10.20378/irb-58491

Improving User Experience through Task Design and Evaluation Metrics in Research Projects

Xiaojun Yuan
College of Computing and Information
University at Albany, State University of New York
135 Western Avenue, Albany, NY 12222
xyuan@albany.edu

Abstract. In UX requirement practice and research, it has been widely acknowledged that usability and user experience should be addressed in every aspect of the product development process. However, challenges exist on how to consistently and explicitly conduct the usability evaluation and improve user experience. In this paper, I will present how we have tried to improve user experience with interactive information systems by employing a consistent task design and a set of evaluation metrics. An integrated interactive information retrieval system is introduced as an example. Task design and user experience metrics used in a user-centered experiment evaluating this system are described. The task design scheme and user experience metrics have been applied in other user experiments to test the effectiveness of some other information systems and they have been proved effective. Successful experience and lessons learned are discussed. Some further thoughts and suggestions on future work are also presented.

1 Introduction

In a study employing integrative approach to requirements analysis in a user-centric design framework, [5] pointed out the lack of consistency among designers could be an important reason leading to the limited usefulness of the designed system. There exist many reasons that may possibly lead to the ill-defined usability/UX requirements, such as constantly changing instructions, vague user needs, budget concerns, and time constraints. In such cases, both stakeholders and Usability/UX researchers/practitioners take the risk of failing the projects. In 2014, the Standish Group [9] found that more than 50% of projects were reported as challenged and over 30% of projects had to be terminated because of major delays, budget concerns, and so on. There are only about 16% software projects were reported successful in terms of time and budget. It has been an urgent issue for user experience researchers to understand how to maximize the possibilities for a system/software project to be able to succeed in subsequent related projects in the field.

Research has shown that the major causes of project failure are requirements inconsistencies and missing user involvement ([1],[6], [9]). As [1], [5] and [7] stressed, user involvement is a critical factor in the product development process. We should get users involved from the very beginning of the project until the end of it: getting their ideas to better understand their requirements and needs before we design the product, getting their experience of the prototype, and asking them to test the final product and give us their feedback. If the product development is an iterative process, our users should be involved in every single stage of this process.

In the user-centered product development process, task design and analysis is a very important step. It has been found in earlier research that task type has an impact on usability and user experience of information systems [11]. Task analysis [3] provides a technique for designers to characterize their understanding of user behavior in completing a task [8]. Besides task design and analysis, another critical step designers should take seriously is selecting appropriate evaluation metrics in measuring user experience and usability. Rogers, Sharp, and Preece [7] gave designers a detailed description of case studies on topics related to user experience and usability requirements.

The paper explores new methods to address the problem of ill-defined UX requirements in the domain of information retrieval systems and applications. It starts with an introduction of an integrated interactive information retrieval system which was tested in a user-centered experiment. Then a discussion about the user experiment details with particular focus on task design and user experience evaluation metrics will follow. The task design scheme and user experience metrics have been applied in the evaluation of other interactive information systems and they have been proved effective.

2 An Interactive Information Retrieval System

In [11], an interactive information retrieval system was introduced. This system was designed based on earlier findings from [10], and can support multiple information seeking strategies in a single system framework (see Figure 1 through Figure 4). A within-subjects experiment was conducted to compare this system with a baseline generic interactive information retrieval (IIR) system, designed to support searching through query specification.

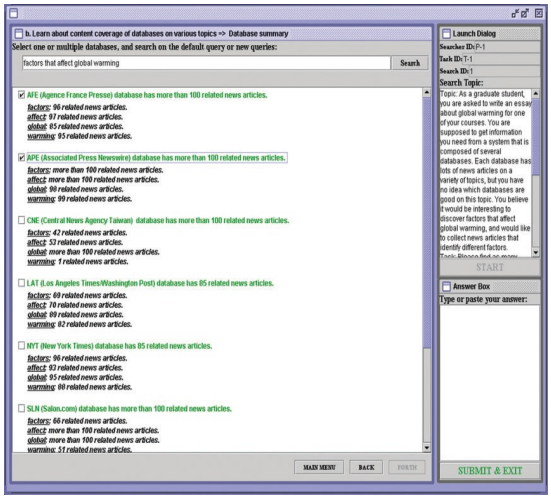

Figure 1. Integrated system, database summary display (c.f. [11])

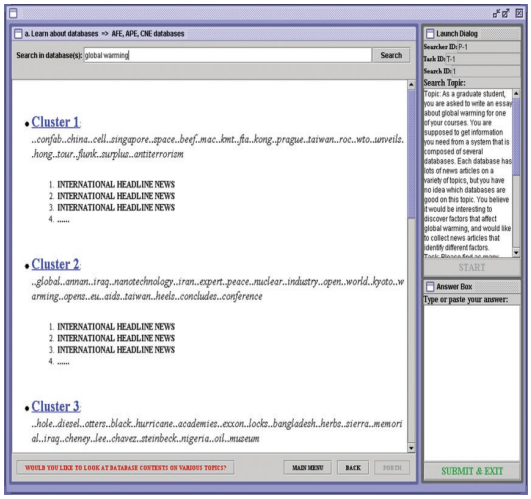

Figure 2. Integrated system, clustered results display (c.f. [11])

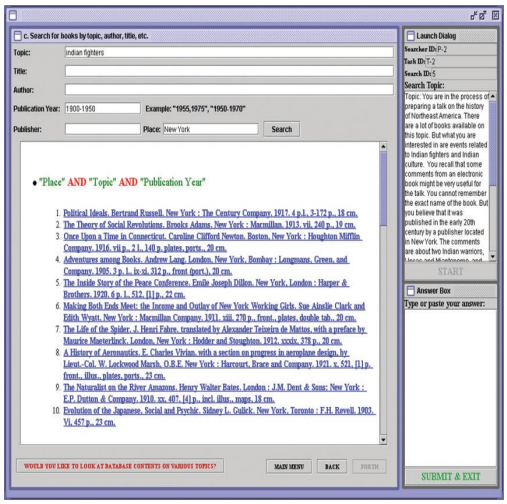

Figure 3. Integrated system, book search results display (c.f. [11])

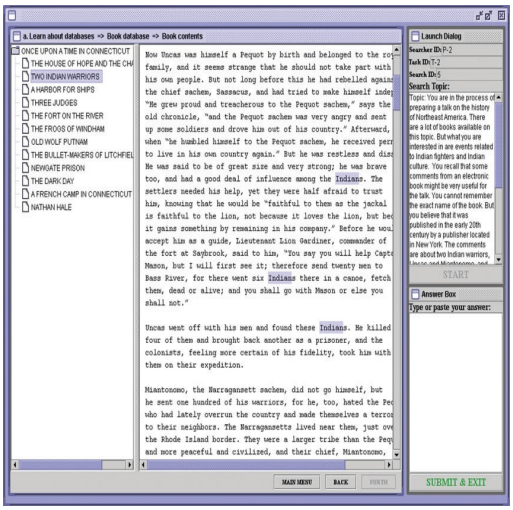

Figure 4. Integrated system, book table of contents display (c.f. [11])

As described in [11], the integrated system begins with a screen containing four options: (a) learning about the databases, (b) learning about the content coverage of databases with respect to a given topic, (c) searching for books on a specific topic, and (d) searching for news articles on a specific topic. Choosing option (a) leads to a screen which lists the names of the databases and the number of documents in each database (see Figure 1). If the user chooses news article databases, the results of the search are displayed as a list of clusters, with highly ranked cluster terms as a label (see Figure 2). If the user chooses a book database, a ranked list of complete citations of the retrieved books will be shown. Clicking on a book leads to a display of the table of contents of the book in a column on the left of the screen. Clicking on one of the items in the table of contents displays that part of the book (see Figure 4). Choosing option (b) leads to a screen which lists 10 queries, each

related to one of the eight test topics, and the two training topics. If the user chooses a news article task, it lists each database and the number of occurrences of each query term in that database (see Figure 1). Choosing option (c), searching for books on a topic, leads to a fielded search/results screen. The user enters values in the fields to perform a query; the results (complete citations) are displayed according to how well they satisfy the Boolean conditions of the query (see Figure 3). Clicking on any citation leads to the display of Figure 4. Choosing option (d) leads the user directly to the search/results screen shown in Figure 2.

3 Task Design

When it comes to real user experiments, sometimes it is challenging to come up with tasks that are similar to real life-time task activities due to the time limitation, lab environment, and diversity of different targeted user groups. Researchers have tried to find a solution to this issue. Simulated Work Task [2] is one of them. Borlund designed the simulated work task situation that can describe users' information need through relevant information, including the source of need, the environment, the information problem, and the goal of the search [2]. We followed the principle of the simulated work task situation model in the user experiment. We categorized the tasks into two different types in order to identify a variety of user behavior patterns. These two types of tasks were designed by following the simulated work task situation model by Borlund [2]. Task type 1 was defined as a task that leads users to engage in scanning, and then searching, while task type 2 was defined as a task that leads users to engage in searching, and then scanning [11].

In the following, we focus on describing how we constructed the tasks. The integrated interactive information system introduced in Section 2 is used as an example.

Before we started evaluating the interactive information retrieval system, we first identified tasks which could lead users to engage in a variety of information seeking strategies (e.g. scanning, and searching) using such a system. We designed the tasks based on our own knowledge of different users' information seeking problems. The specific topics for the tasks associated with book searching were designed so that answers were available to be retrieved from our test collection.

Here is a list of selected task examples we used in the above-mentioned completed user study [11].

3.1 Task type 1: (Finding news articles)

Topic: As a graduate student, you are asked to write an essay about high blood pressure for one of your courses. You are supposed to get information you need from a system that is composed of several databases. Each database has lots of news articles on a variety of topics, but you have no idea which databases are good on this topic. You believe it would be interesting to discover methods that reduce high blood pressure, and would like to collect news articles that identify
different methods.

Task: Please find as many different methods as possible. For each method, please copy the title or link of the article which discusses that method, and paste it to the answer box. For each article that you copy, please type or copy the method(s) that it identifies. If an article discusses more than one method, you only need to copy and paste the article once. If there are several articles which discuss the same methods, you only need to copy and paste one such article.

3.2 Task type 2: (Finding comments)

Topic: You are in the process of preparing a talk on the history of Rome. There are a lot of books available on this topic. But what you are interested in are the wars of Julius Caesar. You recall that some comments from an electronic book might be very useful for the talk. You cannot remember the exact name of the book. But you believe that it was published by a publisher in New York. The comments are about the strategies that Caesar used on the battle field to win the Battle of Pharsalia. You cannot remember the exact comments, but would like to quote them in your talk.

Task: Please find the relevant comments from the book, copy the one best paragraph then paste it into the answer box. Also, please copy the title of the book then paste it to the answer box.

4 User Experience Evaluation Metrics

In the above mentioned user–centered experiment, we measured user experience of an information system as user perception of ease of learning to use the system, ease of use of the system, understanding of the system, and usefulness of the system. These metrics were measured using questionnaires, including post-system and exit questionnaires, all on scales of 1 (low) to 7 (high). The post-system questionnaire elicits user opinions on the system, and

the exit questionnaire gets feedback from users about their experience of the experiment. These metrics helped researchers better understand how users feel about the system, and if they gained satisfactory experience of using the system. Results from the Wilcoxon signed-rank test showed that users found the integrated system to be significantly easier to use, $Z = -2.264$, $p = 0.024$, and significantly more useful, $Z = -2.522$, $p = 0.012$, than the baseline system [11].

It is worthy of mentioning that the above-mentioned metrics were customized in different experiments to better fit the need of system testing. For example, in evaluating the CiteSpace system, a domain knowledge visualization system [4], it was necessary to measure user perception of the various types of knowledge domain visualization graph to better understand if users can understand and use the system well. For this purpose, we added more personalized questions in the experimental questionnaires to get enriched user data for future analysis [12]. This demonstrates that customization or personalization is critical when switching from textual system to knowledge visualization system.

The advantage of using consistent metrics across evaluations of different systems is that researchers can easily compare the findings, and makes it possible to generalize the results. However the disadvantage is that we may fail to collect potentially important data when testing a new type of system. How to design consistent metrics that can be appropriately used in evaluation of various interactive information systems is a critical issue to consider for the UX researchers.

5 Conclusion

This paper described an integrated interactive information system with particular focus on task design and user experience evaluation metrics. Various task types were considered in the task design to better reflect user information needs in using such an interactive information system. The user experience evaluation metrics have been customized for different systems and we are looking to explore how to use the metrics in evaluating mobile applications in the near future. The task design approach and user experience evaluation metrics can be applied to the process of designing and evaluating interactive information retrieval systems. We believe it is critical to know how UX researchers can agree on consistent formats for evaluation metrics, and how such an agreement can better assist researchers in system design and evaluation.

References

[1] Attfield, S., Kazai, G., Lalmas, M., Piwowarski, B. Towards a science of user engagement (Position Paper). WSDM'11, February 9–12, 2011, Hong Kong, China.

[2] Borlund, P. Experimental components for the evaluation of interaction information retrieval systems. Journal of Documentation; 56 (1), 2000, p.71-90.

[3] Card, S., Moran, T. & Newell, A. The Psychology of Human-Computer Interaction. Hillsdale, New Jersey: Erlbaum, 1983.

[4] Chen, S.Y., & Ford, N. Modelling user navigation behaviours in a hypermedia-based learning system: an individual differences approach. Knowledge Organization, 25(3), 1998, 67-78.

[5] Montabert, C., McCrickard, D. S., Winchester, W. W., Pérez-Quiñones, M. A. An integrative approach to requirements analysis: How task models support requirements reuse in a user-centric design framework. Journal of Interacting with Computers, 21, 2009, 304–315.

[6] European software process improvement training Initiative (ESPITI) project: European user survey analysis. Technical Report ESI-1996-TR95104, European Software Institute.

[7] Rogers, Y., Sharp, H., Preece, J. Interaction Design: Beyond Human - Computer Interaction Paperback – June 7, 2011. Wiley; 3rd edition.

[8] Taylor, F., 1991. Scientific Management. Harper & Row, New York.

[9] The Standish Group, 2014. The CHAOS report. <http://www.projectsmart.co.uk/docs/chaos-report.pdf> (retrieved 06.20.15).

[10] Yuan, X.-J., & Belkin, N.J. Investigating information retrieval support techniques for different information-seeking strategies. Journal of the American Society for Information Science and Technology, 61(8), 2010, 1543–1563.

[11] Yuan, X.-J., & Belkin, N.J. Evaluating an integrated system supporting multiple information-seeking strategies. Journal of the American Society for Information Science and Technology, 61 (10), 2010, 1987–2010.

[12] Yuan, X.-J. Investigating the effect of visualization on user performance of information systems. Design, User Experience, and Usability. Web, Mobile, and Product Design: Lecture Notes in Computer Science. Volume 8015, 2013, pp 591-600.