

Project 3: Usability Testing vs. Heuristic Evaluation

Team Members

Jenny Allen

Eric Drewski

Ashley Engelhardt

Junyoup Kim

Date: October 4, 2007

Introduction

The team analyzed the functionality of the OneStart Classifieds system at Indiana University-Bloomington with a usability test and a heuristic evaluation. The usability test involved three subjects. Each subject was given three scenarios that involved completing specific tasks within the OneStart Classifieds. The team then conducted a heuristic evaluation of the OneStart Classifieds. It should be noted that because the heuristic evaluation test was conducted after the usability test, the results of the heuristic evaluation were influenced by the results of the usability test. The following document provides information regarding the problems discovered through the usability test and the heuristic evaluation method. It also documents the differences between the two methods. An analysis of the two techniques reveals that neither technique is superior to the other, but that the two techniques complement each other.

Usability Testing

In order to evaluate the OneStart Classifieds, a usability test was conducted with three participants. After the data was collected and analyzed, the advantages and disadvantages of usability testing became apparent. Table 1 provides a brief summary of the advantages and disadvantages of usability testing that emerged during our study of the OneStart Classifieds.

Table 1. Summary of advantages and disadvantages of usability testing

Advantages	Disadvantages
1. Usability test can easily identify critical problems	1. User testing cannot find certain minor problems
2. Identifies functionality that is missing from program	2. Testing usually performed later in the development process
3. Test provides quantitative data such as number of clicks and number of errors	3. Difficult to find participants
4. User feedback provides qualitative data	4. High cost method in terms of time to develop test, time to find participants, etc.
5. Finds problems that effect real users of the system	5. Finding certain errors is limited based on the type of prototype used

Advantages of Usability Testing

The primary advantage of the usability test is that the test can easily identify critical problems within a program [1]. A critical problem becomes immediately apparent because it stops the user from completing a task. The critical problem that was discovered through the usability test of the classifieds was that the search tool requires the user to enter a string of characters that exactly matches at least part of the string of characters that makes up the title or description listed within the ad. One participant in the study was unable to complete a task because the search tool did not perform a keyword search, which is what the participant expected.

Not only does the usability test identify problems with existing functionality, it is also able to immediately identify functionality that is missing from the program. Two of the three participants in the usability study initially used the OneStart search tool and then, when faced with a large number of results, further narrowed their search using the search functionality embedded within the browser to find specific items on the page.

Another advantage is that usability tests provide quantitative data [3]. The time to complete the task, the number of clicks required to complete the task as well as the number of errors found were analyzed. The OneStart usability study revealed a correlation between the time it took the participant to complete a task and correctness of the task. The tasks that took the longest to complete were the tasks that the users were not able to complete correctly.

Likewise, qualitative data is also collected during usability testing. One form of qualitative data obtained during testing was a video of the participants' facial expressions. Two of the three participants in the usability study had looks of confusion whenever they tried to use the search tool and could not find the correct item. Likewise, one participant appeared especially frustrated when the search tool did not work even after multiple keyword combinations. The debriefing questions and the post test questionnaire also provided qualitative data. The debriefing questions and the post test questionnaire allowed the participant to provide feedback. It was interesting to discover that each participant in the study was generally satisfied with the OneStart Classifieds. It was also interesting to discover that despite the difficulty the participants experienced with the search tool, the majority identified it as the most valuable feature of the OneStart Classifieds. The usability test allows participants to make suggestions as to how to further enhance the program. The majority of the participants expressed a desire to have subcategories available to further refine a search. Each participant also expressed that the ability to view and upload pictures should be provided.

Finally, because the usability test is conducted on volunteer participants, problems that effect real users are discovered [1]. The inability of the search tool to search by keywords limited or stopped two of the three participants in the usability study from effectively finding results.

Disadvantages of the Usability Testing

Although usability testing has many advantages, it has a number of disadvantages that must be addressed as well. First, the usability test is primarily effective at finding major or critical problems within a program. The usability test used to analyze the OneStart Classifieds was only able to identify 6 problems. Only three of the six problems were identified as minor problems, one of which was merely a suggested improvement to the system.

Another disadvantage of the usability test is that it is typically performed late in the development cycle of a program. The usability test for the OneStart Classifieds was performed when the system was already live. It is unlikely that developers will make significant changes to the interface at during the late stage development unless the problem is critical, meaning that minor problems identified during testing will likely be ignored.

Usability testing also has the disadvantage of requiring users. Finding users to participate in the study that fit the specified profile is a difficult task. For our study of the classifieds, the process took several days. Coordinating schedules for the participants and the evaluators was also difficult.

Furthermore, the "cost" of usability testing can be a problem. The development of the usability test and the coordination and execution of the test took several hours. In addition, the usability study required specific equipment to record the user's reactions to the program.

Another disadvantage that we noted was that the usability test is limited in the sense that not all functionality within a program can be analyzed. The problems that were discovered only reflect what was tested in the three scenarios. For example, since we did not ask the user to utilize the “back” button in any scenario, there was no way for us to detect a problem with that functionality.

The final disadvantage of usability testing identified during the OneStart Classifieds study dealt with the format of the prototype. In our case, this was not an issue since the prototype was a live, working version. However, if the prototype had been on paper, we would have had a limited ability to identify navigation problems because going from one “screen” to another would simply involve swapping one paper prototype for another.

Heuristic Evaluation

Heuristic evaluation represents an alternative to usability testing. This method requires the evaluators to examine a system themselves, using ten established heuristics as a guide. In order to examine the validity of this method, a heuristic evaluation was conducted on the OneStart Classifieds after usability testing was completed. As before, certain advantages and disadvantages of heuristic evaluation emerged as the method was applied. These advantages and disadvantages are summarized in Table 2.

Table 2. Summary of advantages and disadvantages of heuristic evaluation

Advantages	Disadvantages
<ol style="list-style-type: none"> 1. Based on 10 established heuristics 2. Finds a high proportion of major problems 3. Identifies many minor problems that would not be identified by a user test participant 4. Method can be applied at any time during the development cycle 6. No users to recruit 	<ol style="list-style-type: none"> 1. Identified problems tend to be dominated by minor problems 2. Difficult to find elements missing from an interface 3. Must have expert evaluators 4. Must have 3-5 evaluators [1]

Advantages of Heuristic Evaluation

One advantage of heuristic evaluation is that it is based on ten distinct criteria, or heuristics. Unlike the usability test, this method requires that the evaluator view and evaluate all different aspects of a program, considering each heuristic. As such, heuristic evaluation results in a comprehensive view of the entire system. Our heuristic evaluation of the OneStart classifieds identified at least one problem for each heuristic category.

Heuristic evaluation also has the benefit of being able to find many problems [1]. In our case, the heuristic evaluation of the classifieds by four evaluators resulted in a list of more than 35 problems. Evaluators can take their time and thoroughly examine the system. The problems can then be ranked based on severity to prioritize which problems need to be fixed first.

Moreover, heuristic evaluation is good at identifying a high proportion of major problems [2]. For example, three problems with the search engine of the OneStart Classifieds were identified, indicating that the search tool should be the top priority for redesign.

In addition to identifying most major problems, heuristic evaluation is effective at finding minor problems. In general, twice as many minor problems are identified than major problems [2]. Heuristic evaluation is particularly good at finding minor problems that would not be described by participants in a usability test. For instance, heuristic evaluation of the classifieds revealed that if a user performs a search that results in no matches, the error message reads “There are no ads to display.” A more appropriate message would be “There are no ads matching your search.” Although this is only a minor problem and was not identified during usability testing, making the slight modification would improve the system.

Another advantage of heuristic evaluation is that the method can be applied at any time during the development cycle. Likewise, there are no users to recruit. Therefore, the method can be applied multiple times to the same system with little associated cost.

Disadvantages of Heuristic Evaluation

Although heuristic evaluation finds a variety of problems, the majority of these problems are minor. In our study of the classifieds, many problems were cosmetic rather than functional, such as having active links that did not actually go anywhere. These problems should ideally be fixed but if there are more pressing problems, it may not be worthwhile to focus on such minor issues.

An additional disadvantage of this method is that it is not good for identifying missing elements in a system [2]. For example, had our usability testing not revealed the poor search tool, it is likely that our heuristic evaluation would not have identified the problems with the search tool. Since we knew it had been a problem for the users, we paid particular attention to the search tool during our own evaluation.

The final disadvantage that emerged from our heuristic evaluation of the classifieds deals with the evaluators. Expert evaluators are needed to truly capitalize on the method. Furthermore, evaluators are ideally needed to perform a comprehensive heuristic evaluation [1]. In our study, four evaluators were used and each found unique usability problems in the system. In some cases, it may be difficult to find enough experts to make the evaluation feasible.

Conclusion

Neither usability testing and nor heuristic evaluation is a superior method for evaluating the usability of a system. Both methods have their advantages and disadvantages. Yet, when used in conjunction, the methods can provide a comprehensive evaluation of a system. The primary ways in which the usability test and the heuristic evaluation complement each other are the following:

- Heuristic evaluation finds lots of different problems in many areas of a system. Usability testing tends to find a smaller number of problems in a specific area of the system.
- Heuristic evaluation tends to find a variety of different kinds of problems throughout the system. Usability testing tends to find problems that have a significant impact on real users.
- Heuristic methods are typically unable to find missing elements in a program. Usability testing is very effective at finding missing elements in a program.

In order to truly capitalize on the strengths of each evaluation method, usability testing and heuristic evaluation should be strategically used. For example, heuristic evaluation would be an effective method to identify potential problems in a system since the entire system is examined. To verify the results of the heuristic evaluation and to gain further insight into particular problems, a usability test could then be conducted with scenarios that concentrate on the areas of interest. Conversely, a usability test could be conducted first to reveal areas that need particular attention and then a heuristic evaluation could be concentrated on that area. Regardless of the order in which they are performed, usability testing and heuristic evaluation should be utilized together to create the best system possible.

References

- [1] Desurvire, Heather and Jeffries, Robin. Usability Testing vs. Heuristic Evaluation: Was there a contest? *SIGCHI Bulletin* 24, 4 (1992) 39-41.
- [2] Nielsen, Jakob. Finding usability problems through heuristic evaluation. *Proc. CHI2000*, ACM Press (2000) 373-380.
- [3] Hvannberg, Ebba Thora and Law, Lai-Chong. Complementarity and Convergence of Heuristic Evaluation and Usability Test: A Case Study of UNIVERSAL Brokerage Platform. *NordiCHI* (2002) 71-79.