

Yes, HCI Can Conciliate XBRL Financial Reporting Software and Their Users

Antônio Ricardo D'Araujo Amâncio Oliveira^{a*}, Artur Henrique Kronbauer^b

^{a,b}*Programa de Pós Graduação em Sistemas e Computação - UNIFACS, St. Dr. José Peroba, 251 -6th floor-Stiep, Salvador - BA, 41770-235, Brasil*

^a*Email: antonoricardo1993@gmail.com*

^b*Email: arturhk@gmail.com*

Abstract

This paper presents the results of a previously proposed research to mitigate the literature gap found between Human-Computer Interaction and Government to Government e-governance regarding the XBRL financial reporting area. This research conducted two usability and User eXperience evaluations with two different versions of an XBRL financial reporting software prototype. Initially the application provided XBRL knowledge abstraction and underwent HCI redesign to improve task efficiency. The results showed HCI design is a valid way to mitigate the XBRL knowledge required to elaborate XBRL financial reports problem, to improve XBRL financial reporting task efficiency, thus improving B2G and G2G e-Governance practices.

Keywords: XBRL financial reporting; HCI; HCI design; e-Government; G2G.

1. Introduction

The Human-Computer Interface (HCI) design is about analyzing the current situation (problem), synthesizing an intervention, and evaluating how it affected the situation in an iterative matter. Such a process allows designers to produce solutions better oriented to their ends (their users' actual needs) [1]. HCI design processes aim to serve the users and the stakeholders. That is why a number of them are user-centered. They also highlight how it is important to allow users to take part in the decision making processes of a solution's development. The earlier the users get involved in a project, the better the final solution's perceived quality and value [1]. "To use an interactive system consists of interacting with the system's interface to meet specific ends within a given context.

* Corresponding author.

In that scenario, Human-Computer Interaction evaluation studies allow accessing whether a system's interaction and interface are adequate or not" [2]. Such studies typically encompass usability and User eXperience (UX). According to [2], "usability means how easily systems can get used regarding learning ability, operability, aesthetics, and other aspects". This author also states that during a UX, the users interact with a product or system in a way that their experience interest is measurable or observable. UX measurement has been a helpful tool for improving software under development regardless of their lifecycle project [3]. More than 50 countries use the eXtensible Business Reporting Language (XBRL) to represent the financial statements regarding their business performance and compliance [4]. Some of their goals encompass providing better financial information comparability, easiness of analysis, and accessibility to issuers, investors, competent authorities, and people [5]. However, due to the XBRL issues as knowledge and uncertain software support, some XBRL users still struggle to elaborate on the financial reports government forces them to send to oversight custody [5]. Those problems also degrade the efficiency of Government to Government and Business to Government e-Governance [6,7,8]. A real example of this problematic situation also happens in Brazil. In this country, 5,570 municipalities, 26 states, the Federal District, and the Federal Union have to submit accounting, financial, and tax statistics information as XBRL financial reports to the Sistema de Informações Contábeis e Fiscais do Setor Público Brasileiro (Siconfi) of the Secretaria do Tesouro Nacional (STN) [9]. It is an example of a scenario in which HCI design practices typically show the potential to bring technology users a solution that actually meets their needs. As contextualized in [10] previous study, "this research originality and social relevance rely on mitigating the gaps and demands found in the literature and the XBRL financial reporting area". This paper presents the results and findings obtained after conducting the case study to verify the following hypotheses validity regarding the methodology proposed in the previous study [10]:

- H1: "Providing financial reporting professionals a software whose design considered Human-Computer Interaction (HCI) matters is enough to increase the task efficiency" [10].
- H2: "The adopted procedures to verify H1 are a valid methodology for similar studies to improve Business to Government (B2G) and Government to Government (G2G) Electronic-Government (e-Government) practices" [10].

The remaining of this paper structures as follows. The Materials and Methods section presents the methodology and the technologies that supported the research conduction. The Results section brings research conduction procedures' results that provided the hypothesis validation analysis. The Conclusion section brings the conclusions and contributions related to the proposed project.

2. Materials and Methods

As discussed in [10], "this search aimed to verify the efficiency improvement obtained in XBRL financial reporting tasks when the users adopted a tool developed under HCI guidelines to mitigate the XBRL knowledge problem". So, it was necessary developed an XBRL financial reporting software prototype (OFR) to support the evaluation sections for the sake of:

- Attaining comparability between the data gathered from users that typically adopt different tools to

perform such tasks.

- Assuring function implementation correctness.
- Preventing software compatibility problems.
- Easing identifying concise relations between evaluation results and source code.
- Easing the redesign processes.

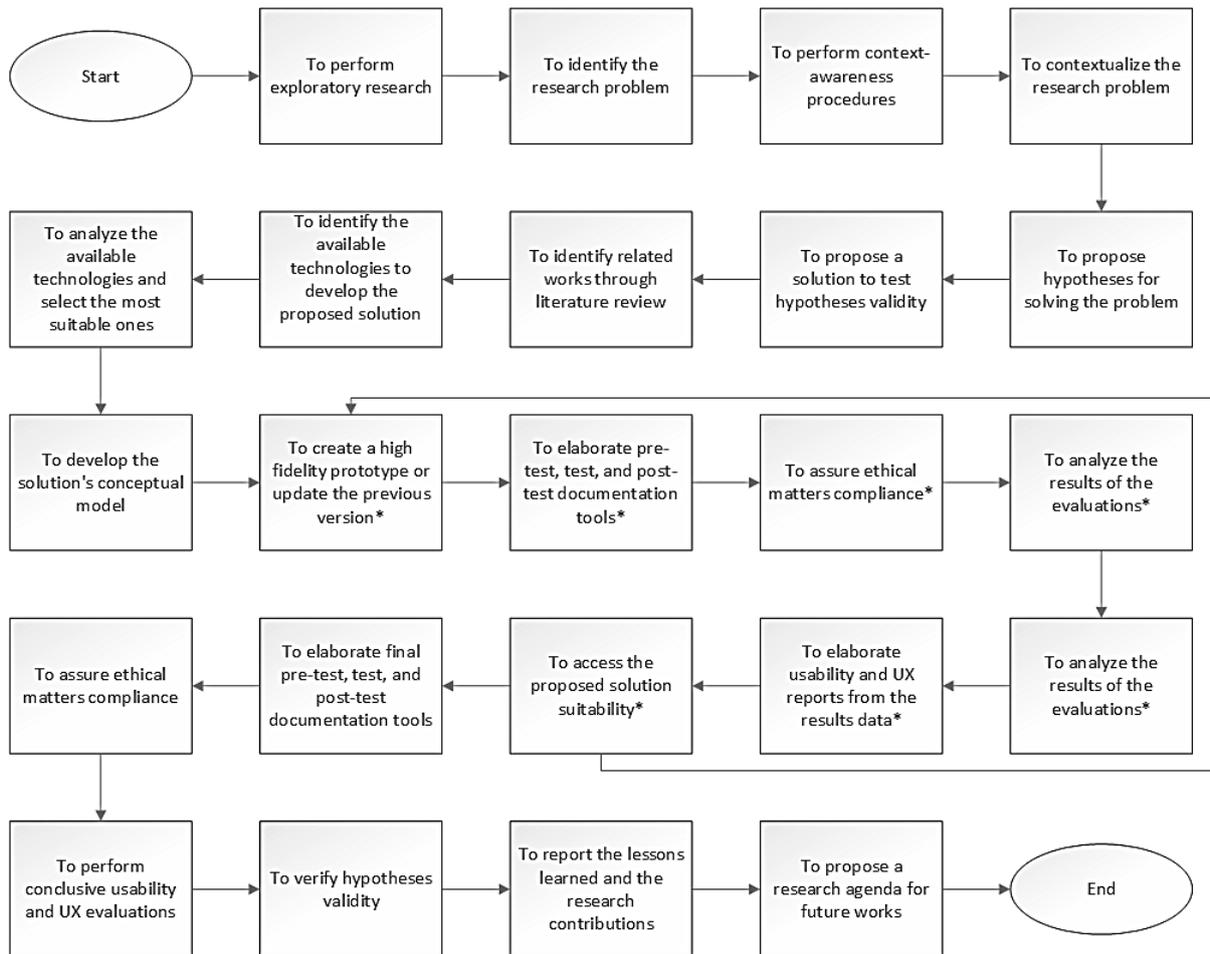


Figure 1: Pre-set methodological procedures for conducting this research (*repeated for each research cycle).

As a single developer played the role of all the stakeholders that were not users or volunteers, geographically distributed, all documentation whose purpose was to communicate concepts, and information among stakeholders from different knowledge areas became unnecessary. The OFR is a monolithic application in which all architecture components' code consist of a single java program source code. Figure 2 shows the OFR's use case diagram.

2.1. The Missing Link Between HCI, G2G e-Governance, and XBRL

The literature reviews conducted in [10] to identify: related works about HCI practices in the XBRL financial reporting domain [11, 12, 13], which of the e-Government digital interactions do the HCI solutions focus on, and what are the current HCI practices regarding the e-Government digital interactions [14, 15, 16, 17, 18, 19,

20, 21, 22, 23, 24, 25] revealed that this is the first study regarding HCI in the XBRL financial reports domain aiming to deploy HCI Design to improve G2G e-Governance, in a half-decade. Among the papers retrieved, there were also no methodologies, referrals, guidelines, or directions to support this research conduction or the technology selection process. The OFR notification features adapted the strategy presented in [13]. All the selected papers focused on Citizen to Government (C2G) e-Government digital interaction, except from the study [17] performed. According to the review conduction results, the current HCI practices regarding the e-Government digital interactions consists of: To deploy quantitative/qualitative usability and UX evaluation methods to assess technical features of e-Government web portals and mobile applications (regardless of redesign purposes) concerning international and national standards of usability, accessibility, and functionality; To provide insights about how citizens interact with government mobile or web services and social media; To provide guidelines, a research agenda, or evaluation models for assuring usability, accessibility, and functionality of e-Government's mobile or web services, and citizen participation within the public sector.

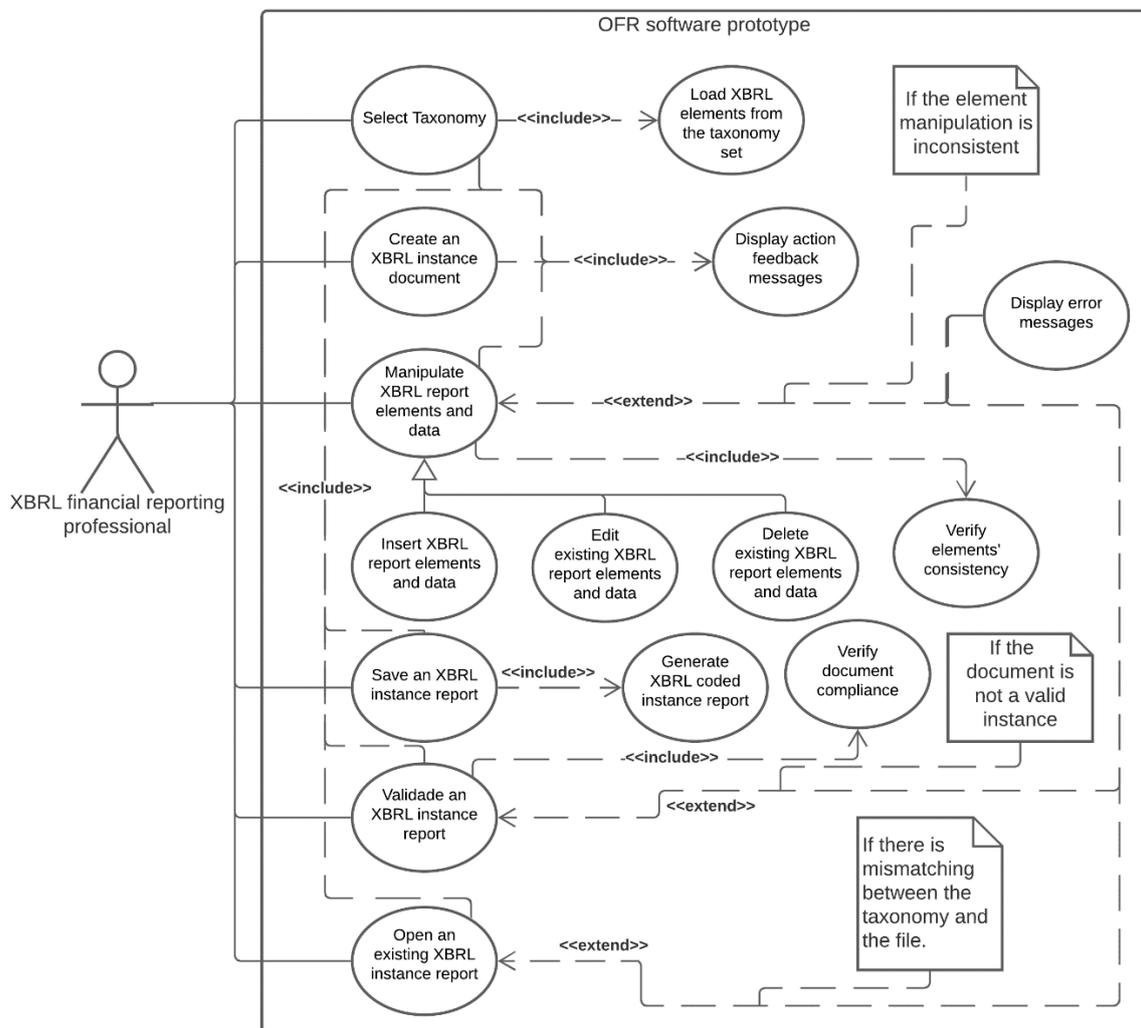


Figure 2: OFR’s use case diagram.

2.2. OFR’s Prototype in the First Research Cycle

The software prototype built for the first research cycle focused on implementing features and components that minimally allowed users to create the Siconfis's RREO Consórcios 2020 [26] report without the need to validate it according to Siconfis's taxonomy Presentation link for that report. The OFR's prototype regarded the implementation of four architectural components: `Communication_Manager`, `XBRL_File_Manager`, `XBRL_Element_Manager`, and `Error_Manager`. The `Communication_Manager` consists of the `user_interface` component and a task manager. It is a taxonomy-based self-adaptative component that gathers user's requests, forwards them to other components, and returns their results to the user. It is also responsible for sending the other architectural components all information they need. Whenever the user takes an incoherent action, the `Communication_Manager` receives the error treatment result from the corresponding component and requests the `Error_Manager` to retrieve the procedure to warn the user from the `error_message_list` component. The `error_message_list` component should contain the user interaction warn procedures generated according to information retrieved from the taxonomy set. The `XBRL_File_Manager` contains a component to create the XBRL instance document file (`instance_creation`) and another to load such files (`instance_load`), allowing the user to edit the report content, for example. It generates the instance files according to the taxonomy's hypercube structure, under the `Communication_Manager` request. This component is also responsible for keeping the taxonomy's hypercube structure in the instance documents. It is also responsible for storing the report under edition within the program, so the `XBRL_Element_Manager` can perform operations on the instance document. The `XBRL_Element_Manager` enables the operations of inclusion (`element_creator`), edition (`element_editor`), and exclusion (`element_exclusion`) of taxonomy elements from the XBRL instance document. As a provisory measure to support the prototype's testing, instead of retrieving the information from the taxonomy with the architectural component, it was necessary to go through the taxonomy and linkbase files organizing the information needed into text files. Whenever the software prototype needed information from the taxonomy, it retrieved the information from a specific text file. In that way, to enable the prototype to generate other reports, one had to fill up the text files with the content for the new Siconfi report. The user interface components for managing an instance document's elements may vary depending on the taxonomy files loaded. For example, the amount of information the user needs to input a context depends on how many hypercube axes it couples, so the software window must display all fields to allow the user to insert it. The components for opening and automatically generating taxonomy valid instance documents, generating error messages, and generating communication messages also vary according to each taxonomy. In that scenario, automatic programming mechanisms are a viable way to handle this peculiarity. However, such feature do not compose early OFR's prototype versions. Even though the validation process prevents inconsistencies in the instance file, this prototype implemented a reduced set of rules for two main reasons. Firstly, it did not prevent the instance creation task. Secondly, Linkbase formulas contained mathematical validations rules for a complete instance document. If the prototype implemented the full validation set, the users could never test this functionality because the scheduled section duration was smaller than the required to compose a complete report. A further version of this software shall fully implement this functionality. Meanwhile, a user that composed a complete financial report with this prototype version had to validate it against the taxonomy's Linkbase formula through external tools. To the best of our knowledge, all instance files generated with this prototype successfully validated against Siconfi's taxonomy Linkbase presentation through the Interstage XWand Toolkit Evaluation Copy (the evaluation purpose version of the same XBRL financial reporting tool Siconfi deploys to generate its

taxonomy and test instance files). In the first research cycle, the OFR undergone HCI formative evaluations, whose results worked as guidelines for redesigning it under HCI perspective. The HCI Design also has the premise of preventing the idiosyncrasies influence over a project. That is why OFR's redesign only considered changes that the UX and usability evaluation techniques allowed identifying and highlighted by at least three of the five testers involved. The redesign process also aimed to fix all faults and bugs identified through the evaluation sections. In the second cycle, the OFR redesigned version underwent the conclusive HCI evaluation tests to provide data for hypothesis validation and future works guidelines. The evaluation results from the second cycle allowed verifying:

- If the software prototype, built under HCI design matters, was enough to increase the reporting professionals' task efficiency;
- If the set of adopted procedures are a potential starting point for similar studies that aim to improve B2G and G2G e-Government practices;
- Whether the academic contribution to the HCI area was an unsuccessful case or not.

Due to information accessibility convenience, this research adopted the Siconfi's Relatório Resumido da Execução Orçamentária – RREO Consórcios 2020 [27] as the financial report representative of Brazilian G2G e-governance for both research cycles tests. This research deployed actual software development methods to develop the prototype for supporting hypothesis validation. Even though this stage of the current research did not aim to provide a fully commercially competitive tool, going on with the development methodology iterative cycles might allow getting to it.

2.3. HCI Evaluation

For the usability tests, this research adopted the performance measurement, the System Usability Scale (SUS) [28] questionnaire, and user feedback methods. The communicability evaluation performed under the procedures presented in [1]. In both cycles, it was necessary to conduct evaluation sections with one user at a time. All the procedures adopted underwent a pre-test with a volunteer XBRL specialist. This research did not consider the interruption time due to volunteers' requests in the overall results. Due to the wide variety of professionals (accountants, system analysts, consultants, IT professionals, business owners, and others) who plays the role of financial reporter, it was not possible to specify a strict set of occupations for the user's profile composition. As there are financial reporters who try to avoid XBRL, and the research was going to provide the users a new software prototype that abstracts the XBRL knowledge need, it did not make sense to require them to have already had in touch with platforms for composing XBRL reports once they were already going to learn how to use the OFR, and they should need only financial reporting knowledge to do so. This study conducted evaluation sections with five users, no backup volunteers, a pilot and a dry-run participants (both of them were the only reused volunteers for both research cycles). It was necessary to offer training to ensure the volunteers could achieve a minimum level of expertise, but to avoid providing information about relevant aspects for the main usability test. Then the participants answered the pre-test questionnaire that incorporated the PrEmo UX capture method. The volunteer recruiting process for both HCI evaluations offered a prize (a tablet HOW HT-705 XS) to a randomly chosen participant. The call for volunteers strictly followed the procedures presented in

[10]. To deal with the eventual need for recruiting volunteers from abroad, all the forms, supportive materials, user guides, and communication messages were available in two languages: English and Portuguese. This research kept the volunteers data for one day after each evaluation section, then everything that could allow identifying the users was deleted. Within that time, their data were not shared or handled under any circumstance. As discussed in [10], this research adopted the Emocards [29,30,31] UX capture methods due to the demand for performing online studies, with functional prototypes, with one user at a time, while gathering qualitative and quantitative UX data gathered before the user's interaction and after the whole task conclusion. This search had to deal with geographically spread users within different global time zones, so the evaluation sections had to perform individually and via the internet, as in online studies. The users had to perform financial reporting related tasks with OFR high-fidelity functional prototypes (first version and redesigned version) while providing pre-task and post-task completion interacting experiences feedback. In that way, the pre-test questionnaire contained a PrEmo and the post-test questionnaire contained a SAM and an Emocards measurement instrument through which the users could express how they felt after the test. The procedures described in this subsection were adopted for both evaluation sections. The translations adopted for the Portuguese UX capture methods explanations were extracted from [32].

3. Results

This section presents the results of all tests performed with both OFR's high fidelity prototypes. All demographic data collected allowed selecting volunteers with the desired user profile and to better understand the data each volunteer provided. It also had no other use than that. All the expected values for the evaluation items were established through the pilot test. The error indications also accounted for the differences between the user's XBRL generated files and the one proposed in the activity sheet.

3.1. Usability Formative Evaluation Results

The five volunteers that participated in the formative evaluation sections fit the following profile: Professionals with a bachelor's or higher college degree, from 31 to 60 years old, that do not struggle to use computers, that perform financial-reporting related tasks for two or more years, regardless of their XBRL knowledge degree. All of the volunteers demanded a training section before starting the experience with the OFR software, and fulfilled the task goals to create an XBRL instance with the OFR. The usability indicators related to this task showed users needed to perform one attempt on average to achieve it. They also committed no errors while performing that task. The OFR did not fall in faulty conditions during the subjects' trials. The dry-run test results allowed estimating values for those indicators: one attempt, no user errors, and no system errors, respectively. It is also important to highlight that the average productive time was 1.11 minutes, the success rate was 100 %, and the error rate of 0 %. On average, users had to invest the following amounts of time to accomplish the XBRL instance creation, to perform the first attempt, and backtracking from an error were 2.01, 2.01, and 0 minutes, respectively. The average quantity of OFR's functions users had to deploy to accomplish the first task was two. They were supposed to use two functions. Only one volunteer requested the evaluator's assistance one time. All users did not consult the software supportive documentation to complete the referred activity. They also did not use any extra time getting how to use the OFR's functions. Every volunteer successfully developed the second

task, all of them could figure it out on the first try, and learned all functions necessitated to accomplish the task. All of the individuals terminated the task to insert data with the OFR. The results assessed for the aforementioned task revealed that users necessitated performing one try on average to conclude it. They underwent a total of 9 failures during that task execution. The OFR had no errors during their attempts. From dry-run test results, the predicted rates for those indicators were one attempt, two errors per user, and no system errors, respectively. The measured average productive time was 7.95 minutes, the success rate was 100 %, and the error rate was 13.85 %. The average time enlistees used to fulfil the data insertion, to perform the first attempt, and to backtrack from an error, were 12.33, 12.33, and 0.48 minutes, sequentially. On average, users needed to use five functions to finish the second task in which they should have used five ones. On medium, participants bid for the evaluator's help one time and only one volunteer went for the OFR's supportive documentation 1 time to accomplish that activity. He/she deployed 2 minutes to figure out how to use the OFR's functions. 60 % of the enlistees accomplished the task to save and validate an XBRL instance with the OFR. However, 40 % of the individuals did not end that task properly. The statistics regarding the aforementioned task revealed that, on average, volunteers performed 1.4 attempts to fulfil the task's goals. They also had a total of one error during that task execution. The OFR did not put the users in trouble with errors during their attempts. Regarding the dry-run test results, the predicted values for those indicators were one attempt, no user errors, and no system errors, respectively. It was also necessary to measure the average productive time (1.68 minutes), the success rate (80 %), and the error rate (1.82 %). The medium time participants had to waste to save and validate an XBRL instance, to perform the first attempt, and to backtrack from an error were 5.3, 3.4, and 0.3 minutes, respectively. The average number of functions the enlistees had to deploy to accomplish the third task was 2 while they were supposed to use 2 functions. On average, volunteers inquired the evaluator's help 0.8 times. Only one user had to consult the OFR's supportive documentation one time in the referred activity. He/she also used 2.5 minutes to understand how to use the OFR's functions. The overall average activity accuracy completion was 91.11 %, and it regarded to how close the reports the volunteers created were from valid Siconfi's reports. The average SUS score is 68, and the OFR score was 70. This matches a C grade, which means OFR does not contain catastrophic usability problems. As it is a good method to distinguish between unusable and usable systems, it is possible to classify this version of the OFR software as usable. The users also provided feedback information not covered in the questionnaires and tools adopted. Even though the user feedback method provides valuable design directions when the testers give concise opinions, there were no common points in the observations and information users provided, so they represented idiosyncrasies that did not influence re-design steps. In the XBRL instance creation task, users suspended their semiosis because they did not have a suitable way to denote their communication four times. A volunteer broke off his/her semiosis one time because he/she was not able to find the means to do the succeeding communication input. And the users attempted to learn the OFR's communication process via testing many assumptions regarding its meaning two times. In the data input task, users interrupted their semiosis seven times because they were unable to identify a suitable way to express their communication. And they suspended their two times semiosis because they did not get a way to do the following communication input. The volunteers endeavored to comprehend the communication process with the OFR by implicit metacommunication three times, and they did notice the communication was not flawless because they did something wrong in the interaction about ten times. One participant could perceive one communication attempt was faulty because he/she interacted in an incorrect

context. The individuals discontinued their semiosis a total of two times because they did not perceive the OFR's communication and put effort to grasp the communication process with the OFR via experimenting with several hypotheses about the OFR's communication meaning four times. The users gave up a semiosis two times before they could attain the wanted effects to perform a new one with the equivalent meaning because they were not able to grasp the HCI proposed solution. The subjects performed an ineffective semiosis four times, but they did not commence a new one because they did not notice something was missing to match the coveted results. They also tried to figure out the communication process with the OFR via explicit metacommunication two times. For the task of saving and validating an XBRL instance, the volunteers discontinued their semiosis three times because they had no proper way to express their communication and stopped their semiosis four times because they did not figure how to perform the following communication input. One of the enlistees broke his/her semiosis one time because he/she did not grasp the OFR's communication. And the subjects put effort into learning the communication process with the OFR by attempting several theories about the software's communication meaning three times. The subjects were inclined to understand the OFR's communication process by explicit metacommunication four times. One user terminated an incomplete semiosis but did not try it again because he/she did not have a mean, the potential, or a wish to keep trying. The users desisted a semiosis three times before they could attain the coveted effects to attempt to do the same thing in another way because they did not recognize the HCI proposed solution. One of the users completed an unfruitful semiosis, but he/she did not go for another one, because he/she did not regard they did not reach the desired results.

3.2. UX Formative Evaluation Results

In the pre-test questionnaire, the users reported how they expected to feel before the experience with the OFR software through the PrEmo UX capture method (Figure 3). In the results, the OFR should provide users with elicitations of hope, pride, admiration, satisfaction, desire, joy, and fascination. Boredom, dissatisfaction, contempt, sadness, shame, fear, and disgust measured intensities should be as low as possible in the users' responses.

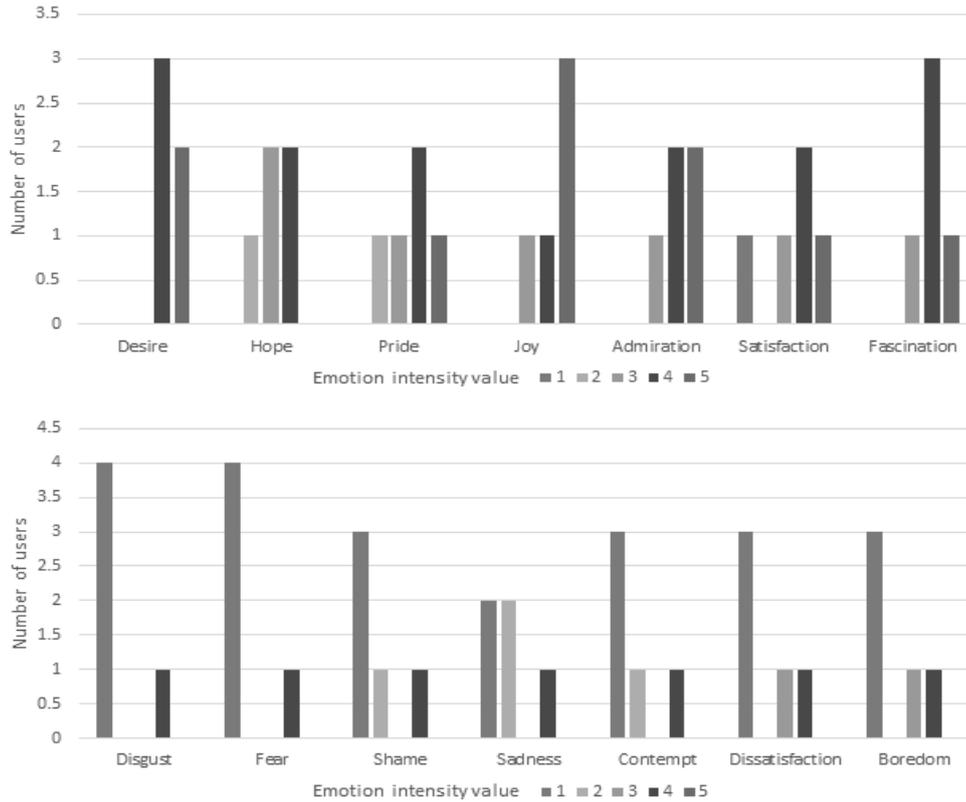


Figure 3: Users' expectations before the experience with the first OFR software prototype through the PrEmo UX capture method.

As predicted, most of the users did not expect to feel the negative emotions of the PrEmo circumplex. Surprisingly all the volunteers expected to feel desire, fascination, and joy with intensity up to three during their experience. Almost all subjects expected to elicit hope, pride, admiration, and satisfaction according to the foreseen expectations. Regarding most of the users' evaluations with PrEmo, the first OFR prototype did not fail to elicit positive emotions in all four emotional dimensions (Social, Material, Expectation, and Well-being) PrEmo can capture about software or product [30]. So, its interface design shall carry through the next prototype development. The UX evaluation through the SAM UX capture method should result in high pleasure and dominance dimensions measurement, while the arousal measurement should attain mediocre levels. Regarding the “Correlations for SAM Ratings and the Relevant Semantic Differential Factor Score with each of the six Adjective Pairs Associated with the Pleasure, Arousal, and Dominance Dimensions” [31] presented, 60 % of the volunteers felt somewhat or strongly in control, dominant or autonomous while elaborating XBRL financial reports, and 40 % of them reported indifference to dominance matters (do not feel in control, but also not cared for) to accomplish the activity. So, the OFR first prototype did not fail providing users the control they needed to perform the task. In the Arousal dimension evaluation all users reported feeling high levels of excitement to accomplish the activity. It might mean volunteers had put greater effort into keeping track of what they were doing or avoiding committing mistakes. So, that levels of arousal might not be desired for the task execution. The Pleasure dimension evaluation revealed about 20 % of the volunteers felt neither satisfied nor unsatisfied while elaborating XBRL financial reports, 20 % of the subjects felt unsatisfied, annoyed or unhappy performing

this task, and 60 % of them reported feeling satisfied, pleased or happy to accomplish the activity. The UX evaluation through the Emocards UX capture method (Figure 4) should result in indicators ranging from average pleasant to calm pleasant. 60 % of the users elicited the expected emotions through the Emocards. However, the high levels of arousal did not show up in the volunteers' answers through this method. It was possible to notice that most of the users' reported emotions did not drastically differ from the ones OFR software should have caused users to elicit. It meant OFR still needed some improvement to provide a better user experience. However, it did not have enough problems to cause catastrophic User eXperience results.

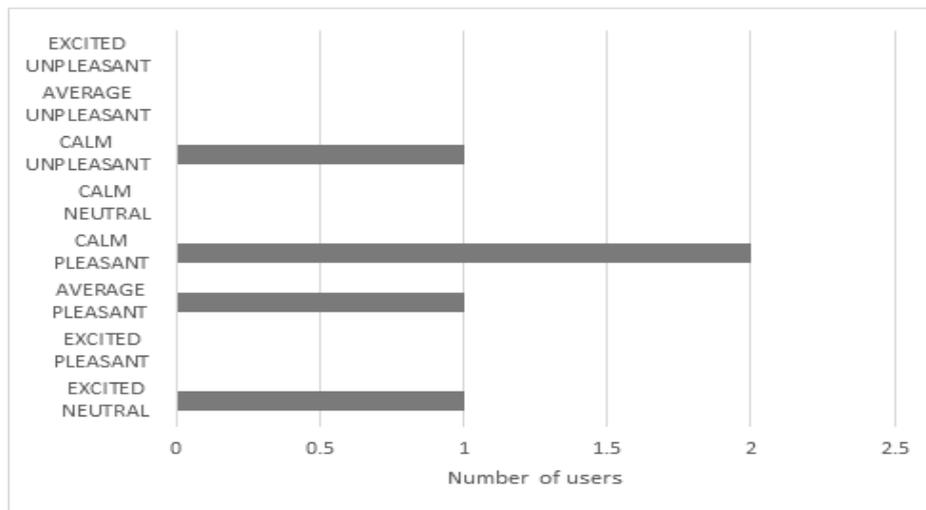


Figure 4: The overall UX after the tests with the first OFR prototype through Emocards UX capture method.

3.3. The OFR Redesigned Version

Even though the evaluation results showed positive evidence that the OFR's prototype was a valid tool to support the XBRL financial reporting task, it had to go under redesign to mitigate the identified problems. The HCI Design must prevent idiosyncrasies from influencing a project. Regarding that premise, OFR's redesign considered the changes identified through the UX and usability evaluation techniques. It also encompassed topics highlighted by at least three or more testers. The redesign process also aimed to fix all faults and bugs identified through the evaluation sections. In that way, the prototype's redesign consisted of:

- To remove the Report Models selection option from the Taxonomy menu and turn it into a new item of the OFR's menu bar;
- To add a new warn window message regarding the instance validation results;
- To remove the Elements menu from the OFR's menu bar;
- To add, in the menu bar, a new menu named Insert with four subitems: Insert Context, Insert Unit, Insert Account, and Insert Non Numeric element (Each subitem triggered the selection of the corresponding checkbox in the element management window);
- To add, in the menu bar, a new menu named Edit with four subitems: Edit Context, Edit Unit, Edit Account, and Edit Non Numeric element (Each subitem triggered the selection of the corresponding checkbox in the element management window);

- To add, in the menu bar, a new menu named Remove with four subitems: Remove Context, Remove Unit, Remove Account, and Remove Non Numeric element (Each subitem triggered the selection of the corresponding checkbox in the element management window);
- To reorder the element type fields position in every element manager screen;
- To update some of the software's messages;
- To implement a new menu item with tips about the procedures users had to perform with the software.
- To correct minor bugs caused due to the aforementioned alterations.

3.4. Usability Conclusive Evaluation Results

The following paragraphs describe the results related to the usability user tests conducted with the OFR's redesigned version. The five volunteers who signed up for the second research cycle evaluation sections were: professional accountants, without any previous XBRL knowledge, with Bachelor's degree (60%) or Graduate course (40%), that have never used an XBRL financial reporting tool, and have been composing financial reports for five or more years. Before starting the evaluation section, the evaluator provided a training section to all of the volunteers at their request. All of the enlistees fulfilled the task to create an XBRL instance with the OFR, and they made it in the first attempt. They also committed no errors to accomplish that task, and the OFR did not cause errors during their attempts. According to the dry-run test estimated values, those indicators should have reached one attempt, no user errors, and no system errors, respectively. The statistics about this task showed an average production time of 0.75 minutes, a success rate of 100 %, and an error rate of 0 %. The medium time users consumed to perform the XBRL instance creation, to complete the first attempt, and backtracking from an error were 1.1, 1.1, and 0 minutes, respectively. Three volunteers used only the two functions needed to fulfill the task, while two of them also used the software's help function. Users did not require the evaluator's help or used the OFR's supportive documentation. They also did not employ any extra time to discover how to apply the software functions. 80% of the volunteers accomplished the task of insert data with the OFR. One user forgot to insert the explanatory note in the report before going through the next task. The indicators also showed that the volunteers performed on average one trial to complete it. The users went through a total of 2 errors performing that task. The OFR incurred in no flaws during their attempts. Regarding the dry-run test results, the predicted rates for those indicators were one attempt, no user errors, and no system errors, respectively. 7.78 minutes was the average fruitful time, the success percentage was 80 %, and the error rate was 3.08 %. The medium time spent to finish the data insertion, to complete the first attempt, backtracking from a mistake were 8.22, 8.22, and 0.28 minutes, respectively. Three enlistees utilized all five functions needed to perform the second task. Only one of them also used the software's help function, while other one forgot to use the function to insert non-numeric elements. They also did not bid for evaluator's help or checked the software documentation during the referred activity. So, they did not spend the activity time learning OFR's functions. All of the users successfully finished the task to save and validate an XBRL instance with the OFR. They also made only one attempt to complete it. The volunteers did not incur in errors while working on that task. The OFR had error no occurrences during their attempts. Regarding the results obtained through the dry-run test, the predicted rates were one attempt, no user errors, and no system errors, respectively. The average fruitful period was 1.14 minutes long, succeeded by a success rate of 100 % and an error rate of 0 %. The

average time users wasted to save and validate an XBRL instance, to perform the first attempt, backtracking from an error consisted of 1.52, 1.52, and 0 minutes, respectively. 60 % of the users had to deploy 3 functions to fulfill the third task in which they should have used 2 functions. All of them did not inquire the evaluator's help or accessed the OFR's supportive documents while striving to perform the referred activity. They also consumed no extra time learning OFR's functions. The average accuracy completion for the whole activity was 95.56 %. As the average SUS score is 68, and the OFR score was 80 that matches an A- grade, it means OFR does not hold catastrophic usability problems. As SUS is a good tool to recognize unusable and usable systems, it is possible to classify this version of the OFR software as usable. The participants also gave the following feedback information that was not covered in the questionnaires and tools adopted:

- "Achei que mensagens do programa são muito grandes. Acho que as mensagens poderiam ajudar melhor a localizar os erros do relatório."
- "Quando o programa ficar pronto, ele vai ajudar a classificar as contas que temos de cadastrar em cada campo?"

In the XBRL instance creation task, one of the volunteers discontinued a semiosis because he/she did not discover a proper way to perform his/her communication. Other among them suspended a semiosis because he/she did not get the means to perform the succeeding communication input. The volunteers strived to grasp the communication process with the OFR through implicit metacommunication two times. Regarding the data inclusion task, the users stopped their semiosis three times because it was not possible to find a suitable way to communicate. The participants attempted to conjecture the OFR's communication process by inexplicit metacommunication four times, and noticed the communication failed because they completed a wrong interaction two times. One of them realized the communication went wrong because he/she interacted in the wrong context. One of the subjects left off a semiosis before accomplishing the aspired results to start another one to obtain the same effect because he/she decided to perform the semiosis in their fashion, even though they comprehended the HCI proposed solution. The users concluded a faulty semiosis, but they did not perform another one to reach the expected results, because they did not regard their achievement did not suit the aspired results four times. In the task to save and validate an XBRL instance, one of the participants broke the continuity of a semiosis because he/she has not got a way to do the succeeding communication input. The volunteers made an effort to understand the OFR's communication process through non-explicit metacommunication two times. One of the users caused the semiosis to stop because he/she did not comprehend the OFR's communication. One subject ceased a semiosis before achieving the aspired results to commence a new one with an identical goal because he/she had the ill to complete the semiosis in their form, despite getting the HCI offered solution. One volunteer ended an ineffective semiosis, but did not go for another try to achieve the wanted results, because he/she did not discern that it was not enough to match the desired results. Regarding the communicability evaluation results, the semiotic profile has not changed. The only updates are the designer's perception about:

- To provide more software function intuitiveness and communication to users who do not have previous XBRL knowledge.
- To provide less information overload on the screen to avoid users getting lost during their activities.

3.5. UX Conclusive Evaluation Results

In the pre-test questionnaire, the participants informed how they presumed to undergo the experience with the OFR software through the PrEmo UX capture method (Figure 5) before interacting with it. Regarding the PrEmo results, the OFR should cause volunteers to elicit hope, pride, admiration, satisfaction, desire, joy, and fascination. Boredom, dissatisfaction, contempt, sadness, shame, fear, and disgust should not be elicited or show up with low intensity in their replies.

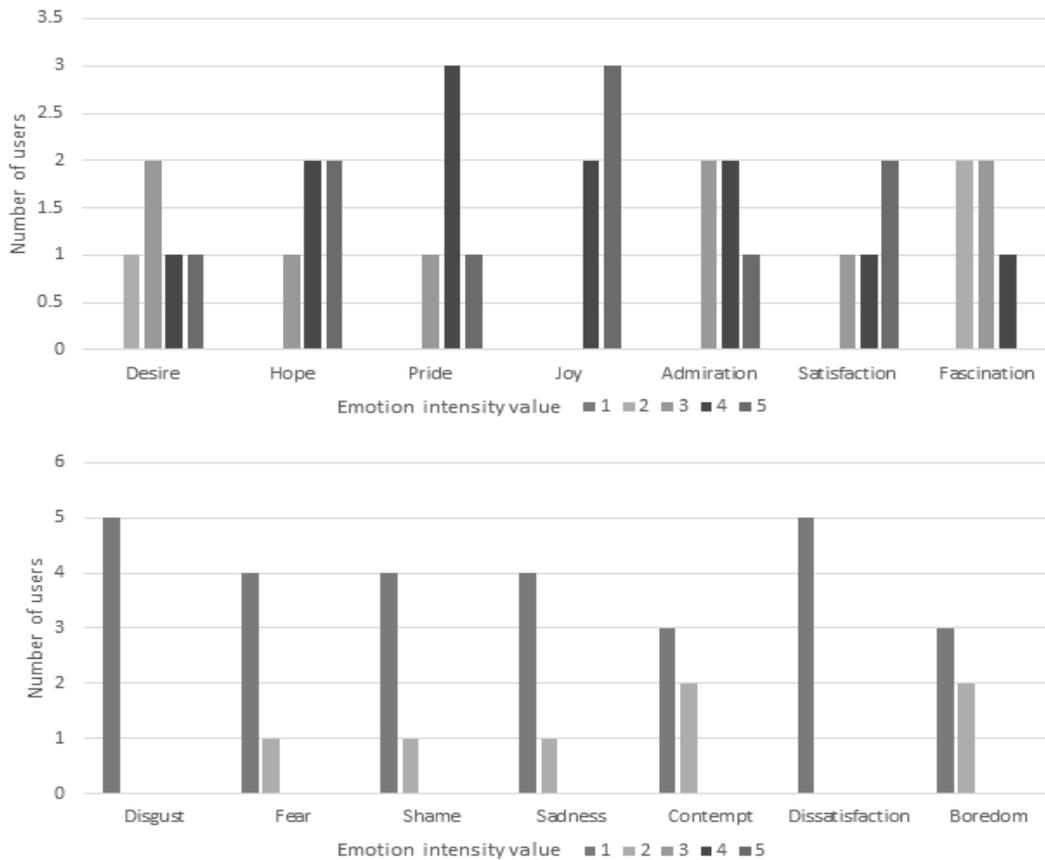


Figure 5: Users' expectations before the experience with the redesigned OFR software prototype through the PrEmo UX capture method.

According to the users' evaluations with PrEmo, the redesigned prototype did not fail to elicit the desired emotions in all four emotional dimensions the method can measure about software or product [30]. The SAM UX capture method should present high the pleasure and dominance dimensions measurement, while the expected arousal measurement was mediocre. The Emocards UX capture method indicators (Figure 6) varied within the expected set of emotions: average pleasant, calm pleasant, and calm neutral. High levels of arousal relate to the user's tension to avoid committing mistakes or to keep track of their actions, so it is not a positive aspect for the task under evaluation and should be avoided. Concerning the "Correlations for SAM Ratings and the Relevant Semantic Differential Factor Score with each of the six Adjective Pairs Associated with the Pleasure, Arousal, and Dominance Dimensions" [31] showed, 60 % of the users felt in control, dominant or

autonomous through the experience. 40 % of the subjects related to be indifferent to dominance matters (do not feel in control, but also not cared for) during their interaction.

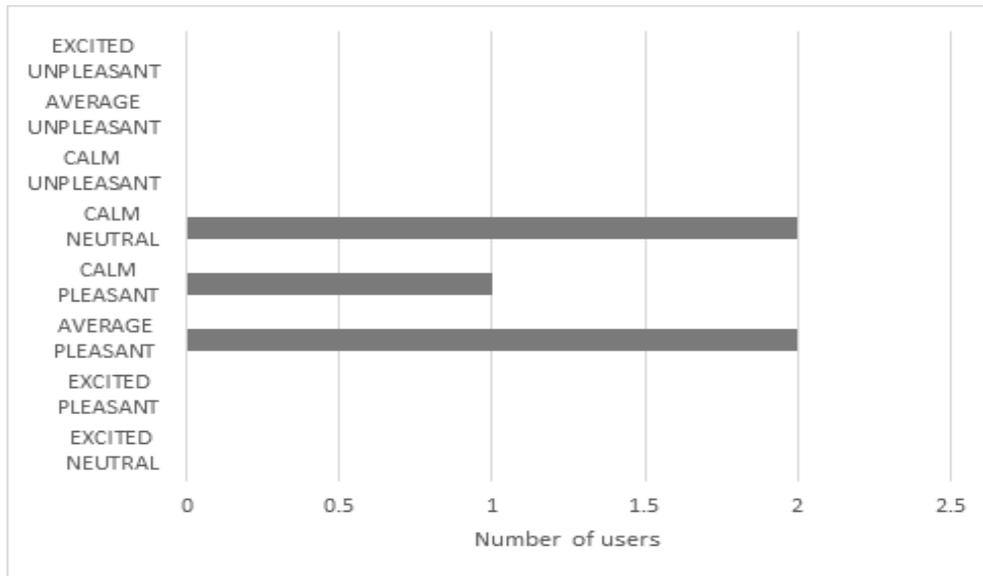


Figure 6: The overall UX after the tests with the second OFR prototype through Emocards UX capture method.

The results regarding the Arousal dimension revealed that nearly 60 % of the enlistees felt neither frenzied nor sluggish through the interactions with the OFR software, and 40 % of them informed feeling frenzied or jittery while performing the interactions. The outcome from the Pleasure dimension assessment exposed that 60 % of the users felt neither satisfied nor unsatisfied while performing XBRL financial reporting tasks with the OFR, and 40 % related feeling satisfied, pleased, or happy for accomplishing the activity goals. The redesigned prototype did not attain higher context control levels than the first one. However, it did not cause users to undergo an unpleasant experience and also allowed lowering the unexpected arousal levels elicited by the first prototype. Once 80% of the users from the first cycle had previous XBRL knowledge, it's possible to infer that their context control perception benefited from that. So, it is also conceivable that the lower context control perceived from volunteers of the second cycle has relation to their lack of previous XBRL knowledge. After analyzing the awaited effects most of the users' related emotions do not deviate from the ones OFR software should have induced users to elicit. It means OFR has provided a good user experience. In all three tasks users performed with the second version of the OFR, all the ISO efficiency indicators and the communicability had a better performance. So, the HCI design allowed volunteers: to spend less time lost or navigating through the software interface, to most of their productive time performing actions related to the task they had to accomplish, to stop needing external help to use the software, to undergo less semiosis interruption, to commit fewer errors while composing financial reports, to attain a higher accuracy rate. Regarding the UX evaluation, the HCI design practices helped out the software stop causing users to elicit undesired emotions or reactions as emotions from the negative part of Premo's circumplex; the unpleasant, high arousal, and low dominance of SAM's ratings; and the excited unpleasant Emocards. These results prove the initial hypothesis validity. HCI design is a "de facto solution" for improving G2G e-governance in the XBRL financial reports domain providing technological complexity abstraction and increasing process efficiency.

3.6. Dos and Don'ts of Conciliating XBRL and e-Governance Through HCI

In the first research cycle, the prototype provided only XBRL knowledge abstraction and some user communication messages. However, the results from the first evaluation proved that it had no catastrophic or severe usability and UX problems. But, it was still necessary to improve its easiness of learning and to make its metacommunication more intuitive for the users with little or no XBRL knowledge. It was not possible to map and delimitate a group of representative personas. Then, the OFR design followed a generalistic approach through the prototyping software development method. Even though that method provided rich information for redesign the first high fidelity prototype into the second one, the number of structures to change from one version to another can lead to excessive workload without a well-defined ending point. The action research approach matched the prototype method of cycle interactions. It shows evidence that combining that research methodology with this software development life cycle might be a good research practice and that they are not incompatible. It was also challenging to design software whose use context does not allow drawing personas. In institutions and entities across this country and the world, several different occupations play the role of XBRL financial reporter. As a result, the Designer had to put great effort into making the requirements for using the software to be the usage instructions, the concern about the financial report the professional has to elaborate on and which data shall it contain within. Choosing appropriated usability and UX tools showed to be very important to keep up the coherence between the users' needs and expectations and the prototype redesign. It avoided not addressing relevant problems that did not appear in the second OFR version. This study adopted only free technological solutions and platforms. It highlights both that it is possible to conduct relevant research without additional costs and compromising conduction results, and how it is important to make solutions available for free to research purposes. It is one of the reasons why OFR is a free and open code solution. Online forms and videoconferencing platforms showed a satisfactory performance to support HCI evaluations' conduction with geographically spread users. Even though the monolithic architecture approach provided easiness of implementation in both designed prototypes, it does not provide easiness to extend the software or implement new features. Future work should include adopting a new industry-standard architecture paradigm frequently used to build scalable and extensible projects, such as The Model-View-Controller (MVC) architectural pattern. As this research deployed a general-purpose architecture description language, there is no problem in documenting architectural changes regarding the language specificity context. So, regarding the MVC pattern, the Communication_Manager and the Error_manager would become View components, each of the task managers would become Controllers components, and the remaining ones would be part of Models components. According to [33], gathering volunteers for UX and usability studies can be challenging without hiring a recruiting agency. The OFR tests were not an exception to it. Even working with a small population (five people) and offering a prize draw as an incentive, it was not possible to have backup participants. So, no-show rates could have compromised the study results. Even the awareness quiz had to stay available online for over a month to attain a more representative number of respondents. UX and usability data validity regarding a product or solution is related to the measured indicators' nature and how representative the tests users' tasks and context are of the real use situation. It was already a challenging task to perform the tests within the use context of each geographically spread volunteer. However, COVID 19 pandemic turned it into an impossible task. The pandemic situation imposed new contexts with different impacts on how people perform their job tasks all

around the world (e.g., exchange the office context by the home office one). Within that period, it is not possible to perform usability or UX tests regarding the regular volunteer's context in non-pandemic days. Thus, this is a threat to this research data validity. Another context issue that treats the data validity was the need to run tests with XBRL financial reporting professionals from other countries in a testbed that represents Brazil's G2G e-governance. It incurred in the lack of internationality and universality of XBRL taxonomies, that is not addressed in this research. The information to implement several mechanisms in the source code for preventing the user from inserting inconsistencies in the instance file came from taxonomy's provider documentation external to the discoverable taxonomy set (e.g., date's input format, account value's precision, institutional code). The lack of taxonomy metadata might pose another relevant challenge to implement the OFR's complete version. In laboratory HIC evaluation sections, the setup time is part of the planning and preparation step and not a part of the section. But, not every remote user was using a computer properly prepared for the evaluation section. So, the users needed help to set up their computers before the training section within the section time. Even though Java is multiplatform, MAC OS internal security prevents users from executing Java files from unknown sources, as the OFR's prototype, and there is no way the user can override the security settings. As a result, it was not possible to conduct sections with volunteers who used MAC OS. It was challenging to train users and to help them with setup procedures within a half-hour in the evaluation sections. Future studies with geographically spread users shall consider reserving a specific moment to perform setup procedures. However, a researcher must avoid increasing the total section-time over 90 minutes (it is not a good practice according to the HCI evaluation literature [33]). All the volunteers scheduled their sections out of their work time and used their personal computers to download, execute and test the software prototypes. In the context of geographically spread users, this would pose a challenge in future works that aim to compare the OFR complete version with existing tools due to the following aspects:

- The lack of volunteers available to perform evaluation tasks in their job environments through with they have access to their entity's XBRL financial reporting tool;
- The bureaucracy to obtain permission from volunteers' entities to download and execute the OFR on their computers;
- The bureaucracy to obtain permission from volunteers' entities to record users' screens during their interaction with both software;
- The need to make one evaluator available at any time regardless of the volunteers' time zone.

People from abroad demanded a more intense training time to understand concepts related to Siconfi's financial reports. However, as they could successfully compose the XBRL proposed reports, regardless of their familiarity with Siconfi's financial reports, it was not possible to state that OFR did not match its purpose of mitigating the XBRL's knowledge complexity problem. Even though the redesigned OFR is still not commercially competitive, the evaluation results highlight the importance of HCI matters to make products and solutions characteristics and capabilities closer to the users' actual needs. As the second version attained higher indicators related to XBRL financial reporting task efficiency, it is possible to validate the research hypothesis. Because the OFR is in an early prototype development stage, it was not possible to compare it with completely functional existing XBRL financial reporting tools. However, the evaluation results proved that the OFR is a valid tool to support the XBRL financial reporting task even as a prototype. So, HCI matters are capable of

providing better task efficiency in the financial reporting area, thus showing its potential to improve e-Governance practices related to presenting institutions' accounting information to government oversight entities. Through a valid hypothesis and the research gap found in the literature review, it is possible to concern that the methodological procedures adopted for this research conduction consist of a valid way to perform similar studies regarding HCI, e-Government, and the XBRL financial reporting area. Or it could also be a starting point for such work development.

4. Conclusion

This paper presented the results and procedures adopted for this multidisciplinary research conduction in the areas of HCI, XBRL financial reporting, and e-Government. First, consulting professionals related to XBRL financial reporting through a questionnaire provided a better understanding of the HCI problems and demands the area faces. Then conducting a literature review should have allowed to identify how the academic community was addressing the existing problem, but it showed signs of research gap instead. Next, this research investigated if providing financial reporting professionals a software prototype whose design considered HCI matters (regardless of the user geographical dispersion) was enough to increase the task efficiency. To do so, it was necessary to compare the results of two usability and UX evaluations with two versions of the same financial reporting tool that aimed to abstract the XBRL knowledge need to create XBRL financial reports. This research considered developing both versions of the software prototype. While the first one focused only on minimal functional requirements, the second version focused on mitigating the HCI problems found through the formative evaluation section. According to results obtained in the conclusive evaluation, the measured task efficiency was greater with the second version of the software. It corroborates the validity of the research hypothesis, so this shows evidence that HCI design is a valid approach to improve financial reporting creation and e-Governance relations. The OFR prototype is also an initiative or a starting point to bring HCI practices to the financial reporting area. As it is an open-source tool, other researchers can adopt, adapt, extend, and explore its potential to improve XBRL financial reports creation (it is available in GitHub at <https://github.com/araao93/OFR/blob/main/OFR.zip>). Current XBRL financial reporting software providers could benefit from this study to optimize their products and provide better solutions to their clients through HCI practices. This research showed evidence that incorporating HCI design to XBRL financial reporting tools has the potential to increase task efficiency by bringing users tools that do what they need in the way they need. It also shows signs of the HCI design potential to improve the B2G and G2G e-Governance areas because bigger task efficiency prevents human and material resources waste and results in a more efficient exchange of financial information between government entities and businesses. Government institutions can benefit from the study's insights for evaluating their current G2G practices and find ways to improve them through HCI. Through this research, it was possible to identify successful practices to conduct HCI studies in the context of geographically spread users and some challenges to overcome. The detailed methodological description in this document and the corroboration of the hypothesis imply that adopted procedures have high reproducibility and might represent a valid starting point for similar studies that aim to improve B2G and G2G e-Government practices. Regardless of the hypothesis validity results, this research is already a stimulus for academic researchers to approach this problem under other perspectives or conduct similar studies that aim to improve B2G and G2G e-Government practices through HCI design practices.

4.1. Research Limitations

The study was performed only with five volunteers per section. So it might be necessary to conduct studies with larger populations to gather more representative and coherent data to support the hypothesis validation. With such a small population, it was not possible to assure the representativeness of the enlistees against the wide variety of professionals that plays the role of XBRL financial reporter. The pandemic situation imposed the world new realities and contexts. This study was performed under the "new normal" context of the home-office professionals and might be valid only within it. There is no way to assure the validity of the gathered data for the pre-pandemic contexts. The developed tool is still a prototype and could not be compared to actual XBRL financial reporting software to provide a more coherent efficiency impact analysis.

4.2. Recommendations for Future Work

As future works, this research highlights the following topics: to optimize the research approach on problem through the lessons learned, to reach a complete version of the OFR and perform comparative HCI evaluation studies with other XBRL financial reporting tools, to perform a field study in a government organization or business accessing the HCI adoption benefits that improve B2G and G2G e-Governance, to research ways to overcome the challenges highlighted in section 3.6, and to extend this research to other areas of G2G e-Governance out of the financial reporting domain while mitigating the research gaps found.

Acknowledgements

The authors Acknowledge the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financing this research.

References

- [1]. B. S. d. Silva and S. D. J. Barbosa, *Interação Humano-Computador*. 2010.
- [2]. I. N. da Costa. "USERBILITY: UMA TÉCNICA PARA AVALIAÇÃO DA EXPERIÊNCIA DO USUÁRIO E USABILIDADE EM APLICATIVOS MÓVEIS". MSc. dissertation, Instituto de Computação Programa de Pós-graduação em Informática, Universidade Federal do Amazonas, Manaus, 2016.
- [3]. Luis Rivero and Tayana Conte. "A Systematic Mapping Study on Research Contributions on UX Evaluation Technologies" in Proc. BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2017, pp. 41-50.
- [4]. XBRL INTERNATIONAL. "An introduction to XBRL: The basics of XBRL for business and accounting professionals" Internet: <https://www.xbrl.org/the-standard/what/anintroduction-to-xbrl/>, [Jun. 17, 2020].
- [5]. EUROPEAN SECURITIES MARKETS AUTHORITY, "Consultation paper: consultation paper on the regulatory technical standards on the european single electronic format (ESEF)," tech. rep., Paris, France, 2015.
- [6]. R. W. Caves. *Encyclopedia of the City*. New York, Usa: Routledge, 2005. 596 p.

- [7]. C. Hai and I. Jeong. *Fundamental of development administration*. S. L: Scholar, 2007.
- [8]. S. Bose and M. R. Rashel. "Implementing e-governance using oecd model (modified) and gartner model (modified) upon agriculture of Bangladesh" in *Proc. INTERNATIONAL CONFERENCE ON COMPUTER AND INFORMATION TECHNOLOGY*, 2007.
- [9]. TESOURO NACIONAL. "Apresentação" Internet: <https://siconfi.tesouro.gov.br/siconfi/pages/public/conteudo/conteudo.jsf?id=21>. [Jul. 17, 2020].
- [10]. A.R.D.A.A. Oliveira and P.C. da Silva. "Conciliating XBRL Financial Reporting and HCI". *American Scientific Research Journal for Engineering, Technology, and Sciences*, vol 74, pp.46-60, Nov. 2020.
- [11]. M. Deru and R. Neßelrath. "AutoUI-ML". in *Proc. INTERNATIONAL SYMPOSIUM ON PERVASIVE DISPLAYS*, 2015.
- [12]. M. Xiaofeng, W. Xufei, Z. Xiucui, L. Zhenliang and J. Ruisheng. "Research on IETM authoring platform architecture based on S1000D specification". in *Proc. IEEE INTERNATIONAL CONFERENCE ON ELECTRONIC MEASUREMENT & INSTRUMENTS*, 2015.
- [13]. S. S. Bhowmick, C. Dyreson, B. Choi and M. Ang. "Interruption-Sensitive Empty Result Feedback: Rethinking the Visual Query Feedback Paradigm for Semistructured Data". in *Proc. ACM INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT*, 2015.
- [14]. L. M. Garcia-Garcia, J. R. Gil-Garcia, and V. Gomez, "Citizen centered e-government? the case of national migration institute in the southern Mexican border," in *Proc. ANNUAL INTERNATIONAL CONFERENCE ON DIGITAL GOVERNMENT RESEARCH*, 2015.
- [15]. D. Shwartz-Asher, S. A. Chun, and N. R. Adam, "Social media user behavior analysis in e-government context," in *Proc. INTERNATIONAL DIGITAL GOVERNMENT RESEARCH CONFERENCE ON DIGITAL GOVERNMENT RESEARCH*, 2016.
- [16]. C. Maciel, C. Cappelli, C. Slaviero, and A. C. B. Garcia, "Technologies for popular participation: a research agenda," in *Proc. INTERNATIONAL DIGITAL GOVERNMENT RESEARCH CONFERENCE ON DIGITAL GOVERNMENT RESEARCH*, 2016.
- [17]. I. M. de Souza, C. Maciel, and C. Cappelli, "The model of accessibility to electronic government: applicability in dataprev," in *Proc. INTERNATIONAL DIGITAL GOVERNMENT RESEARCH CONFERENCE ON DIGITAL GOVERNMENT RESEARCH*, 2016.
- [18]. F. E. M. Quispe and M. M. Eler, "Accessibility recommendations for mobile applications: a contribution to the brazilian digital government standards," in *Proc. BRAZILIAN SYMPOSIUM ON INFORMATION SYSTEMS*, 2018.
- [19]. P. P. Adinda and A. Suzianti, "Redesign of user interface for e-government application using usability testing method," in *Proc. INTERNATIONAL CONFERENCE ON COMMUNICATION AND INFORMATION PROCESSING*, 2018.
- [20]. R. Baguma, "Usability evaluation of the etax portal for uganda," in *Proc. INTERNATIONAL CONFERENCE ON THEORY AND PRACTICE OF ELECTRONIC GOVERNANCE*, 2018.
- [21]. N. H. Basri, W. A. W. Adnan, and H. Baharin, "E-participation service in malaysian e-government website: the user experience evaluation," in *Proc. INTERNATIONAL CONFERENCE ON E-EDUCATION, E-BUSINESS, E-MANAGEMENT AND E-LEARNING*, 2019.

- [22]. N. Almrezeq, R. Alhamdan, M. Mahyub, and M. Alfayad, "An ex-ploratory study to investigate citizens' experience with e-government mobile services in saudi arabia," in Proc. INTERNATIONAL CONFERENCE ON INFORMATION MANAGEMENT, 2019.
- [23]. F. E. d. Oliveira and A. P. Freire, "Aceitação e uso de tecnologias interativas de governo eletrônico por pessoas com deficiência" in Proc. BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2017.
- [24]. L. A. Rodrigues and S. S. Prietch, "Analysis, redesign and validation of accessibility resources applied to an official electronic journal for the promotion of equal access to public acts," in Proc. BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2018.
- [25]. C. Maciel, L. Roque, and A. C. Garcia, "Mensurando a maturidade na tomada de decisão em sistemas de e-participação," in Proc. SIMPOSIO BRASILEIRO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS, 2019.
- [26]. TESOURO NACIONAL "Documentação" Internet: <https://siconfi.tesouro.gov.br/siconfi/pages/public/conteudo/conteudo.jsf?id=581> [Jul. 17, 2020].
- [27]. FUJITSU. "Interstage XWand Toolkit Evaluation Copy Download" Internet: <https://www.fujitsu.com/global/products/software/middleware/application-infrastructure/interstage/download/xwand-toolkit-evaluati-on-download.html> [Jul. 17, 2020].
- [28]. J. R. Lewis and J. Sauro. "The Factor Structure of the System Usability Scale" in Proc. INTERNATIONAL CONFERENCE ON HUMAN CENTERED DESIGN, 2009, pp. 94-103.
- [29]. P. Desmet, K. Overbeeke and S. Tax. "Designing Products with Added Emotional Value: development and application of an approach for research through design". The Design Journal, vol. 4, pp. 32-47, Mar. 2001.
- [30]. D. G. Caicedo. "Designing the new PrEmo: an empirical research on how to improve the emotion measuring tool" MSc. dissertation, Faculty Of Industrial Design Engineering, Delft University Of Technology, Delft, 2009.
- [31]. Margaret M. Bradley and Peter J. Lang. "Measuring emotion: the self-assessment manikin and the semantic differential", Journal of behavior therapy and experimental psychiatry, vol. 25, no. 1, pp. 49–59, 1994.
- [32]. J. B. Silva, and A. H. Kronbauer. "A Study of Hedonic Experience Related to UX Capture Techniques" in Proc. BRAZILIAN SYMPOSIUM ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2018. p. 1-10.
- [33]. DEBORAH HINDERER SOVA and JAKOB NIELSEN. (2020) 234 Tips and Tricks for Recruiting Users as Participants in Usability Studies. [On-line]. Available: <https://www.nngroup.com/reports/how-to-recruit-participants-usability-studies/> [11 Aug, 2020].