, et al., 2014). The usability inspection can be applied as preventive action allowing the identification of a large number of interaction problems which could adversely affect the user performance, at a relatively low cost. Among the several techniques that categorises the usability problems, the set of heuristics proposed by Jacob Nielsen is one of the most used to support the heuristic evaluation (HE) (Fernandez, et al., 2011).

Based on the authors' findings, (i) we have investigated usability problems of different market ERP systems focusing on discovering new usability issues. Furthermore, (ii) we have conducted a survey in order to explore what is the relevance of the usability aspects for users during the interaction in ERP systems. From the preliminary outcomes (i) and (ii), we proposed a method to guide the usability inspection supported by the ten Nielsen's heuristics under the perspectives of presentation and task support. Our works have not the goal of creation new heuristics; we have fitted the Nielsen's heuristics on the perspectives which we have called perspective-based ERP heuristics. The perspective-based inspection improves the understanding of novice inspectors and developers on the key concepts within the domain and maximizes the detection of usability issues (Zhang et al, 1999). In this paper we aim at validating the proposed perspective-based ERP heuristics, answering the following research questions (RQ):

- RQ1: Can the proposed perspective-based ERP heuristics (on the perspective of presentation and task support) improve the performance of novice inspectors?
- RQ2: Can the perspectives of presentation and of task support effectively identify the major usability issues in ERP systems?
- RQ3: What are the inspectors' opinions about

Choma J., Quintale D., A. M. Zaina L. and Beraldo D..
A Perspective-based Usability Inspection for ERP Systems.
DOI: 10.5220/0005346700570064
In *Proceedings of the 17th International Conference on Enterprise Information Systems* (ICEIS-2015), pages 57-64
ISBN: 978-989-758-098-7

the inspection of interactive interfaces regarding the aspects of ease of use; ease of understanding and usefulness of heuristic evaluation technique applied in ERP systems?

In this direction, we have conducted an experimental study to observe the indicators of efficiency and efficacy of two groups of inspectors who evaluated ERP systems interfaces – the first group used the essential Nielsen's heuristics and the second group used the perspective-based ERP heuristics. Both groups were further subdivided to inspect two different artefacts: a medium-fidelity prototype and a functional module. Then, we analyzed the experimental study answering the RQ and raising conclusions about the proposal.

The rest of the paper is organized as follows: Section 2 introduces related work about heuristics, usability criteria and the heuristic evaluation in ERP systems; Section 3 presents the proposal of perspective-based ERP heuristics; Section 4 describes the experimental study, the data analysis, and outcomes discussion; and Section 5 presents conclusion and future work.

2 RELATED WORK

For Human-Computer Interaction (HCI) community heuristic is defined as a general principle or rule used for both, to forward a decision in the process of designing an interactive system, and to support the critical analysis of a design decision already performed. To supply the heuristic-based evaluation process, usability experts use the HE that is an inspection technique to evaluate the usability of interactive systems based on analysis of a set of heuristics previously defined (Hollingsed & Novick, 2007).

Traditionally, the ten heuristics defined by Jacob Nielsen are the most used to guide the design of interactive interfaces (Nielsen, 1995). However, the Nielsen's heuristics are known as general rules rather than specific usability guidelines, since they are not entirely suitable to address for particular use's characteristics of different interactive systems (Mirel & Wright, 2009). In these cases, the definition of an adherent set of heuristics can often fulfill the particularities of the system domain (Singh & Wesson, 2009). Virtual reality applications (Sutcliffe & Gault, 2004), systems in bioinformatics (Mirel & Wright, 2009), and web applications (Conte et al, 2009) are examples of works that proposed specific heuristics for a domain of system. Other possibilities to ease up the detection of usability problems would be adding criteria to known heuristics gathered from the study of the application domain (Singh & Wesson, 2009); or systematic steps easing the cognitive load of inspectors (Law & Hvannberg, 2004). Aiming at easing the cognitive load of inspectors Zhang et al. (1999) proposed one inspection technique based on perspective of usability. According to the proposal, each usability perspective has its own usability objectives and a list of general usability issues that the inspector should consider during the inspection. For the authors, the general issues can be fitted regarding the characteristics of a particular type of interaction.

Efforts have been undertaken regarding the usability evaluation in ERP systems. While Oja & Lucas (2010) explored qualitative methods of evaluation involving end-users in the usability tests; Scholtz et al. (2013) conducted a experimental study aiming to check the most appropriate criteria for evaluating of medium-sized ERP systems in higher education domain; and Faisal et al. (2012) concentrated the attention on the identification of other usability impacts in another market segment ERP systems. Singh & Wesson (2009) have mapped the recurrent usability issues of ERP systems to the most common usability criteria used to evaluate such systems. They reduced the fourteen criteria they had previously identified to five criteria which were the basis for proposing five specific heuristics to ERP system: navigation, presentation, learnability, task support, and customization.

3 PROPOSAL

Motivated by the work of Zhang et al. (1999) our proposal adopts the ten Nielsen's heuristics by ERP perspectives, instead of adding new heuristics or another interpretation of them, as suggested by Singh & Wesson (2009). We chose to use Nielsen's heuristics because they are widely known by user experience and usability experts. Singh & Wesson (2009) state that the heuristics proposed by them can identify problems significantly different than Nielsen's heuristics can do. However, the tests that were performed with their proposal were made with a small number of experts, not bringing sufficient information and results that could support the experiment's replication. What Singh & Wesson (2009) called specific heuristics we have fitted them in perspective. Thus, the perspectives may be used to guide the inspection; and there is no need of previous training on the use of specific heuristics.

We conducted an exploratory study with the purpose of selecting the most appropriate perspectives for usability of ERP system. The study was splitted into two stages. In the first step, using ten Nielsen's heuristics we inspected four market ERP systems collecting the common usability problems. Mapping the same usability problems to the heuristics proposed by Singh & Wesson (2009), we noted that all usability problems occurred in terms of presentation, navigation, and task support.

We listed 23 issues that representing the most common usability problems in ERP systems. Table 1 shows issues about usability aspects matched to the Nielsen's heuristic and grouped by four categories Presentation (P) Feedback (F), Navigation (N), and Task Support (S).

Learnability and customization criteria weren't added because we have considered that these could be best measured when the observation is focused on the user interaction with the system for instance in usability testing rather than usability inspection. On the other hand, we added a new category denominated “Feedback”, because we observed (in the first step) that ERP systems had many violations

Table 1: Issues to identify the degree of importance assigned for usability aspects by users ERP systems.

Categories	# Issue	Nielsen’s Heuristic	Usability aspects issues	i	ii	iii	iv	v
Presentation	P-1	H1	Visual identification of user location in the system: recognize where you are and where you go.	14	12	10	1	0
	P-2	H2	The position and the arrangement of contents on the screen follow a logical order.	19	15	3	0	0
	P-3	H8	The content on the screen is irrelevant or rarely required information.	7	10	14	5	1
	P-4	H5	Right indication of input data (e.g. date, zip code, and phone).	23	11	2	1	0
	P-5	H5	Required form fields are highlighted.	24	7	5	1	0
	P-6	H6	The vocabulary used on the screen can easily be remembered.	8	14	10	5	0
Feedback	F-1	H1	Information messages of what is happening in the system.	12	13	10	2	0
	F-2	H2	The use of familiar language which can easily be interpreted.	11	16	9	1	0
	F-3	H4	Standardization of messages, shapes, symbols and system colors.	14	11	10	2	0
	F-4	H5	Alert messages to incorrect or inconsistent data formats.	26	8	3	0	0
	F-5	H9	Error messages are visible, simple and easy to understand.	29	4	3	1	0
Navigation	N-1	H9	Support to undo and redo actions (navigability).	5	17	14	0	1
	N-2	H3	Control to return to the starting point or leave an unexpected state.	8	13	13	3	0
	N-3	H5	Messages that prevent problems from occurring in the case of misguided actions.	15	7	12	2	1
	N-4	H5	The action buttons are identified clearly and define the state that will be reached after pressing them.	14	13	9	1	0
	N-5	H7	Information arrange alphabetically or by known logical.	4	14	16	3	0
	N-6	H7	Interface elements are arranged as to minimize the effort of physical actions and visual searches.	11	13	10	2	1
Task Support	S-1	H9	Support to undo and redo actions (task support).	11	13	13	0	0
	S-2	H5	The system provides a sequence of steps to complete tasks.	11	16	7	3	0
	S-3	H7	Automation of routine and redundant tasks.	12	14	7	4	0
	S-4	H7	Search filters are appropriate for the number of items and information.	9	16	10	2	0
	S-5	H9	The messages which aid in error recovery and show how to access alternative solutions.	14	14	8	1	0
	S-6	H10	Support for complex tasks providing visible and accessible instructions.	16	13	8	0	0
Legend: Presentation (P) | Feedback (F) | Navigation (N) | Task Support (S)								

regarding feedback messages to user actions.

In the second step, we conducted a survey in order to identify the degree of importance assigned by ERP systems' users for these usability aspects. In each issue, we posted a briefly example of application to aid the user interpretation. The survey was answered by 37 users of ERP systems who attributed the degree of importance for each question using the Likert scale – (i) *very important,* (ii) *significantly important,* (iii) *important,* (iv) *somewhat important or* (v) *not important.* See the results on Table 1.

We observed that of the six questions related to the heuristic of error prevention (H5), three were assigned as *very important* by at least 62% of participants. Overall, participants attributed greater importance to usability aspects that are linked to error prevention (H5); and recover from errors (H9).

Based on the analysis of the two steps outcomes - the ERP systems inspection and the survey - we reviewed the four usability categories concluding that the issues could be grouped by only two perspectives: Presentation and Task Support. The main reason for reducing the four categories to the two perspectives was to avoid overlaps or redundancies on the definition of perspectives. First, we grouped issues on Presentation categories and Feedback by Presentation perspective, because both categories may represent the characteristics related to application layout and arrangement of interface elements; defining how information is presented to users; focusing on what the user sees and understands. Moreover, the survey revealed us that most participants have considered very important the issues on category of Presentation (42.79%), as also the issues on category of Feedback (49.73%). In the same direction, we clustered issues on Navigation categories and Task Support by perspective of Task Support. Conte et al. (2009) state that the navigation usability issues can be consider satisfactory when the navigation options allow the users to track their own tasks effectively, efficiently and pleasantly. Therefore, the effective task support may be achieved from a good navigability. Moreover, most participants have deemed significantly important the issues on category of Navigation (34.68%), as also the issues on category of Task Support (38.74%).

Table 2 shows the proposal of perspective-based ERP heuristics mapping the Nielsen's heuristics into the ERP perspectives - Presentation (P) and Task Support (S), in a total of thirteen subperspectives: eight heuristics in the perspective of presentation and five heuristics in perspective of task support.

Following the Zhang et al (1999) recomendation we suggested three key-questions to guide the inspector through the perspective of Presentation - (i) “Am I seeing?”; (ii) “Am I understanding?”, and (iii) “Is the message clear for me?”, and through the perspective of task support - (iv) “Can I complete the task without obstacles?”.

Furthermore, based on the survey's questions, we created tips for each heuristic to guiding the inspector during the inspection of an ERP system. For example, for the heuristic of “Prevention of errors” in the perspective of Task Support (S5) the tip is: “Verify if the system can guide the user through the correct sequence of operations to complete a business process”.

The heuristics were mapped from the tips that were created as orientations for inspection. We analyze the impact that each orientation had on the perspectives of presentation and task support; so it some heuristics have been mapped to the perspective of presentation; others to the perspective of task support tasks; and others to both perspectives.

Table 2: Perspective-based ERP heuristics.

#	Nielsen's Usability Heuristics	ERP Perspectives	
		P	S
H1	Visibility of system status	✔	
H2	Match between system and the real world	✔	
H3	User control and freedom		✔
H4	Consistency and standards	✔	
H5	Error prevention	✔	✔
H6	Recognition rather than recall	✔	✔
H7	Flexibility and efficiency of use		✔
H8	Aesthetic and minimalist design	✔	
H9	Help users recognize, diagnose, and recover from errors	✔	✔
H10	Help and documentation		✔
Legend: [P] Presentation | [S] Task Support			

4 EMPIRICAL STUDY

We carried out an experiment to compare the efficiency and effectiveness of usability inspections between the perspective-based ERP heuristics (HPERP) and the traditional Nielsen's heuristics (HN). It is noteworthy that in this study we did not address issues of severity of usability problems

identified. Two sub-modules of an enterprise management web-based system were selected for inspection: the first one was a medium-fidelity prototype of Holiday Planning of the Human Resources module; and the second one was a functional sub-module of Retail Sales of the Sales module.

4.1 The Subjects

Two group of undergraduate and graduate Brazilian students in Computer Science at UFSCAR - Sorocaba campus enrolled in the course of Human-Computer Interaction of course - 11 undergraduates and 8 graduate students - were selected by convenience. All participants signed a consent term and filled out the participant's characterization form. The characterization form responses revealed that: only one participant (5%) had greater familiarity and experience with ERP systems, 11% of participants had some familiarity and the others either had no familiarity (42%) or had a very superficial knowledge of the subject (42%). Regarding experience with heuristic evaluation, the majority of participants (84%) did not know the technique, 11% had some knowledge on the subject and only one participant (5%) had experienced the technique. The fact that most participants were novices in inspection technique has allowed us to verify the effectiveness of learning of technique.

All participants received the same training during 4 hours which was splitted into: (i) the explanation of the concepts of heuristic inspection; and (ii) the warming up in which all participants, using the two sets of heuristics (HN and HPERP), revised a finantial module of the ERP web system called "GestãoJá" (Betalabs, 2013).

4.2 Inspection Procedures

We conducted the inspection activity a week after the training and all participants performed the tasks on the same day and time. Participants were divided into four inspection groups with 4 or 5 members: the first and the second group (G1 and G2, respectively) would work with a medium-fidelity prototype of Holiday Planning sub-module; while the third and the fourth group (G3 and G4, respectively) would inspect the functional sub-module of Retail Sales. Aiming to compare the Nielsen's heuristics to our proposal, we assigned to G1 and G3 the HN and to G2 and G4 HPERP proposal.

The groups were accommodated in two laboratories and the inspections were conducted individually. Participants were given as support material: (1) inspection instructions guided by task-based; (2) a spreadsheet to record usability problems found and the corresponding violated heuristics; and (3) the web links to medium-fidelity prototype and functional sub-module. Beyond the support material, the groups that used the HPERP (G2 and G4) received a summary table of the heuristics with the tips to guide the inspections through the perspectives of presentation and task support. The other groups G1 and G3 followed the Nielsen's heuristics guidelines. Both HN and HPERP groups should accomplish the same tasks (task-based) in the system that they would inspect, avoiding that they had different degrees of difficulties concerning on the tasks of interaction. Following the Nielsen recommendations (Nielsen, 1995), the experiment was undertaken in a controlled environment and a predetermined time-limit - two hours - avoiding the tiredness of the participants and the misunderstanding of violations, factors that could affect negatively the inspection's results. Upon concluding of the inspection activity, the participants, individually, answered a questionnaire whose objective was to catch from the participants' perceptions concerning the ease of use, the ease of understanding and usefulness of both heuristic evaluation technique (HN and HPERP).

4.3 Data Analisys

The lists of violations identified by inspectors individually were integrated into a single list by group of inspection (G1, G2, G3 and G4). The four lists were analyzed by three HCI researchers to eliminate duplicate violations (same violation pointed out by more than one inspector) and false positives (violations that were not considered real problems). After the inspection consolidation by the HCI researchers, we noted that the total violations indicated by the inspection of the medium-fidelity prototype from the use of perspective-based ERP heuristics (102) was almost twice the number of violations indicated from the inspection with Nielsen's heuristics (52). Regarding the functional sub-module, although the total of violations indicated from the use of the perspective-based ERP heuristics (115) were also higher; the difference in the total number of violations pointed out by Nielsen's heuristics (106) is very small. Considering the time to inspection activity, we found that the groups who inspected the functional sub-module was on average 22% less than the time spent by groups that inspected the medium-fidelity prototype.

Table 3 shows the efficiency and effectiveness indicators calculated by group of inspection. The efficiency for each inspection group was determined by the ratio between the average of confirmed violations and the average time spent on inspection activity. The best indicator of efficiency (approx. 23 violations per hour) was obtained by G3 who inspected the functional sub-module using HN; and the smallest indicator (approx. 7 violations per hour) was obtained by G1 which also had used the HN but the inspection was on medium-fidelity prototype.

The effectiveness per group was calculated as the ratio of the average confirmed violations and the number of known-problems. Known issues were previously identified by two usability experts who inspected the same objects. One of the experts sho used the HN identified 41 violations in the medium-fidelity prototype and 43 violations in the functional sub-module. The second expert who used the HPERP pointed out 43 violations in the medium-fidelity prototype and 50 violations in the functional sub-module. Regarding the effectiveness indicators, it is observed that the group G3 had the highest efficacy indicator detecting 61.63% of known-problems, while the group G1 had the lowest indicator of effectiveness because it was detected only 24.19% of known-problems.

Table 3: Efficiency and effectiveness indicators.

Group	Average #V.C.[1]	Average Time (h)	Average #V.C. / hour[2]	#V.C. / know problems[3]
G1	10.40	1.50	6.95	24.19%
G2	20.40	1.60	12.75	47.44%
G3	26.50	1.18	22.55	61.63%
G4	23.00	1.22	18.90	46.00%

[1] Number of confirmed violations
[2] Efficiency indicator
[3] Effectiveness indicator

In order to confirm significant differences between the efficiency and effectiveness of each group, we analyzed the results of inspection activity by statistical tests considering a confidence interval of 90% (α= 0.10) due to the small sample size (Dybå, et al., 2006). Null hypotheses and corresponding alternative hypotheses were tested:

- H_{01}: There is no difference between the efficiency of usability inspection technique in ERP systems performed with the HPERP or with the HN.
- H_{A1}: The usability inspection in ERP systems using HPERP is more efficient than usability inspection in ERP systems using the HN inspection.
- H_{02}: There is no difference between the effectiveness of usability inspection technique in ERP systems performed with the HPERP or with the HN.
- H_{A2}: The usability inspection technique in ERP systems using HPERP is more effective than usability inspection in ERP systems using the HN inspection.

Based on the hypotheses and in the research questions previously defined in the section Introduction the next subsections states the outcomes gathered from the statistical analysis.

4.3.1 Efficiency and Effectiveness

The statistical analysis of efficiency sought to answer the first research question RQ1. The efficiency for participants was individually calculated as the ratio between the number of confirmed violations and the time spent on inspection activity. When we compare the samples with the nonparametric Mann-Whitney test (Juristo & Moreno, 2010) we found a significant difference between G1 and G2 ($p = 0.037$), ie, the G2 who used the HPERP was more efficient than the G1 who used the HN considering the object of medium-fidelity prototype. However, the Mann-Whitney test revealed no significant difference between G3 and G4 ($p = 0.54$). In this case, we can consider that G3 and G4 had similar performance to inspect the functional sub-module. The results of the groups who performed the inspection of the medium-fidelity prototype of Holiday Planning sub-module (G1 and G2) confirm the alternative hypothesis H_{A1}, therefore the H_{01} null hypothesis is rejected. While the results of the groups who inspected the sub-module functional of Retail Sales (G3 and G4) support the null hypothesis H_{01}, consequently the H_{A1} alternative hypothesis is rejected, because statistical analysis revealed no significant differences between the groups.

Likewise efficacy data were statistically analyzed to answer the second research question RQ2. The effectiveness of participant was individually calculated by the ratio between the number of violations indicated and the number of known-problems (previously identified by usability experts). The Mann-Whitney test revealed a significant difference between G1 and G2 ($p = 0.044$), reaffirming that G2 - who used HPERP - was more effective than G1 - who used the HN - to inspect the medium-fidelity prototype. However, we found no significant difference between G3 and G4 ($p = 0.39$) suggesting that both groups had a similar

efficacy to inspect the functional sub-module. The results of the groups who performed the inspection of medium-fidelity prototype of Vacation Planning sub-module (G1 and G2) confirmed the alternative hypothesis H_{A2}, and then we can reject the null hypothesis H_{02}. The results of the groups that inspected the functinal sub-module of Retail Sales (G3 and G4) give support the H_{02} null hypothesis, and we can reject the H_{A2} alternative hypothesis, since statistical analysis revealed no significant differences between groups.

4.3.2 Usability and Usefulness Perceived

The questionnaire applied after inspection activity was elaborated based on model called TAM (Davis, 1989) to identify factors involved in satisfaction of individuals regarding the acceptance and use of the inspection technique, answering the third research question RQ3.

The participants answered eight questions indicating their response on a Likert scale with six tracks of agreement (*strongly agree, largely agree, partially agree, partially disagree, largely disagree e strongly disagree*). The responses *strongly agree* would receive score 6 (total agreement - 100%) and the responses *strongly disagree* would receive the score 1 (total disagreement - 0%), for purposes of calculating. Figure 1 presents a graph with the level of agreement of the inspection groups who had used of HN in relation to groups who had used the HPERP on the perception of ease of use, ease of understanding and usefulness of the technique.

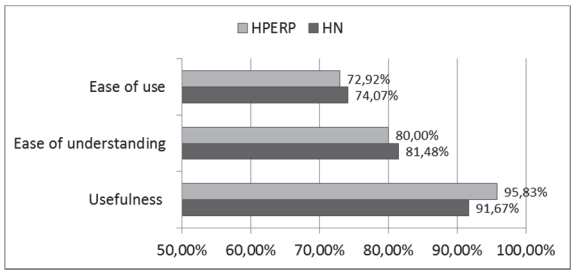

Figure 1: Degree of agreement.

There is little difference in the level of agreement between the groups because higher percentages for ease of use, ease of understanding have been assigned to groups who had made use of HN. However, the percentage of agreement about the usefulness of the technique was greater in groups who have used HPERP. Importantly most of the answers are within the range of concordance which *largely agree* (between 70% to 99%), so participants, regardless of the set of heuristics used, demonstrated a good perception about ease of use, ease of understanding and the utility of heuristic evaluation technique.

4.3.3 Validity of Results

We analyze the validity of the results of this study considering four levels of threat (Wohlin, et al, 2012.): internal validity, external validity, construct validity and conclusion validity.

Internal validity is concerned to issues that could affect the performance of the participants in the inspection activity, such as the kind of applied training and the lack of motivation of the participants (once the inspection activity was mandatory item of the course). The same type of training was applied simultaneously to the four inspection groups on the inspection technique with the two set of heuristics. Regarding motivation of the participants we observed that the degree of agreement among participants about the usefulness of the technique was a mean 93.75%.

The major threats to construct validity were related to the following issues: (i) inspection of the medium-fidelity prototype; and (ii) experience with concepts of ERP systems and the inspection technique. The groups who did the inspection in the medium-fidelity prototype reported that they had difficulties to analyze usability problems, given that the prototype did not reproduce properly all the user actions and the feedback of system. However, we know the inspection in the design initial phase is a good practice to prevent the usability problems in the software development. Regarding participants' experiences, although many participants had some knowledge about ERP systems they had no knowledge of the inspection technique. Nevetheless, this was not a significant concern, since novice inspectors are able to inspect interactive interfaces starting from basic training in heuristic evaluation.

External validity means examining the possibility of the results being generalized beyond the academic environment. The present study has characteristics with the potential to generalize results. First, the heuristics by perspectives have been proposed based on the results of an investigation of major usability problems seen in consolidated products in the ERP software market. Furthermore, the functional module and medium-fidelity prototype used in the experiment were developed by a developer of ERP software industry who is partner of our research project.

The conclusion validity refers to issues that affect the ability to draw correct conclusions, eg the

choice of appropriate statistical methods for analysis. We attend to this issue and we take care to verify if the samples were normally distributed. As the normality test result was negative for a sample, we opted to run the Mann-Whitney which is one alternative nonparametric test to the t-test in cases where data are not normally distributed.

5 CONCLUSIONS AND FUTURE WORK

The experiment results showed us that the perspective-based ERP heuristics are better for inspect medium-fidelity prototypes because the tips created for each heuristics can lead better the inspectors during the inspection of ERP systems.

In future work, new experiments in ERP industry will be conducted in order to refine the process of usability evaluation. We intend to accomplish further experiments in order to verify if there is significative variance that should be considered on the use of perspective-based heuristics in prototypes of different fidelities. We also intend to include an analysis of usability problems by severity, and check if the benefits identified in this paper can be held with experienced evaluators.

ACKNOWLEDGEMENTS

We thank the company for providing the prototypes used in the validation of the proposal.

REFERENCES

Betalabs, 2013. Gestão Já. *Betalabs Technolog Ltd., São Paulo, Brazil*.

Conte, T., Massolar, J., Mendes, E. & Travassos, G. H., 2009. Web usability inspection technique based on design perspectives. *IET software*, pp. 106-123.

Davis, F. D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340.

Dybå, T., Kampenes, V. B. & Sjøberg, D. I., 2006. A systematic review of statistical power in software engineering experiments. *Information and Software Technology ,* 48(8), pp. 745-755.

Faisal, C. M. N. et al., 2012. A Novel Usability Matrix for ERP Systems using Heuristic Approach. *International Conference on Management of e-Commerce and e-Government.*

Fernandez, A., Insfran, E. & Abrahão, S., 2011. Usability evaluation methods for the web: A systematic mapping study. *Information and Software Technology,* 53(8), pp. 789-817.

Hollingsed, T. & Novick, D. G., 2007. *Usability inspection methods after 15 years of research and practice.* Texas, ACM, pp. 249-255.

Juristo, N. & Moreno, A. M., 2010. *Basics of software engineering experimentation*. Springer Publishing Company, Incorporated.

Lambeck, C., Müller, R., Fohrholz, C. & Leyh, C., 2014. *(Re-) Evaluating User Interface Aspects in ERP Systems - An Empirical User Study.* IEEE, pp.396-405.

Law, E. L.-C. & Hvannberg, E. T., 2004. *Analysis of strategies for improving and estimating the effectiveness of heuristic evaluation.* New York, ACM Press New York, p. 241–250.

Lucas, W. T., Xu, J. & Babaian, T., 2013. *Visualizing ERP Usage Logs in Real Time.* ICEIS, pp. 83-90.

Mirel, B. & Wright, Z., 2009. Heuristic evaluations of bioinformatics tools: a development case. *Human-Computer Interaction*, pp. 329-338.

Nielsen, J., 1995. *How to Conduct a Heuristic Evaluation.* http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation.

Oja, M. K. & Lucas, W., 2010. Evaluating the Usability of ERP Systems: What Can Critical Incidents Tell Us? *Fifth Pre-ICIS workshop on ES Research.*

Scholtz, B., Calitz, A. & Cilliers, C., 2013. Usability Evaluation of a Mediumsized sized ERP System in Higher Education. *The Electronic Journal Information Systems Evaluation ,* Volume 16, pp. 148-161.

Singh, A. & Wesson, J., 2009. Evaluation criteria for assessing the usability of ERP systems. *Proceedings of the 2009 Annual Research Conference of the South African,* pp. 87-95.

Sutcliffe, A. & Gault, B., 2004. Heuristic evaluation of virtual reality applications. *Interacting with computers,* 16(4), pp. 831-849.

Uflacker, M. & Busse, D., 2007. Complexity in Enterprise Applications vs. Simplicity in User Experience. *HCI'07 Proceedings of the 12th international conference on Human-computer interaction: applications and services,* pp. 778-787 .

Wohlin, C. et al., 2012. *Experimentation in software engineering.* Springer.

Zhang, Z., Basili, V. & Shneiderman, B., 1999. Perspective-based usability inspection: An empirical validation of efficacy. *Empirical Software Engineering,* 4(1), pp. 43-69.