ONESTART BOOKMARKS: A Comparison of Heuristic Evaluation and Usability Testing

Shruti Bhandari
sbhandar@indiana.edu

Ronald Simamora
rsimamora@indiana.edu

Nina Onesti
nonesti@indiana.edu

Andrew Trus
atrus@indiana.edu

INDIANA UNIVERSITY SCHOOL OF INFORMATICS
Human Computer Interaction Design Graduate Program
1900 East Tenth Street
Bloomington, Indiana 47406
USA
812.856-1802

ABSTRACT

The goal of this paper is to compare and contrast heuristic evaluation and usability testing as methods to document the usability or user-friendliness of a system or application. By using a specific application (OneStart Bookmarks channel), our group as heuristic evaluators, and a number of usability participants, we hope to create a overall view of how heuristic evaluation and usability testing limited and expanded our study.

KEYWORDS

Heuristic evaluation, usability test, comparison, OneStart, usability problem, comparison

TABLE OF CONTENTS

OneStart Bookmarks: A Comparison of Heuristic Evaluation and Usability Testing
Shruti Bhandari, Nina Onesti, Ronald Simamora, Andrew Trus
Indiana University

1

INTRODUCTION

Usability testing and heuristic evaluation both prove to be essential to the proper testing and evaluation of any application or system that involves a high level of user interaction. Our team was interested in finding which of the two methods would prove to be the most beneficial to understanding usability, as well as to identify what value each type of testing can bring to the field of usability. We conducted our testing and compared our results and overall feelings about each type of testing, and ended with our final analysis.

APPLICATION TESTED

Our team tested the Bookmarks channel within the OneStart portal. OneStart is a portal where an Indiana University student can access a number of different campus resources that include but are not limited to: their updated bursar information, class registration, and financial aid information. All of these services and resources are accessible through what OneStart refers to as "channels". These channels are located within "tabs" that organize the channels according to category.

In addition to these more traditional channels, there are customizable channels and tabs that can be added depending on the user's needs and wants. One such customizable channel is the OneStart Bookmarks. Users are able to add bookmarks from their personal web browsers so that they may access them from any computer terminal via OneStart. Our team was testing the overall usability on the Bookmarks channel as well as how the user interacted with OneStart in general.

METHODS

In order to test the OneStart Bookmarks feature, we first conducted a formal usability test (UT) with three different Indiana University students. Once we had our results and analysis of the usability testing, our team proceeded with individual heuristic evaluations (HE). Our team was charged with completing the same tasks as our testers, and thus ended with data that we could compare and contrast between each type of testing.

The UT occurred during one day at the Indiana University Herman B. Wells Main Library. There, our team recruited testing subjects on a participatory basis, to work through several tasks that we assigned to them. The basic tasks were:

- Log into OneStart
- Locate the Bookmarks channel
- Import bookmarks from your favorite browser into the channel

OneStart Bookmarks: A Comparison of Heuristic Evaluation and Usability Testing
Shruti Bhandari, Nina Onesti, Ronald Simamora, Andrew Trus
Indiana University

2

- Modify the properties of the bookmarks (title, etc…)
- Organize the bookmarks (make folders, move, delete, etc…)

Each of our team members conducted an individual HE. We all followed the same task list and compiled our results to finalize our comparison between the testing methods.

RESULTS

Problems identified with Usability Testing

1) Importing bookmarks: Importing bookmarks proved to be extremely difficult. The users struggled to figure out the proper procedure for importing bookmarks. (Major Problem)
2) Help not very useful: The help menus do outline step by step procedures, but are difficult to follow and are not written as clearly as they should be. (Major Problem)
3) Difference between channel and tabs: No distinction is made between a channel and a tab. A user figures out the difference on their own after they gain experience using the system. (Minor Problem)
4) Locating Bookmarks: Users are unable to locate "My Bookmarks" channel easily. "My Bookmarks" channel can be added to a new tab and is present in one of the already existing tabs. Still, a user has difficulty finding the channel. (Major Problem)

Problems identified in Heuristic Evaluation

1) Navigation in HELP documentation: User cannot navigate a page back in HELP documentation. (Major Problem)
2) Navigation within the channel: The Bookmarks are organized in a tree structure of folders and bookmarks. If a user navigates in a folder, they cannot go up a level to other folders in the Bookmarks channel. (Major Problem)
3) A user cannot add a bookmark in a folder. A new link has to be added to the main page of the channel and then has to be manually repositioned in the folder. (Minor Problem)
4) HELP for editing, adding, deleting, and moving bookmarks are unnecessarily lengthy. (Minor Problem)
5) Inconsistent symbols and icons: The cursor turns into a clickable hand if hovered over a web link. However, the cursor turns into editable when it is hovered over folders. This inconsistency may cause confusion to a novice user. (Minor Problem)

COMPARISON

Below is a chart which quantitavly compares our results from our UT and HE, classified as major and minor problems.

OneStart Bookmarks: A Comparison of Heuristic Evaluation and Usability Testing
Shruti Bhandari, Nina Onesti, Ronald Simamora, Andrew Trus
Indiana University

3

	Major Problems	Minor Problems
Heuristic Evaluation	2	3
Usability Testing	3	1

From both types of testing, our team was able to identify a few areas that overlapped.   Our usability participants and team members found the OneStart help menu to be extremely unhelpful. On the other hand, our team found that specific symbols and icons were not consistent, something that the UT participants either overlooked, or were not aware of.

Usability testing (UT) gave our team a clear idea about what the user feels and how the system would be received. We asked that the testers use verbal protocol (spoken cues and thoughts) to indicate their thoughts and feelings about the interface, ease of use, and over all design on OneStart Bookmarks. We were also able to record their facial expressions and mouse clicks to help further analyze where potential "user problems" may arise.

UT is more focused on the system's functionality and brings out the problems that may make the system unusable or inefficient. HE tests the interface and tries to find out what "looks out of place" in the system. It is also concerned with the visual layout and the relative user ease to adapt to the system.

The heuristic evaluation (HE) was conducted on a personal basis according to each team member. All of the HE data was documented and collected by each individual group member, as he/she encountered them. Our goal was to use HE to identify problems within the system. Hence, we expected HE to bring out the finer problems that the user might not even face.   Additionally, three of the four group members have had prior experience with OneStart as a University system, which was helpful in identifying problems that advanced OneStart users might come upon during interaction with the Bookmarks Channel.

Testers can devote more time on HE and test each feature individually. While UT is more time intensive and labor intensive, it promotes test of the main functionalities/features of the system. Hence, HE tests every aspect of the system, and UT tests the most used/accessed features of the system.

The test setup matters in UT while the HE is undertaken in optimum conditions. This may affect the results of UT. In fact, we faced this problem in our UT. One of our users was very frustrated with the computer setup. These things though do not reflect the quality of the system but the

OneStart Bookmarks: A Comparison of Heuristic Evaluation and Usability Testing
Shruti Bhandari, Nina Onesti, Ronald Simamora, Andrew Trus
Indiana University

4

frustration is carried out forward. A user may have problems in logging in the system which is plausible scenario, the repeated password failure can lead to frustration can be carried out to the next step of the system.

According to Law and Hvannberg, "HE is easy to apply and quick to produce results, some of which may be irrelevant from the user's point of view, whereas UT is slow and laborious to apply but the results are accurate." (72). Our team drew similar comparisons of both HE and UT since several of us had previous experience therefore giving us more prior knowledge of the application.

HE and UT helped us to find different types of problem. Several problems we found out in HE were not even pointed out by the users of the system. Though, these problems might be labeled as "false alarms", our team would not like to classify them as such since our user sample was small (3 users).   Law and Hvannberg also came to similar conclusions stating, "that UT yields more **accurate** and **objective** results than HE. HE is likely to misidentify UPs (i.e. 'false alarms')" (72).

Law and Hvannberg also state, "that HE is more likely to **fail to identify positive features** of the product than is UT" (72). Our team had a comparable experience as we found ourselves looking for more things that would be "wrong" or "flawed" within the system, rather than just having an experience with the application and finding the problems that way. Thus, it became difficult for us to be positive about our experience OneStart bookmarks.

REFLECTION

After all of the UT and HE was finished, our team came together to discuss our findings and overall experience using each method to evaluate the usability of OneStart Bookmarks. Initially our team was extremely opposed to doing a HE on the very same system that we had just conducted UT on, but found that the HE proved to be helpful. We were frustrated because we felt that we had conducted an extremely informal HE prior to the UT in order to familiarize ourselves with OneStart Bookmarks in order to be helpful to our participants during the UT's.

We found that it was easy to find the flaws during our HE as we had already witnessed our usability testing sessions and seen what our participants had come across. We also found that we looked for other things that were flawed and were almost "digging" to find other things that were usability problems. In reflection we feel that the HE would have provided to be more helpful if it were completed prior to the UT.

OneStart Bookmarks: A Comparison of Heuristic Evaluation and Usability Testing
Shruti Bhandari, Nina Onesti, Ronald Simamora, Andrew Trus
Indiana University

5

We feel this way since we would not have the user's experience to weigh upon our actions and decisions. This would also allow an "expert" or someone with more experience to test the application's usability first, identifying any major problems that the user may encounter, and provide a solution. We also feel that it would have proved to be more effective, that is if the HE occurs after the UT, to have another set of "experts" conduct the HE as to have unbiased experience.

CONCLUSION

In conclusion, our team conducted both usability tests as well as heuristic evaluations on the OneStart Bookmarks channel. We found that our results from each testing session were similar in a more basic, cosmetic sense. On the other hand we found that through heuristic evaluation, our team was much more particular and attentive to finding problems, more so than if we were to not have already conducted our usability testing sessions.

REFERENCES

(1) Ellerston, Peter. The IMRAD Research Paper Format. January 16, 2006. September 24, 2006. http://www.uta.fi/FAST/FIN/RESEARCH/imrad.html

(2) Neilson, Jakob. Ten Usability Heuristics. September 24, 2006. http://www.useit.com/papers/heuristic/heuristic_list.html

(3) Law, Lai-Chong; Hvannberg, Ebba Thora. Complementarity and convergence of heuristic evaluation and usability test: a case study of universal brokerage platform. September 24, 2006. http://doi.acm.org/10.1145/572020.572030

OneStart Bookmarks: A Comparison of Heuristic Evaluation and Usability Testing
Shruti Bhandari, Nina Onesti, Ronald Simamora, Andrew Trus
Indiana University

6