Scenario and Task Based Interview to Evaluate Usability of Computer Assisted Data Collection

Luiz Agner[1], Patricia Tavares[1,2], and Simone Bacellar Leal Ferreira[2]

[1] IBGE - Instituto Brasileiro de Geografia e Estatística,
Rio de Janeiro, RJ, Brasil
[2] UNIRIO, Departamento de Informática Aplicada,
Universidade Federal do Estado do Rio de Janeiro,
Rio de Janeiro, RJ, Brasil
{luizagner,pztavares}@gmail.com, simone@uniriotec.br

Abstract. This article aims to present the method of usability evaluation called Scenario and Task Based Interviews (STBI). The method was proposed to add flexibility to field usability testing, so that they could be applied to the context of The Brazilian Institute of Geography and Statistics (IBGE). IBGE is the institute of Brazilian central Administration that performs the Census and other important official demographic and economic data collection. This evaluation technique was specifically designed to be implemented with the participation of interviewers who use PDA (personal digital assistants) to perform data collection for statistical research in Brazil. The authors analyzed the usability of the application developed for PDA to support the Continuous National Household Sample Survey (Continuous PNAD). The method proposed in this paper represented a mix of four approaches to usability evaluation.

Keywords: usability, PDA, method, data collection, interaction design, statistics.

1 Introduction

With the advancement of information technology, people and organizations increasingly need systems built with quality. Considering it is through the interfaces that people communicate with systems to perform their tasks, they must increasingly be designed with focus on usability [10].

There are different methods of usability evaluation: there are those without the presence of users - the methods of inspection or analytical methods - and there are those that involve users, called observation methods or tests with users. These can be made in the context of use or in a monitored environment such as usability labs [7].

The method proposed in this paper represents a mix of four approaches to evaluation: ethnographic observations, usability testing using a portable lab, semi-structured interviews and heuristic inspection.

In the first phase of the proposed method, the users were observed in their primary context of use: using the PDA in field to interview citizens in their homes. The second phase consisted of an observation made in a semi-controlled environment, where users

A. Marcus (Ed.): Design, User Experience, and Usability, Pt II, HCII 2011, LNCS 6770, pp. 349–358, 2011.

were interviewed and observed in their workplace, performing tasks that simulate real situations that frequently occur in their field activities. These tasks were performed using a portable usability laboratory, where all interactions were videotaped.

The evaluation method was developed and applied for the assessment of the usability of software developed to support interviewers during the Continuous National Household Sample Survey (Continuous PNAD) conducted by The Brazilian Institute of Geography and Statistics (IBGE), the statistical institute of Brazilian central government. IBGE is the institute that performs the Census and other important official demographic and economic data collection.

The applied method, called Scenario and Task Based Interviews (STBI), sought to preserve essential characteristics of the scientific method to lend credibility to its findings and to presentation of results. It was similar (but not identical) to field usability testing presented in human-computer interaction literature [3].

This method was created to meet the conditions and peculiarities of the specific context and had been previously experienced in evaluating the usability of a 2010 Census application [2].

2 Continuous PNAD Survey

Continuous PNAD is a new survey which aims to permit a continuing research on labor and income of the population. It is the result of the merger of two polls: the Monthly Employment Survey and the National Household Sample Survey. IBGE began testing its system and methodology since October 2009 in the states of Pará, Pernambuco, Rio de Janeiro, Sao Paulo, Rio Grande do Sul and Distrito Federal, continuing in 2010, in Rio de Janeiro. In 2011, the survey will be officially launched across the country [9].

Fig. 1. Graphical interface for the Continuous PNAD mobile device: preliminary design – version 1.0.9

Continuous PNAD will be conducted from a sample of approximately 179,000 households and has several questions about work and income. It also will investigate additional subjects such as: adult and youth education, migration, vocational education, child labor, fertility, social mobility, marriage, health, food safety, information technology and communication, income and time use.

An electronic questionnaire has been created to support data collection and management. Our study examined the ease of use of the application developed for mobile device in order to generate design recommendations to make its interface more efficient and appropriate for users (Fig. 1). This study was also part of a master degree dissertation carried out in Universidade Federal do Estado do Rio de Janeiro.

3 Technology in Context of Data Collection

According to Greene [8], the introduction of a technology can transform the context of an interview to collect statistical data contributing to the feeling that this is an important event for the informant. In some cases, curiosity about the technology can draw more people to observe or participate in interviews. Sometimes it helps to "break the ice", making it easier to talk with other members of family or community. In some rare situations, technology can inspire fear or anxiety.

The methods of data collection assisted by computers are known by the terms CADC (Computer-Assisted Data Collection) or the European term CADAC, CASIC (Computer-Assisted Survey Information Collection), and CAI (Computer-Assisted Interviewing). Traditional methods of paper and pencils are often denoted by PAPI (Paper-And-Pencil Interviewing).

The main feature of the interviews supported by computer is that questions are oriented in the correct order, following an algorithm available in the interactive program on computer screen. The software has the intelligence to reconfigure the order and type of questions, based on previous answers or the information it already has on the respondent. Responses are directly entered on the computer by interviewer or interviewee [11]. As an alternative to paper questionnaires, the CADC is well accepted by interviewees and interviewers. It permits data quality improvement, especially when complex questionnaires are used. In general, respondents react positively to computer use during an interview: they attribute a greater degree of professionalism to the survey. Social interaction with the interviewer is described as comfortable [11].

4 Methodology

STBI study is basically an applied qualitative research and has six steps: (a) Literature and document research, (b) Ethnographic observation, (c) Sample selection, (d) Scenario and task based lab sessions, (e) Analysis, (f) Communication of results. These steps will be briefly described as follows:

4.1 Literature and Document Research

Initially, the team sought to understand the principles of usability with a focus on the usability of mobile devices. We studied and selected some tools and software that support professional interface evaluation, such as video recorders and applications that sync PDA with the notebook, as well as the Continuous PNAD questionnaire manuals, and videos on best practices to approach the informant.

4.2 Ethnographic Observation

Authors have scheduled a field work follow-up, where the application could be observed in real use. Our intention was to identify problems and understand the difficulties and advantages of collecting data using the PDA. To understand how the users perform their field work, the authors followed three teams of interviewers in the household interviews during the Continuous PNAD test.

Firstly, authors participated in a field visit in the Leblon neighborhood, a high middle class residential area. Then we followed a work team in various districts in Rio de Janeiro suburbs. Then the authors undertook a visit to an IBGE data collection agency in Centro to interview the employees. After that, a field visit was undertaken to Tanguá (Fig. 2), within the State of Rio de Janeiro, an area with rural characteristics, difficult transportation and no cell service.

Moreover, the authors followed a 2010 Census interviewer using a similar device in middle class area and slum of Belo Horizonte. Ethnographic observations were recorded through written notes, video and photos, presenting the view of the interviewer about his/her experience using the electronic questionnaire.

According to Cooper [6], contextual studies developed in the process of interaction design should bring the spirit of ethnographic research and apply it on a micro level.

Ethnography is not a method, is more properly a category of research in Human Computer Interaction. The ethnographic study is a powerful means of identifying the true nature of the work. It is very common for users to perform their tasks differently from what was prescribed. Ethnographic methods can discover valuable and unusual facts that would never be identified by in-house methods [4].

Fig. 2. Record of ethnographic observation: Continuous PNAD interviewer focuses on a selected household in Tanguá, a rural characteristic sector in Rio de Janeiro

4.3 Sample Selection

To analyze the target audience of Continuous PNAD workers, it was developed an online questionnaire with ten closed questions on the user's profile, their experience with technology and other surveys. It was asked whether or not he/she would like to participate in usability interviews. We chose to define our sample of participants with six individuals carefully recruited from the population of actual users. The online questionnaire was included in the Continuous PNAD management system. Users were informed about the benefits that would be obtained with usability research.

The online questionnaire was available for 30 days. After this period, results were aggregated and analyzed, yielding a total of 57 responses, where 29 users (51%) said they would like to participate. The general profile was predominantly male (70%) aged between 18 and 29 years (66%) with incomplete higher education (44%) or complete (39%), with high experience in PDA (33%) or moderately high (33%).

Based on given profile, we selected two females and four males who were coursing graduation in history, social welfare, veterinary, geography or biology. Scheduling was conducted by telephone and the interviews occurred in the Rio de Janeiro unit.

4.4 Scenario and Task Based Interview Lab Sessions

The term Scenario and Task Based Interview (STBI) was coined to avoid creating anxiety in the participants and to avoid suggesting that they were being tested. In addition, participants are accustomed to the term "interview" because of their own work in Continuous PNAD.

Comprehension of informant approach situations represented a valuable aid to build scenarios of use (common situations that occur in field work, written in the vocabulary of the user). Meetings with application development team were held to consolidate eight scenarios/tasks addressing the main field situations (Table 1). It was also provided a fictitious database to help performing the tasks.

The type of usability lab used to support lab sessions was the minimalist portable setup [13]. There is no specific room dedicated to testing: the equipment and software are taken to different locations in a notebook. Among the advantages, it is easier to recruit participants. This configuration is considered by Rubin and Chrisnell [13] as the most suitable for organizations that are beginning to test because it presents the best cost-benefit, dispensing a physical plant.

In STBI lab sessions, participants were encouraged to externalize their thoughts as they worked on tasks and activities were recorded (think-aloud protocol). As the device was synchronized with the notebook, an image of the PDA screen appeared on the screen of the notebook. A camera captured the user reactions as the whole process of usability evaluation was recorded by audio and video software.

During sessions, users were provided with the option to stop and freely register their storytelling, comments, criticisms or suggestions. So the method differs from traditional usability testing which is focused on the measurement of performance. It is essentially a method of evaluation that generates qualitative data and insights.

Table 1. Scenarios and tasks created for STBI lab sessions

Scenario	Task
You are in the field, conducting Continuous PNAD interviews. You opened a home and were informed that there lives a family with three people.	Create in the application a family formed by these three persons and confirm the information in the PDA.
By continuing to interview, you discovered that there is one extra resident to be recorded in this household.	Change the relationship of people and add grandmother Iracema with the given profile.
On the same day, you continued your work and opened another home, now with five people forming two families.	Create two families formed by these five people.
You spoke with another informant and concluded that one of persons should be excluded from the list.	Delete Octavio's brother Severino from the list of people who lives in the household.
You are now going to continue an interview that had begun in the previous month. Thus, you will need to open that file in the backup area.	Save up the current interview and open the previous month file.
You noticed that you would need to correct the information already recorded.	Change the income data of residents according to given figures.
During the same interview, you had to consult some answers to make sure they were correct.	Visualize what was the answer to the question about the number of hours worked by Mauro during reference week.
You had to stop the interview at the request of the informant. You have to call another day to complete your questionnaire.	Note the informant telephone number and type in the appropriate field a brief reminder to call him/her next Thursday between 8:00am and noon.

Upon completion of all tasks, participants answered a post-test questionnaire consisting of eleven closed and three opened questions. Based on these responses, the researchers performed an open interview, recorded on audio, giving the user the opportunity to explain their suggestions further.

In each session, two researchers were present observing and noting problems, comments and behaviors. Some sessions were attended by an invited member of application development team.

4.5 Analysis

STBI study adopted top-down analysis: six user sessions produced 48 videos which were carefully reviewed by researchers who conducted heuristic inspection to identify all problems and related suggestions by the users.

After registering all problems and suggestions, researchers used a top-down data grouping strategy beginning with Nielsen's ten heuristics categories [12] which describe general usability principles. Top-down approach began from his well-known

range of established principles that could provide consistency to the analysis and to the interpretation of data [1].

5 Research Results

5.1 Ethnographic Observations Results

The authors could verify that PDA user needs to withstand harsh environmental conditions (heat, cold, humidity, drought, and light) depending on local usage. It is true that mobility imposes physical, visual and cognitive limitations to users [5]. Added to these factors, the difficult access to households (high-risk areas, distant places) and upper middle class condominiums where more and more restrictions to interviewers are imposed due to security concerns. To register these circumstances and facts the authors wrote ethnographic reports pointing out main events [14]. The color contrast of screen in sunlight is the most often cited usability problem: researchers photographed the device under the incidence of sunlight to emphasize the severity of the problem (Fig. 3).

We could also observe that ethnographic study is a powerful mean of identifying the true nature of the work. It is very common for users to perform their tasks differently from what was prescribed. In the case of computer-aided data collection ethnography can be useful to show if users spell out their questions using informal or popular language, so different from what is proposed in the official electronic questionnaire. Or, again, you will notice if users try to make a bypass of the normal operation of the application in order to avoid slowdowns, crashing or other technical issues that may impact usability.

Fig. 3. Field use observation: the device under the sunlight has legibility impaired

5.2 Portable Lab Sessions Results

Using post-task questionnaires, application was thoroughly evaluated by users. They considered it was simple, easy to learn and use. Also, 67% pointed out the simplicity, 67% rated its menus, buttons and functions as simple to use, and 50% answered that the messages are well written. Despite these figures, interactions in videos showed that some usability principles were violated to varying degrees.

Usability problems recorded in these videos were also analyzed through a heuristic inspection and were grouped together according to Nielsen´s heuristics categories [12]. Violated heuristics were identified as the following (figures show the percentage of occurrences): Match between system and the real world (13%); User control and freedom (11%); Visibility of system status (10%); Consistency and standards (8%); Flexibility and efficiency of use (6%); Aesthetic and minimalist design (6%); Error prevention (3%); Recognition rather than recall (3%); and Help users recognize, diagnose, and recover from errors (2%).

In addition to these usability principles, the analysis and sorting of results showed us that four extra categories should be created to permit a better classification of our findings. Thus, the previous list was added with 4 specific categories in order to better describe our usability issues and emphasize some important findings. See table 2 which also indicates their percentage of occurrence. It is important to note that these issues were detected during the testing phase and they were carefully considered for subsequent correction by development team.

6 Final Considerations

The impact of computer-aided interviews on data quality has been systematically evaluated by statistical institutes in diverse countries. This is an evaluation study based on Human-Computer Interaction theory. This article presented a method for usability evaluation consisting of scenario and task based interviews designed to be applied along the interviewers who used PDA (personal digital assistants) during the experimental phase of the Continuous National Household Sample Survey (Continuous PNAD).

The proposed method had essentially two major steps: observing users in their context of use (ethnographic observations) and performing common tasks while users are interviewed and interact with the interface in a notebook lab. Our experience has shown that ethnography can be considered a very important phase for the interface design because it reveals the true nature of work and avoids misunderstanding or idealization.

In our ethnographies it became evident that the users - not infrequently - conduct their work in a different way from what had been prescribed. For example, they tend to formulate questions using popular colloquial words, instead of reading the official version. They also reverse the order of questions as posed by the software, because they intend to reduce its time duration. They may also want to make it less tedious and more natural to the informant. Evidences from field studies reinforce the conclusions reached in STBI sessions in a semi-controlled environment. This allows us to affirm that both techniques have great potential when applied in a complementary manner as shown.

The method was considered quite satisfactory, since it involved usability testing of low cost. Problems were detected during the early testing phase and carefully considered for subsequent correction by development team. Moreover, it is important to note that data analysis led us to create four extra heuristic categories, where the most outstanding is concept usability - that addresses difficulties of interviewers and informants to deal with specific questions of Continuous PNAD.

Table 2. Additional heuristic categories in the Continuous PNAD - testing phase

Additional Categories	Examples of Associated Problems
Response Time (11%)	Slow response of the virtual keyboard.
	Slow opening of the residents frame.
	Response time beyond acceptable (30 seconds).
Explicitness, grammar and spelling (2%)	Messages in English (not in Portuguese) and containing misspellings.
	Some abbreviations, titles or text labels are incorrect or incomplete.
Bugs and crashes (19%)	System crash.
	Bug occurs when a resident is deleted.
Concept usability (6%)	Confusion between the concepts of the family chief and principal head of household.
	Interviewer and informant have difficulty to answer the question about color or race.

At the phase of reporting the results, besides scientific papers, the conclusions of this study were presented at seminars for other software development teams, aiming to spread and institutionalize usability methodology in order to contribute to make the procedures for data collection more efficient, comfortable and safe.

Acknowledgement

The authors acknowledge the support of the IBGE Executive Board, the teams of the Rio de Janeiro unit and DI-Codes, without whom this research would not have been possible. The opinions expressed in this article are the sole responsibility of the authors and do not necessarily express the views of the IBGE.

References

[1] Agner, L.: Arquitetura de Informação e Governo Eletrônico: Diálogo Cidadãos-Estado na World Wide Web – Estudo de Caso e Avaliação Ergonômica de Usabilidade de Interfaces Humano-Computador. Rio de Janeiro. Tese de Doutorado – Departamento de Artes e Design, Pontifícia Universidade Católica do Rio de Janeiro, 354p (2007)

[2] Agner, L.: A experiência do usuário no dispositivo de mão do Censo Demográfico de 2010. 9º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design. Universidade Anhembi Morumbi. São Paulo. 13 a 16 de outubro (2010)

[3] Barnum, C.M.: Usability testing and research, 428 p. Pearson Education, New York (2002)

[4] Burke, J., Kirk, A.: Choosing Human-Computer Interaction (HCI) Appropriate Research Methods – Ethnographic Methods, University of Maryland (2001), `http://otal.umd.edu/hci-rm/ethno.html`

[5] Burzagli, L., Billi, M., Palchetti, E., Catarci, T., Santucci, G., Bertini, E.: Accessibility and Usability Evaluation of MAIS Designer: A New Design Tool for Mobile Services. In: Stephanidis, C. (ed.) HCII 2007, Part II. LNCS, vol. 4555, pp. 275–284. Springer, Heidelberg (2007)
[6] Cooper, A., Reimann, R., Cronin, D.: About Face 3: The Essentials of Interaction Design. Wiley Publishing, Inc., Indianapolis (2007)
[7] Dias, C.: Usabilidade na Web: criando portais mais acessíveis, 2ª edn. Alta Books, RJ (2007)
[8] Greene, P.D.: Handheld computers as tools for writing and managing field data. Field Methods 13(2), 181–197 (2001), http://fmx.sagepub.com/content/13/2/181
[9] IBGE - Instituto Brasileiro de Geografia e Estatística. Quinto Fórum do Sistema Integrado de Pesquisas Domiciliares por Amostragem, http://www.ibge.gov.br/home/estat'istica/indicadores/sipd/quinto_forum/PNAD_Continua_Questionario.pdf
[10] Leal Ferreira, S., Nunes, R.: e-Usabilidade. LTC, Rio de Janeiro (2008)
[11] Leeuw, E.D., Hox, J.J., Snijkers, G.: The effect of computer-assisted interviewing on data quality. A review. In: Blyth, B. (ed.) Market Research and Information Technology. ESOMAR Monogaph, pp. 173–198. Esomar, Amsterdam, http://igitur-archive.library.uu.nl/dissertations/2006-1024-200101/c3.pdf
[12] Nielsen, J.: Ten usability heuristics, http://www.useit.com/papers/heuristic/heuristic_list.html
[13] Rubin, J., Chrisnell, D.: Handbook of usability testing: How to plan, design and conduct effective tests. Wiley Publishing, Inc., Indianapolis (2008)
[14] Tavares, P., Agner, L., Leal Ferreira, S.B.: Avaliação de Usabilidade de Dispositivos Móveis de Coleta de Dados Domiciliares Através de Entrevistas Baseadas em Cenários e Tarefas. In: IHC 2010 – IX Simpósio Sobre Fatores Humanos em Sistemas Computacionais, Belo Horizonte, MG, Brasil, October 5-8 (2010)