Development of a Generator System of Parameterized Questionnaires to Evaluate the Usability of Web Pages

William J. Yamada[1(✉)], Marcelo Morandini[1], Thiago A. Coleti[2], and Cinthyan Renata S. C. Barbosa[3]

[1] Princeton University, Princeton, NJ 08544, USA
[2] Informatics Department, State University from North of Parana State, Bandeirantes, Parana, Brazil
[3] Computer Department, State University of Londrina, Londrina, Parana, Brazil

Abstract. The effectiveness of using websites have a strong dependence on their ergonomic qualities, which affect directly the business. This is more significant when related to the e-commerce, especially considering the little known companies. To verify ergonomic features in a website, the evaluator may conduct inspections based on questionnaires that are answered by him/her or by final users to provide information about the interface.

In this paper, our goal was the development of an application that automatically generates ergonomic questionnaires to evaluate the usability of web sites. The main idea is to select specific questionnaires and specific questions that must be parameterized according to the users choice. Among these parameters, are the interaction elements present in the pages that compose the website, such as scroll bars, buttons, hyperlinks, windows, text fields and others.

The analysis of the questionnaire responses, provides data to build better usability interfaces, resulting in less time and effort to achieve their goals, and thus, delivering positive results. These questionnaires and their responses can be stored in a database for future reference and may support new usability assessments.

The tool was used by 22 evaluators. These users are aged between 19 and 30 years, and have some connection with the University of Sao Paulo. Correlations were also made between some questions to see if there was any relationship between them.

Keywords: Usability · Evaluation · Survey

1 Introduction

Usability is defined as the measure that can be used to achieve specific objectives with effectiveness, efficiency and satisfaction in a given context of use (ISO) and is an important factor in human-computer interaction [4]. With the development

of the internet and technology, the society has more access to information and as a result, more websites are created and used.

On the web, usability is fundamental. If a site is difficult to use, the users may give up using it, after all, there are several website options that can be used for the same purpose. If the information on a website is difficult to find, or does not answer the key questions of users, or if they get lost on a website, they will give up accessing it.

There is no tutorial for a website. If a user does not feel comfortable, his first line of defense is to leave, and will not access it again, negatively recommending the site to friends and acquaintances [2].

In the current web scenario, more and more, we find applications that seek to exploit to the maximum the interactivity with their users. However, this interaction does not always occur natural or intuitively, what can generate dissatisfaction, either by not finding the desired information, by the difficulty of navigation or by other usability problems. These problems can cause the user to avoid or even quitting the use of the website [2].

A common way to evaluate the usability of a web interface is through the use of usability questionnaires [19], since can be used to collect information at the initial stage of the project and also, to evaluate an existing site. One advantage of using questionnaires is due to their low cost and time saving. However, these questionnaires can be very comprehensive, containing several questions that do not apply to the page or interface in which it is being evaluated.

In this scenario, aiming to optimize these evaluations, we decided to develop a tool that will receive as input, one or more elements of interaction and generate a specific ergonomic questionnaire, which will be based on some questions related to the chosen elements.

Selecting only the most relevant questions for the required interaction elements and based on the application domain provided by the user, we could generate more specific questionnaires that optimize the usability evaluation of the web pages.

2   Related Work

Questionnaires are easy to use and allow the data collection even by professional or traditional users, and could be applied to identify subject aspects of user, interfaces and systems such as usability, acceptance and users engagement. For supporting ergonomic inspections, questionnaires are used as checklists and checkpoints, that guide the evaluator in how he/she must analyze and which are the interface's parameters that must be considered.

However, some kind of questionnaires used for ergonomic inspection have great amount of questions that can be not used for that evaluation context. It happens because the questionnaires are based on guidelines that provide prepared questions, answers and parameters to be followed in the inspection. As an example, the website ErgoList (www.labiutil.inf.ufsc.br/ergolist/) is a widely used ergonomic inspection survey, but it presents about 180 questions based

on eighteen ergonomic criteria with the followed answers options: Yes, No, Not Applied and Answer Later.

Due to the amount of questions, conducting an inspection can be an unproductive activity since evaluators must read, analyzed and maybe answer questions not applied for the inspection because the questionnaire is related to a guideline. Also, using practices proposed by several different guidelines can provide more inputs to support inspection and analysis of interface's quality. In this sense, the evaluator should complete two or more questionnaires making the inspection complex even as difficult to analyze data.

Besides ErgoList, other questionnaires can be found such as ISONORM 9241/10 and WCAG 2.0 all of them used to evaluated user interface ergonomic.

The ISONORM 9241/10 is based on norm 9241 and aims to evaluate the compliance of the software products with the recommendations contained in the part of the standard. The questionnaire is applied in two extremely opposing phases of the software project: the first is more traditional, for the evaluation of products already completed and in the commercialization phase, and the second, which is much rarer when it comes to the application of satisfaction assessment questionnaires, for the measurement between the team of designers and the future users of a new module of a software product [23].

The Web Content Accessibility Guidelines (WCAG) are part of a series of web accessibility guidelines published by the Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C), the main international standards organization for the Internet. WCAG 2.0 was published as a W3C Recommendation on 11 December 2008 [11,12]. It consists of twelve guidelines (untestable) organized under four principles (websites must be perceivable, operable, understandable, and robust). Each guideline has testable success criteria (61 in all) [13]. The W3C's Techniques for WCAG 2.0 [14] is a list of techniques that help authors meet the guidelines and success criteria. The techniques are periodically updated whereas the principles, guidelines and success criteria are stable and do not change [15].

2.1 Related Questions

As mentioned earlier, questions regarding the feedback and readability questionnaires were analyzed. All questions of each questionnaire can be found in Annex A, which were based on the ErgoList website questionnaires.

3 Materials and Methods

A tool to support the development of personalized questionnaires that could be automatically deployed for user, can improve the process of create ergonomic inspections to ensure a more targeted evaluation according to interface features. So our goal was to develop an application that automatically generates ergonomic questionnaires to evaluate the usability of web sites.

The idea is to select specific questionnaires that generate specific and predetermined questions based on parameters, that must be parameterized according to the user's choice. This parameters are the entry of the respective application domain, for example, if the application domain is social media, then the evaluator could choose as input parameters: user experience, readability, promptness and conciseness, and the corresponding questionnaires to the input parameters will be generated. In this way, it is possible to generate parameterized and customized questionnaires according to the application domain, showing just relevant questions that can be totally useful for the evaluation context.

It is expected that future website developers can use this tool to assist in the implementation of practical environments to be used and that are efficient enough to use, always aiming at the ergonomics and usability of the website.

3.1 Evaluation of the Usability of the Tidia-Ae Portal

The usability assessment can be understood as a set of methods and techniques that must be applied in order to inspect the design quality of an interface. When performing an activity of evaluating the usability of a site, problems can be verified and with this, strategies can be developed so that new projects do not present these problems observed [6].

Tidia-Ae. Ae is an e-learning environment that allows you to manage courses, projects, and collaborative and group learning activities. The environment offers a set of resources or tools to support communication, content distribution, knowledge building, and participant management.

The Ae environment can be used both as a support for face-to-face courses and for courses or projects entirely online. For each virtual space created, called a worksite or site, it is possible to:

– Provide and solve exercises and problems in a collaborative or individual way;
– Set up, make available and take courses of any level of education: primary, secondary, secondary, higher, specialization, training, plan, develop and report on group projects;
– Conduct online meetings and store minutes and documents;
– Perform activities of collaboration and interaction of the participants;
– Support the teaching-learning process in person;
– Carry out teaching and learning activities completely online;

The Ae environment is the result of the efforts of the Tidia-Ae project funded by FAPESP (Foundation for Research Support of the State of Sao Paulo) and associated with the IMS - Global Learning Consortium and the Sakai Foundation, both international institutions that discuss in a collaborative way the use of technology and its results in educational activities [8].

3.2   Development of the Tool

To build the application, was used Java 1.8.0_181 based on Oracle to create the questionnaire interface and MySQL server 14.14 Distribution 5.7.23, as the database manager system, to store the information in the database.

First the tool was developed and hosted on a local computer for some functionality tests and then, was created an online version of it, using google forms, following the same logic of the application (selecting only relevant questionnaires that generates specific and predetermined questions for the user according to the input parameters) with the benefit of sending a link, so that people can access and respond remotely wherever they are, like in a web application. The input parameters are the type of questionnaire available, according to the application domain, so for example, if the input parameter is user experience, the tool will generate the predetermined questions related to the user experience questionnaire.

To each questionnaire (total of 7), which can be: promptness, feedback, readability, user experience, conciseness, error message or location grouping (one questionnaire for each application domain), that necessarily have one to two interaction elements associated with each question. And as a basis for each question, the tool have some types of possible answers, such as: text field, 0 to 10, likert (terrible, bad, reasonable, good and great), not apply and, yes or no. https://www.overleaf.com/project/5bb51feb56227625b1ebb18b It is an initial version of the prototype, that still needs improvements, so that it can be better used by professionals and researchers in the field.

3.3   Data Collection Procedure and Validation

For tool validation, was counted on the participation of 22 volunteers from 19 to 30 years, which have some connection with the University of Sao Paulo and with this, there was no need to explain the site to be evaluated (Tidia-Ae), since it belongs to the university and all students have access to the platform.

The contact with the volunteers and the research guidelines were given in person and via e-mail. The volunteers had to answer all the questionnaires, simulating an application domain that involved all available parameters, with the purpose of qualify the various elements of interaction and usability of the site to be evalued (Tidia-Ae) and validating the tool, whose respective questions, are listed in Annex A. The questionnaires were answered independently (each volunteer on their respective computers and without a defined location), with an average duration of 15 to 20 min.

3.4   Transformation of Variables

In our research, was analyzed the questions belonging to the feedback and readability questionnaires.

For the calculation of the correlations and a better analysis, the variables were standardized and categorized according to the following descriptions:

1. Variables that presented their items arranged in likert scale (terrible, bad, not apply, reasonable, good and great), were transformed as follows in Table 1.

Table 1. Likert factors

Answer	Value
Good/Great	2
Reasonable	1
Not Apply	0
Bad	-1
Terrible	-2

2. Variables with items in 3-point scale such as 'yes', 'no' and 'not apply' were transformed as follows in Table 2.

Table 2. 3-point scale factors

Answer	Value
Yes	1
No	-1
Not Apply	0

We did not analyze text type responses because they were not easily standardized, as there is a wide possibility of possible answers (each user can give different answers).

3.5   Data Analysis

To analysis the data, were calculated correlation between the obtained results, using Spearman's Correlation [3]. The correlation result R is computed from the formula1, that measures the strength and direction of the relationship between two variables. In terms of the strength of relationship, the value of the correlation coefficient varies between $+1$ and -1, where the value $R = 1$ means a perfect positive correlation and the value $R = -1$ means a perfect negative correlation.

For a sample of 'n', the 'n' raw score 'Xa', 'Ya' - a raw score is an original datum that has not been transformed - are converted to ranks 'XRa' and 'YRa'. A ranking is a relationship between a set of items such that, for any two items, the first is either ranked 'higher than', ranked 'lower than' or ranked 'equal to' the second [9]. By reducing detailed measures to a sequence of ordinal numbers,

rankings make it possible to evaluate complex information according to certain criteria [5].

$$R = \rho_{XRa,YRa} = \frac{Covariance(XRa, YRa)}{\sigma_{XRa} * \sigma_{YRa}} \tag{1}$$

where:

- ρ denote the usual Pearson correlation coefficient [7] but applied to rank variables
- XRa denote the ranks of X values
- YRa denote the ranks of Y values
- σ denote the standard deviation

The calculated correlation is important to understand which usability factors were connected and if together, they have any influence on the usability of the website, whether positive or negative, thus providing important points to be considered by the website developer.

4   Results

From the answers obtained from the feedback and readability questionnaires that were applied to the volunteers, were made correlations between some questions, in order to identify important association.

First, from the answers obtained, a brief discussion will be presented, and then, all the results of Spearman's correlation will be presented in Table 3, for further analysis.

4.1   Feedback Questionnaire

In this questionnaire, was tested the correlations between question 1 and all other questions, except questions 6 and 12, because the answers were in text form, then not standardized. For each correlation, will be presented the respective results.

Question 1 $\times$ Question 2. The percentage of response obtained is shown in Fig. 1.

The R value in correlation between question 1 and question 2 is 0.167. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation, because the question 2 shows that the system provide a feedback to the user, due to some long processing, but the result of question 1 is that the system doesn't provide feedback for all users actions. Then the result obtained is as expected.
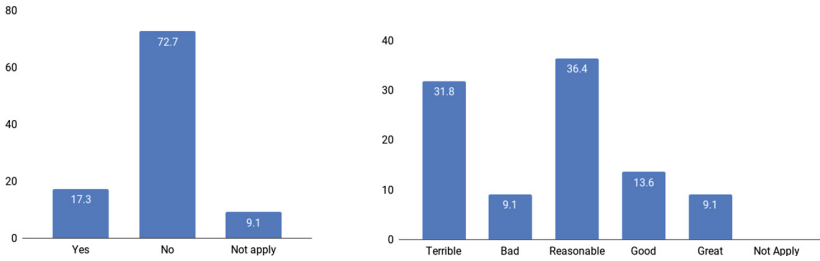
Fig. 1. Percentages - Question 1 and 2 respectively

Question 1 × Question 3. The percentage of response obtained is shown in Fig. 2.

The value of R in correlation between question 1 and question 3 is 0.30941. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation between question 1 and question 3. Although 31.8% of the answers obtained in question 3 were negative, the majority was positive (in the sense that it provides feedback), which contradicts the answer obtained in question 1.
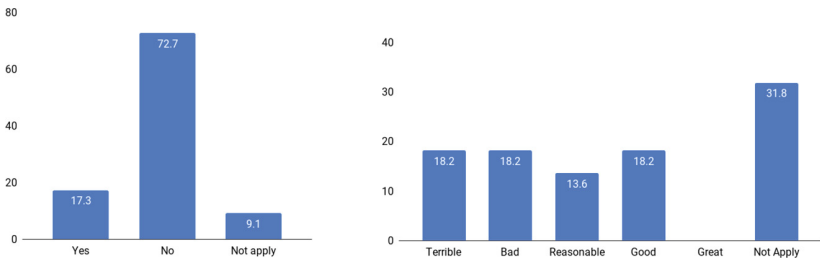

Fig. 2. Percentages - Question 1 and 3 respectively

Question 1 × Question 4. The percentage of response obtained is shown in Fig. 3.

The value of R is 0.12314. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation between question 1 and question 4. Although 31.8% of the answers obtained in question 3 were negative, the majority was positive (the system highlight the items selected - provide a feedback), which contradicts the answer obtained in question 1, that the system doesn't provide feedback for all users actions.
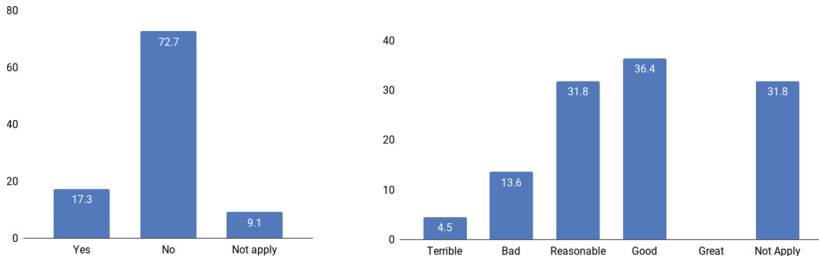
Fig. 3. Percentages - Question 1 and 4 respectively

Question 1 × Question 5. The percentage of response obtained is shown in Fig. 4.

The value of R in correlation between question 1 and question 5 is 0.36478. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation, because the question 5 shows that the system provide dynamic and contextual feedback on direct manipulation (provide a feedback), but the result of question 1 is that the system doesn't provide feedback for all users actions.
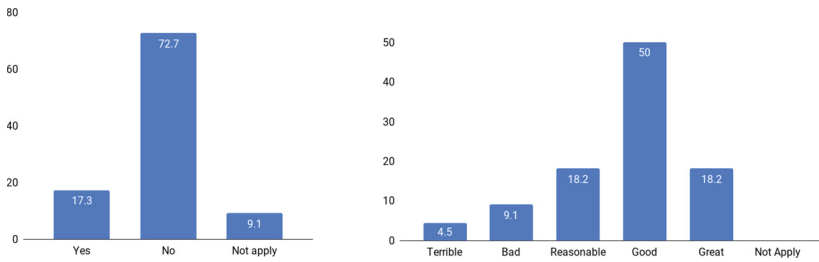

Fig. 4. Percentages - Question 1 and 5 respectively

Question 1 × Question 7. The percentage of response obtained is shown in Fig. 5.

The value of R in correlation between question 1 and question 7 is 0.25645. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation, because the question 7 shows that the system display a message stating the success or failure of time-consuming processing (provide a feedback), but the result of question 1 is that the system doesn't provide feedback for all users actions.
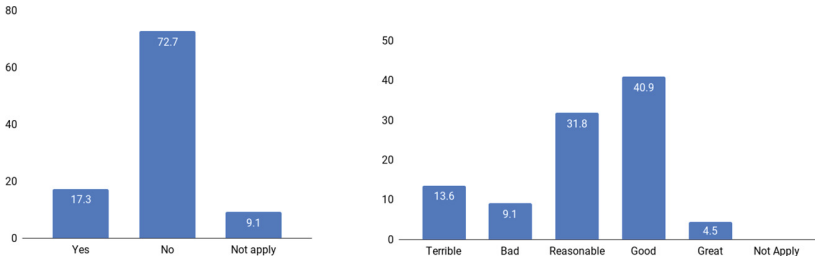
Fig. 5. Percentages - Question 1 and 7 respectively

Question 1 × Question 8. The percentage of response obtained is shown in Fig. 6.

The value of R in correlation between question 1 and question 8 is 0.27015. By normal standards, the association between the two variables would not be considered statistically significant.

This means that there is no important relationship between the system providing feedback with the system provide immediate and continuous feedback from direct manipulations. What is strange, since most of the answers in question 1 (72,7%), was that the system does not provide feedback for all the user's actions and that it corresponds with the answer obtained in question 8, in which more than half (51.8%) of the responses were that do not provide immediate and continuous feedback from direct manipulations. Therefore, some positive correlation was expected between these questions.
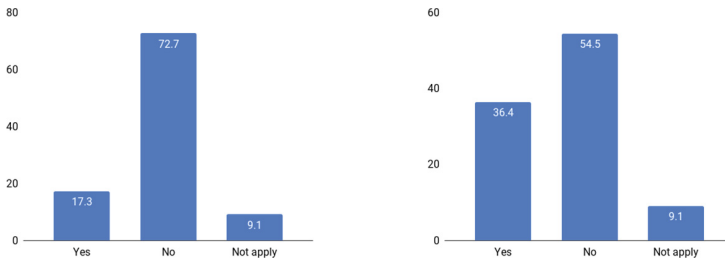

Fig. 6. Percentages - Question 1 and 8 respectively

Question 1 × Question 9. The percentage of response obtained is shown in Fig. 7.

The value of R in correlation between question 1 and question 9 is 0.22679. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation, between question 1 and question 9. Although 13.6% of the answers obtained in question 9 were negative, the majority was

positive (the system set a focus of actions for newly created or newly opened objects - provide a feedback), which contradicts the answer obtained in question 1, that the system doesn't provide feedback for all users actions.

Fig. 7. Percentages - Question 1 and 9 respectively

Question 1 × Question 10. The percentage of response obtained is shown in Fig. 8.

The value of R in correlation between question 1 and question 10 is -0.02164. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation, between question 1 and question 10. Although 13.6% of the answers obtained in question 10 were negative, the majority was positive (the system provides feedback on object attribute changes - provide a feedback), which contradicts the answer obtained in question 1, that the system doesn't provide feedback for all users actions.
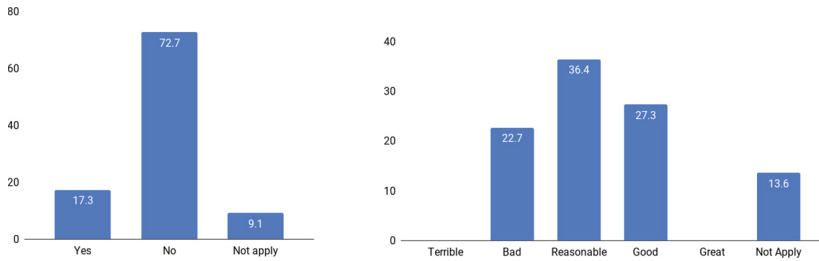

Fig. 8. Percentages - Question 1 and 10 respectively

Question 1 × Question 11. The percentage of response obtained is shown in Fig. 9.

The value of R in correlation between question 1 and question 11 is 0.3942. By normal standards, the association between the two variables would not be considered statistically significant.

Was not found a correlation, between question 1 and question 11. Although 13.6% of the answers obtained in question 11 were negative, the majority was positive (the system notify the user to any change in the current situation of control objects - provide a feedback), which contradicts the answer obtained in question 1, that the system doesn't provide feedback for all users actions.
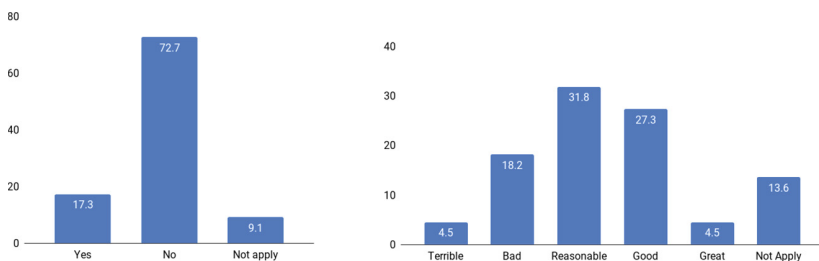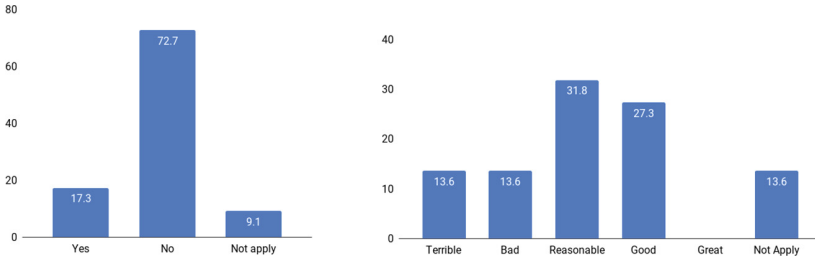

Fig. 9. Percentages - Question 1 and 11 respectively

4.2   Readability Questionnaire

In this questionnaire, was tested the combination of correlation between question 6, 7, 8 and 9. However, a brief analysis will be performed only between the correlations of questions 8 and 9, since no significant correlation was found between the others questions.

The value of R in correlation between question 8 and question 9 is 0.61237. By normal standards, the association between the two variables would be considered statistically significant. This means that there is an important correlation
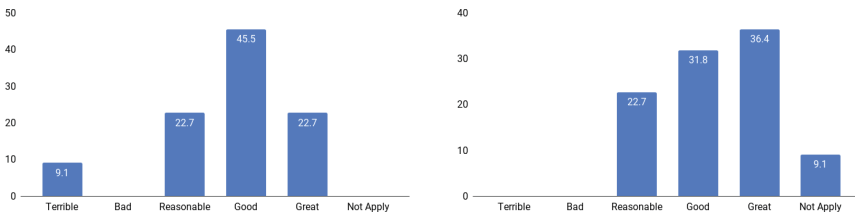

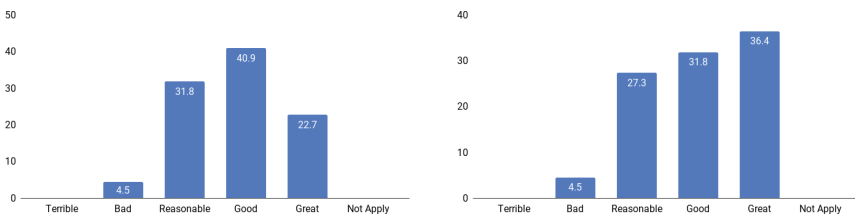**Fig. 10.** Percentages - Question 6 and 7 respectively

Fig. 11. Percentages - Question 8 and 9 respectively

between the use of bold and the use of underlining in the texts. It is therefore suggested that you use both bold and underline in moderation to have a better quality and usability of the website.

4.3   Discussion

All the results of the Spearman's correlations are presented in the Table 3. In the first line, the results of the correlations between the feedback questionnaire (question 1 and all others, except for questions 6 and 12, as explained above). The other values correspond to the correlations of the readability questionnaire (questions 6, 7, 8 and 9).

Regarding the Feedback questionnaire, no correlation was found between the compared questions, which is a bit strange, as we can see in question 1, more than 70% of the responses were "No", "the system doesn't provide feedback for all user actions", and one of the questions proves this correlation, the question 8, where 54.5% of the responses were that the "system doesn't provide immediate and continuous feedback from direct manipulations". For the other's comparison, it was to be expected that no correlation would be found, since the answer obtained in question 1 contradicted the answers obtained in the other questions, except question 8, as stated above.

So for the feedback correlation, it was expected to find only a correlation between question 1 and question 8, since question 8 proves the answer to question 1 - the system doesn't provide feedback for all user actions - but it was not what happened. This can be a problem related to a small amount of data (answers), since our database contains only 22 answers for each question.

Also, we can conclude that this inconsistency is maybe due to the question 1 have 3 point scale response (yes, no and not apply) and most of the other questions have likert responses, what is a 5 scale point response (terrible, bad, reasonable, good, great and not apply), so the values of the categorical variables compared were often different, that is, not standardized between them. Besides that, how the value of the factor "not apply" used in the Spearman's Correlation was "0", maybe, for the correlation, the "not apply" was may be considered closer to "yes" where the factor is "1", since 0 and 1 is a positive number and the "no" factor is "−1", that is, a negative number. So perhaps if we change the type of answer from question 1 to likert, that inconsistency does not recur, or associating negative responses with negative values, or even, comparing only questions that have the same range of possible answers.

On the other hand, analyzing the Readability questionnaire, it was found a correlation between the use of bold and the use of underlining in the texts, then it is suggested that who is developing a website, to use bold and underline in text with moderation to have a better readability of it and consequently a more ergonomic website. This correlation can be explained by the response graphs, where the question 8 and question 9 has a very similar answer, remembering that the answers "good" and "great" use the same categorical value "2" in the Spearman's Correlation. But we could not find a correlation between the others

Table 3. Spearman's correlation on readability questionnaire

	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8	Question 9	Question 10	Question 11
Quest. 1		0.167	0.309	0.123	0.365		0.256	0.27	0.227	−0.022	0.394
Quest. 2											
Quest. 3											
Quest. 4											
Quest. 5											
Quest. 6							0.094	0.159	−0.161		
Quest. 7								0.159	0.216		
Quest. 8									0.612		
Quest. 9											
Quest. 10											
Quest. 11											

questions, which is a bit strange, since all the questions had very similar answers, most of them belonging to "good" and "great".

This inconsistency can be explained by the fact that only questions 8 and 9 had the answers belonging to the same categorical variables "bad", "reasonable", "good", and "great", while questions 6 and 7 had answers belonging to the categorical variables "horrible" and "not apply" respectively.

So, this paper shows us that there are still many things to improve, such as perhaps standardizing the types of responses for a better comparison, as well as make comparisons with the other questionnaires, among them: user experience, concision, error message and location grouping and, of course, expand the number of responses obtained to reduce some user inconsistency, and the main one, make the tool available in a web environment, so that responses are stored directly in the application database, instead of hosted on google forms.

5   Conclusions

The present study analyzed the correlation between the responses of the questionnaires, obtained from the evaluation of the website Tidia-Ae, using the tool developed (still in its initial version), which generates specific questionnaires for the individual application domain. The questions analyzed involved aspects related to the usability features and issues of web sites, mainly referring to feedback and readability.

From the data of the 22 responses, which were correlated from questions 8 and 9 referring to the readability questionnaire, it was verified that the result of Spearman's correlation was equal to 0.61237, that is, a significantly positive correlation. Then, from this correlation, the following inference has been established: the use of bold, along with the use of underlining, brings benefits to the legibility of the website, as long as they are used in moderation. We consider this result is very important for designers, developers and usability evaluation team as this can prove that there is a correlation between these two usability features.

It is hoped that this research can be used as a basis for future researches, and when the tool is completely developed, we plan to allow it to be downloaded free to professionals and researchers, to be used as a tool to help achieve better ergonomics and usability environments.

References

1. Figueredo, A.J., Wolf, P.S.A.: Assortative pairing and life history strategy - a cross-cultural study. J. Hum. Nat. **20**, 317–330 (2009). https://doi.org/10.1007/s12110-009-9068-2
2. Nielsen, J., Loranger, H.: Usabilidade na Web - Projetando Websites com Qualidade (2007)
3. Myers, L., Sirois, M.J.: Spearman correlation coefficients, differences between. In: Encyclopedia of Statistical Sciences (2006). https://doi.org/10.1002/0471667196.ess5050.pub2

4. Liu, F.: Usability evaluation on websites. School Art Des. (2015). https://doi.org/10.1002/0471667196.ess5050.pub2. Wuhan University of Technology
5. Malara, Z., Misko, R., et al.: Wroclaw University of Technology graduates career paths. Econ. Sociol. Biol. (2016). University of Grodno
6. Cybis, W., Betiol, A.W., et al.: Ergonomia e Usabilidade, 3 edn. Novatex Publisher (2015)
7. SPSS Tutorials: Pearson Correlation. https://libguides.library.kent.edu/SPSS/PearsonCorr. Accessed 07 Dec 2018
8. Tidia-Ae, Sobre o Ae. https://ae4.tidia-ae.usp.br/portal. Accessed 9 Nov 2018
9. Merriam Webster - Ranking. https://www.merriam-webster.com/dictionary/ranking. Accessed 9 Dec 2018
10. Web Content Acessibility Guidelines (WCAG) 2.1. http://www.w3.org/TR/WCAG21/. Accessed 15 Nov 2018
11. Web Content Acessibility Guidelines (WCAG) 2.0 - W3C Recomendations, 11 December 2008. http://www.w3.org/TR/2008/REC-WCAG20-20081211/. Accessed 15 Nov 2018
12. W3C Web Standard Defines Acessibility for Next Generation Web. http://www.w3.org/2008/12/wcag20-pressrelease.html. Accessed 15 Nov 2018
13. Web Content Acessibility Guidelines (WCAG) 2.0. http://www.w3.org/TR/WCAG20/. Accessed 15 Nov 2018
14. Techniques for (WCAG) 2.0. http://www.w3.org/TR/WCAG20-TECHS/. Accessed 15 Nov 2018
15. Understanding Techniques for WCAG Sucess Criteria. http://www.w3.org/TR/UNDERSTANDING-WCAG20/understanding-techniques.html. Accessed 15 Nov 2018
16. Stair, R., Reynolds, W.: Princípios de Sistemas de Informação. Centage Learning Publisher, São Paulo (2015)
17. Pressman R.: Engenharia de Software, Uma Abordagem Profissional, 7th edn. AMGII Publisher, Porto Alegre (2011)
18. Colouris, G., Dollimore, J., et al.: Sistemas Distribuídos. Conceitos e Projeto, 4th edn. Bookman Publisher, Porto Alegre (2007)
19. Barros, V.T.: Avaliação da interface de um aplicativo computacional através de teste de usabilidade, questionário ergonámico e análise gráfica do design. In: Dissertação (Pós-Graduação Engenharia de Produção) - Faculdade da Universidade Federal de Santa Catarina (2013)
20. Vanderdonckt, J., Bodart, F.: Encapsulating knowledge for intelligence automatic interaction objects selection. In: Proceedings of the ACM Conference on Human Factors in Computing Systems, INTERCHI 1993. ACM Press, New York (1993)
21. Smith, L.S., Mosier, J.N.: Guidelines for designing user interface software. In: MITRE Corporation and Funded by the U.S Air Force's Eletronic Systems Division (1986)
22. ISO. ISO/DIS9241-16.: Ergonomic requirements for office work with visual display terminals (VDT's) - direct manipulation dialogues. Genebra
23. Prumper, J.: Software-evaluation based upon ISO 9241 part 10. In: Vienna Conference, VCHCI 1993 (1993)