What is the relationship between task-based and open-ended usability testing, in terms of measuring satisfaction?

by

Srinivas Reddy Boddu

LIU-IDALITH-EX-A--13/008--SE
Final Thesis

What is the relationship between task-based and open ended usability testing, in terms of measuring satisfaction?

by

Srinivas Reddy Boddu

LIU-IDA/LITH-EX-A--13/008--SE

Supervisor: Johan Åberg
Examinor: Stefan Holmlid
Upphovsrätt

Detta dokument hålls tillgängligt på Internet – eller dess framtidiga ersättare – från publiceringsdatum under förutsättning att inga extraordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnd som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida http://www.ep.liu.se/

Copyright

The publishers will keep this document online on the Internet – or its possible replacement – from the date of publication barring exceptional circumstances.

The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/hers own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its www home page: http://www.ep.liu.se/

© Srinivas Reddy Boddu.
Abstract:
Usability is one of the most important aspects of Information Technology. Usability plays a vital role in this industry, where organizations strive to ensure utmost satisfaction of their end-users in regard to the experience of using their product. The systems may be a website or a software application. To measure user satisfaction, the method of usability testing can be performed. Performing usability testing gives a clear picture of difficulties that would be faced by potential target users. There are different types of usability testing such as Task-based usability testing, open ended usability testing, remote usability testing etc. The important point here is about deciding upon the most appropriate type of testing technique to get the accurate user satisfaction level.

This study is mainly focused to answer the following research question: What is the relationship between the task-based and open ended usability testing, in terms of measuring satisfaction? System Usability Scale (SUS) has been used to measure the satisfaction of the users in this study. For this we used two websites performing task-based usability testing and open ended usability testing respectively.

This study had involved twenty eight different participants. Participants are divided into two groups, one group to perform open ended usability testing and another for task-based usability testing for both the websites. This study has produced following results; Open-ended testing tended to produce higher SUS-ratings for the tested system. The results in this study showed that users performing open-ended usability testing gave positive responses for both the websites in terms of user satisfaction. Open-ended usability testing is an exploratory testing, where the testing is based on different aspects such as user interface of the system, design etc. Task-based usability testing is goal based where users have to complete the given task without fail. This method drew lower scores when compared to open-ended usability testing for the tested systems from the attained results. Nevertheless that task-based testing attained lower SUS scores, it is fairly straightforward than open-ended testing to measure efficiency and effectiveness. The above results have been discussed in detail. This study has finally concluded that to measure the usability of a system it is recommended to practice both the open-ended and task-based usability testing techniques.

Key Words:
Usability, Usability testing, Task-based usability testing, Open ended usability testing, SUS, Time on task, Task success, effectiveness.
Acknowledgements

This study was carried out in the Department of Computer and Information Science, Linköping University under the Human centered systems division.

I would like to extend my sincere gratitude to my supervisor Dr. Johan Åberg, researcher at IxS group, Department of Computer and Information Science, Linköping University for his precious time and invaluable suggestions throughout this study. His guidance, comments and support has been extremely helpful for gaining very useful knowledge in the subject.

Also I would like to thank my friends: Karteek Maddela, Koushal Kumar Gopi and Mohan Sure for their constant support throughout my stay in Sweden.

I’m very grateful to my family back home in India, who were always there to attend my needs and I’m happy for their love, best wishes and blessings.
# Table of Content

Abstract: .......................................................................................................................... IV

Key Words: ...................................................................................................................... IV

Acknowledgements ......................................................................................................... VI

Table of Content ........................................................................................................... VIII

1. Introduction ................................................................................................................. 1
   1.1. Background ............................................................................................................ 1
   1.2. Aim ....................................................................................................................... 2
   1.3. Disposition ............................................................................................................ 2

2. Theory .......................................................................................................................... 4
   2.1 Usability ................................................................................................................ 4
       2.1.1 Characteristics of Usability ........................................................................ 4
       2.1.2 Usability Evaluation Methods .................................................................... 4
   2.2 What is user experience? ..................................................................................... 5
   2.3 What is usability testing? .................................................................................... 6
       2.3.1 Different methods of Usability Testing ....................................................... 7
       2.3.2 Data Representation Techniques .................................................................. 11
       2.3.3 Benefits of Usability Testing ....................................................................... 12
       2.3.4 Usability Testing Issues .............................................................................. 12
   2.4 System Usability Scale (SUS) ............................................................................. 12

3. Method .......................................................................................................................... 15
   3.1 Type of Study ........................................................................................................ 15
   3.2 System Selection .................................................................................................. 15
   3.3 Participant Selection ............................................................................................ 17
       3.3.1 Participants for Pilot test ............................................................................ 17
       3.3.2 Participants for Real test ............................................................................. 17
   3.4 Test Duration ....................................................................................................... 18
   3.5 Task Design ......................................................................................................... 18
   3.6 Test Procedure .................................................................................................... 19
       3.6.1 Pilot Test Procedure .................................................................................. 19
       3.6.2 Main Test Procedure ............................................................................... 20
   3.7 Dependent Variables ......................................................................................... 21
4. Results: ..................................................................................................................22
5. Discussion..............................................................................................................25
  5.1. Metric Selection ...............................................................................................25
  5.2. SUS Score .........................................................................................................25
  5.3 Efficiency and effectiveness .............................................................................26
6. Conclusion: ............................................................................................................28
7. References ............................................................................................................29
Glossary: ..................................................................................................................31
Appendix: ..................................................................................................................32
1. Introduction

1.1. Background

Developed systems need to be evaluated based upon their quality of performance. Evaluation of the systems is one of the most important concept needs to be implemented on the systems in the initial stages of development or after the system is released into the market. Implementing this concept during the development phase would avoid loss in time, improving the developed design to compete with others. If it is implemented after the product is released into the market then, it will be used to know if the system is performing as required or not. For evaluating the products usability concept is majorly being used. There are many more approaches other than usability, but using usability concept would yield accurate results directly from the users.

There are different approaches in usability evaluation concept; usability testing is one of the major approaches used in evaluating fully developed system. Nowadays IT experts prefer usability testing to other methods for testing the user experience and satisfaction. There are many types of usability testing methods such as comparative usability testing, open-ended usability testing, task-based usability testing, remote usability testing etc. Any type of usability testing aims to solve the usability issues by measuring the user experience level. The detailed explanation about usability testing is given in section 2.3.

In [22], a survey of research studies focused on user experience is presented. One important finding from this study is that studies aiming at evaluating user experience mostly focus on open use situations, i.e. just putting the user in front of the system to be tested and observe what the user does. Given that usability testing as an evaluation method is focused on concrete tasks, this begs the question of whether usability testing with open use situations gives different results in any significant way. Is it reliable to test with open use situations? Hence the present study compares usability testing with concrete tasks with open use situations, in terms of user satisfaction.

For performing this study there are different types of scales to measure the user satisfaction of a system. Some of them are System Usability Scale (SUS), Questionnaire for User Interface Satisfaction (QUIS), Computer System Usability Questionnaire (CSUQ), and Post-Study System Usability Questionnaire etc. But, for this study SUS has been considered over other usability scales. This is because when we consider PSSUQ, CSUQ these are the questionnaires with all positive items or questions using which is there is the chance of acquiescent bias [22]. Due to this reason there was a need to introduce some negative worded items for balancing this bias and get accurate responses from the user. SUS was the better scale to be used to overcome this problem and this was the appropriate usability
scale for this study. SUS was developed by John Brooke in 1986. This consists of ten statements with half of the statements with positive words and other half with negative words. This scale is a 5 point scale of agreement with the score range from 0 to 100 [23]. The detailed description for SUS is explained in the theory section.

1.2. **Aim**
The main aim of this study is to answer the following research question.

- What is the relationship between task-based and open ended usability testing, in terms of measuring satisfaction?

As there are many types of usability testing methods, most of the usability experts will be in a dilemma of using the type of the testing method. This is because; one testing method would yield better results in one area and other in different area. For example, if we consider task-based usability testing measuring time on task, task success rate can be measured accurately. This is not possible with open-ended usability testing method. So for comparing these two testing methods, System usability scale has been used for measuring the user satisfaction.

SUS is one of the most preferred scales using which usability experts can easily draw the conclusion about the type usability testing methods to be used to test a system with the users. So this study would answer about what is the preferred usability testing method task-based or open-ended and also its relation in terms of measuring satisfaction.

1.3. **Disposition**

*Chapter 1: Introduction*

This part mainly deals with the introduction of the research that has been carried out in this study and the research question.

*Chapter 2: Theory*

This chapter mainly includes the entire literature review that has been used in this entire study taken from different articles, books and websites.

*Chapter 3: Methodology*

This chapter explains about the type of study and how this study has been carried out.

*Chapter 4: Results*

This part presents all the results attained in this study which are presented graphically.
Chapter 5: Discussion

This chapter analyses all attained results shown in the results part and also shows what conclusion the analysis leads to.

Chapter 6: Conclusion

This part explains about what conclusions have been drawn from this study.
2. Theory

2.1 Usability
Usability is the ease of use, which generally refers to quality of being able to provide efficient service to the product. The definition of usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use (ISO 9241-11).

There are different characteristics and evaluation methods of usability.

2.1.1 Characteristics of Usability
The characteristics of usability are denoted by 5 ‘E’s [2].

- **Effective**: This is the integrity and preciseness with which the users can fulfill their desired goals. It can be decided by considering whether user’s goals and work were met successfully.
- **Efficient**: Efficiency can be explained as the speed at which the users can finish the task of a product that they are using. Metrics that are included for efficiency can be number of clicks or key strokes required or the total time on task.
- **Engaging**: If the interface of the product is pleasant and satisfying to use by the user, then it is said to be engaging. This includes the style of visual presentation, the number, functions and types of graphic images or colors mainly on websites.
- **Error tolerant**: As a developer the main goal will be to minimize the errors, but as we are humans, errors will occur at any point of time. Error tolerant system is developed to avoid the errors caused by user’s interaction and make sure that bugs are recovered for the user.
- **Easy to learn**: If the interface of the website is simple for the user to learn, it would avoid them from cautious efforts to get familiarize with that product.

In this study usability is mainly focused on the websites rather than the other systems like software products. Usability testing has to be performed to measure if the website is user friendly.

2.1.2 Usability Evaluation Methods
There are mainly three types of usability evaluation methods. The usability of the product is mainly evaluated based on time, cost and way of approach constraints.

2.1.2.1 Testing:
In this approach, representative users work on the system using the typical tasks designed by the evaluators. From the results obtained by testing will help the developers to improve the interactivity of the system with the users [15].

There are different types of usability testing. They are mentioned below:

1. Question-Asking Protocol
2. Remote Testing
3. Think Aloud Protocol
4. Retrospective Testing

2.1.2.2 Inspection:
In usability inspection method, along with usability professionals the developers, users and also system related specialists test the usability related things of user interface [16].

Inspection approaches are as follows:

1. Cognitive walkthroughs
2. Feature Inspection
3. Heuristic evaluation
4. Perspective-based Inspection
5. Pluralistic Walkthrough

2.1.2.3 Inquiry:
Evaluation of the system is done by taking the feedback from the user in terms of likes, dislikes, needs and understanding of the system. This can be done by speaking to the users, observing them using system in real work or from their verbal opinions about their experiences [16].

Different Inquiry methods are:

1. Field observation
2. Focus groups
3. Interviews
4. Questionnaires
5. Proactive Field study
6. Logging actual use

2.2 What is user experience?
User experience (UX) is the way a person feels about a product, system or service. It has emerged as an umbrella phrase for recent ways of understanding and studying the quality in use of interactive products. User experience is defined as all aspects of the user’s experience when interacting with the product, service, environment or facility (ISO CD 9241-210) [6]. When we compare the terms usability and user experience, usability is considered as the user’s ability to finish the task successfully, whereas in the case of user experience it is different. It takes the extended view, looking at the participant’s complete interaction with the product, as well as the emotions, attitude that result from interaction. When we want to mention about the measuring usability, we have to measure user experience (UX) [24].
2.3 What is usability testing?
Usability testing is the technique for ensuring that intended users of system can carry out the intended task efficiently, effectively and satisfactorily. It is a technique used in user-centered interaction design to examine the product by testing on users of the product. This technique is performed by the specific department of the product, who are experts in evaluating the product by testing the users. Here, usability testing is performed to measure the satisfaction level of the users. Results that are achieved by performing this will be represented in the form of different usability metrics. The different usability performance metrics are Task success, Time-on-Task, Errors, Efficiency and Learnability.

- **Task Success:** This metric is used to measure how efficiently users are able to complete the given set of tasks.
- **Time-on-Task:** It is used to calculate the time taken by the user to complete a task.
- **Errors:** Errors are occurred during the task is being performed. They are used to indicate the complex and difficult parts of a product and its interface.
- **Efficiency:** This is used to measure the amount of effort a user spends to complete the task.
- **Learnability:** This will measure to what extent the user is getting adapted to the changes in the product.

Usability testing has to be performed during the initial stages of the product at regular intervals of time. During the development phase, developers will get to know to what extent users are getting satisfied with the product. Also, they will have clarity about what is to be improvised in it.

The classic process of usability testing suggested by Jeff Rubin and Dana Chisnell in the Handbook of usability Testing is to perform formal usability test as follows [9]:

1. **Develop a test plan:** In order to develop a test plan, one has to work with team and accept the test objective, questions to be used to conduct the test and characters of participants who have to test the design of the system. This also includes methods and measures that evaluators use to learn regarding the research questions.
2. **Choose testing environment:** In usability testing choosing test environment is crucial. System analyzer has to decide if the test is to be conducted in lab or not because user will definitely have an impact of the environment. For example if the test is conducted in open areas might disturb the user. In the other case if the test is being conducted in a closed lab this might give accurate results. Choosing of test environment should be decided within the team.
3. **Finding and selecting participants:** When participants are being selected by the usability professionals for testing a system, they have to know the behavior of the users. This makes the life of the analyzers easy than trying to select market segmentation or demographics. If the testing is being conducted on hotel reservation process on a website, we want people who do their own bookings.
4. **Prepare test plan:** For conducting usability test of a system on a user, testers have to prepare a test plan initially. A test plan should consist of the research questions that
are being translated to task scenarios that represent realistic goals. A test material should include specific interview questions that testers might want to ask, prompts for follow-up questions, as well as closing, debriefing questions that user have to answer.

5. **Conduct the session**: For every team there will be a moderator who manages the test sessions. He would see the safety and comfort of the participants, manages the team members observing and handles the data collected. Though moderator is the head of a team, all the members in the team have to observe usability sessions. If there are number of individual sessions conducted, each member of the team should watch at-least two sessions.

6. **Debrief with participants and observations**: Once the session is complete, make sure to step back with the participant and ask, “How’d that go? “ Invite the experience observer to pass follow up query to the moderator or to ask questions themselves. At the end thank participant, compensate him/her and say good-bye. Now, the team observing should talk briefly about what they saw and what they heard.

7. **Analyze data and observation**: What you observed: What you your team saw and heard is, what you know at the end of the usability test. When observers have close look at the observation together, the weight of evidence helps observers to examine why particular things have happened. From these achieved results, we can develop theories about the causes of frustrations and problems. Once these theories are developed, usability experts can use their expertise to determine the solutions to fix problems. Then all the changes can be implemented and test theories in later usability tests.

8. **Create findings and recommendation**:
   - **What you get**: If the entire process is followed linearly, this ends up with concrete planning, solid controls, huge data, and rigorous analysis and finally results. In the real world usability tests have to be lighter and faster. Best user experience teams perform usability testing for only few hours every month. In short we can say that as long as it involves real people using the system, its usability testing.
   - **Someone, Something, Someplace**: In usability testing everything that is required is someone who is a user of a system, something to test (a design in any state of completion), and some place where the user and design can meet and can observe. Once the usability experts gets experience of performing user research and usability testing, they will get to the shortcuts and squeeze the process down to very small number of steps that works fine and reduce the time required to analyze the system.

### 2.3.1 Different methods of Usability Testing

There are many methods of usability testing. Here are some of different types of usability testing.
2.3.1.1 Comparative Usability Testing

As the name implies, comparative usability testing is meant for comparing two different systems or products. This technique evaluates the pros and cons of two or more systems or prototypes based on user’s experience. This testing method is also known as competitive testing. It can measure to what extent the product measures up the competition or focus on competitor’s system to give new ideas about new features and functionality to be implemented [11]. Comparative usability testing can also be used at high level, where a product is tested against competitor’s product, or at low level where two web page prototypes are tested to establish which provides best user experience [10].

Here are some of the goals that are being achieved by comparative or competitive usability testing [12]:

- Understand the marketplace: Who are your Competitors? What are they offering? How can we maintain our competitive advantage?
- Build domain knowledge: What content and functionality do our competitors offer that we have not met.
- Identify best practices: What customer needs and preferences are we competing to meet? Beyond functionality, how can we offer a best quality user experience?
- Fine-tune your strategic direction: what are the unique strengths our product in market place? Is there a great match between the user’s experience and branding?

The results that are achieved from competitive analysis could help us to:

- Get new ideas and leverage wisdom to continue where our competitors left off.
- Discover strengths and weakness of products or process.
- Make difficult decisions about strategic direction.

2.3.1.2 Explorative Usability Testing

This is performed in very early stage of product development cycle. It helps to establish the new content and functionality that a new product should consider to improve beyond its competitors. Testing of the product is done with targeted audience. In this audience are given real scenarios to know gaps in the market. It will help the developers to focus on design aspects. Explorative testing is usually performed as a part of user requirements capture exercise. Performing this method will avoid wastage of time, cost and efforts after the product is being released. This is because, as the test is being done during the initial stage of the product, one will get to know the requirements of the users and develop it accordingly [8].

2.3.1.3 Remote Usability Testing

Apart from traditional usability testing, now the focus has turned towards remote usability testing in which usability experts need not to be present near users when they are performing usability test. The diffusion of screen sharing software and remote control applications is slowly changing how usability evaluations are conducted. This is because usability experts can now interact with users who are located geographically far away. Now-a-days internet
and fast broadband services improve the ability to install and test high quality videos and make the life of usability evaluators easy and accurate by observing users. Remote usability testing can be defined as “a technique that exploits user home or office, transforming into a usability laboratory where user observation can be done with screen sharing applications” [9]. Performing remote usability testing is not that different from normal usability testing except some facts which are need to be considered.

- **Participant recruiting:** For recruiting participants to remote usability testing can be done through any of the available channels used for traditional usability testing. Cheapest and fastest way will be using online questionnaires. But before selecting participants to test, it is required to see that participants have fast internet connection and knowledge about screen sharing and remote control applications. This entire process of preparing questionnaires from scratch will take three to four days to complete. So it will be very good practice to use existing questionnaires adapting them to current situation thus reducing work to one day.

- **Environment set up:** Performing remote usability test requires remote control software for both usability experts and participants. While setting up the application, installing, configuring will be easy for analyzers but, the case might be different on user side if they do not have enough knowledge or experience. This can be avoided by sending simple documents to user explaining about configuring the system and how to set up the application for usability session. Other solution will be to launch collaborative conferencing service like WebEx that allows meeting anyone, anywhere in real time over web.

- **Test execution:** Entire test will be conducted based on dialog with participant and is performed over phone or using chat and messenger applications. This can avoid misunderstanding between testers and users in foreign countries. In this user interaction can be viewed in facilitator’s monitor, thanks to the use of screen sharing software developed. In all the cases user interaction has to be recorded for future reference and analysis.

Here are some advantages and limitations of remote usability testing [9]:

**Advantages:**

1. Reduces usability testing costs
2. Makes usability tests more real
3. Permits strategic participant recruiting
4. Increases the parallelism of usability tests
5. Reduces time needed to perform usability evaluation

**Limitations:**

1. Security and performance issues
2. Only limited visual feedback is attained
3. Makes life of analyzer difficult to build trust and relation
2.3.1.4 Task based Usability Testing

Task based usability testing is all about observing users how they interact with the website to browse the required information and complete all tasks. This is one of the major methods of usability measurements available. There are many aspects need to be considered in this type of testing. Initially it is required that a task is to be defined for the user. In other words a participant has to perform something with the product/system, which very much essential for attaining required result. Usually, usability testers gain nothing by simply giving the system to the user and then expect some outcome about its usability. There can be a variety of tasks being designed for users depending on the factors that are to be measured. Let us consider an example of a task installing a software product by the user. It measures that part of software, but it does not calculate its use. If we request the user to use the particular portion of that software, then the required task is being completed and can get accurate result in terms of usability. This type of test is bit narrow, but the results that are achieved will be very rich in quality. Here as mentioned using software is just an example one can use this technique to test other products like websites.

Once all the interested tasks are devised, next step is selecting the participants and this process really goes for all measures of usability. Participants who are being selected for testing the system should resemble the real users of the item in questionnaire. If the selection of participants is not genuine, then results obtained will be faulty. Final part of this study is actual test. Conducting actual test requires many things to be considered, but many of these considerations depend upon what a usability expert is actually interested in finding out. Some of these considerations will be:

- Audio and video recording of the participants, as they complete the task.
- While performing a task, facilitator often should ask the participant to read aloud.
- If they find the participant saying “I don’t understand this” at a particular point of the task, then better understanding can be obtained of where an actual user of the system might get confused.

Now a days to conduct task based usability testing there are many software products available in the market. Most of the software performs some type of recording of the key strokes and navigation, a participant employs during the test. For example, if the user has to perform a specific task of selecting an icon on the screen in order to do something, but if the user can’t find the icon then it will be recorded by special usability software [23].

2.3.1.5 Open-ended Usability Testing

Open-ended usability testing is one of the testing methods in which testing is based on open ended questions. These open ended questions or tasks allow participants to organically explore the system based on a scenario we provide. This is one of the major testing methods with which an on-going narration can be established from the test participant [15 & 28]. In this type of testing user is given a system to use and allot them a specific time period to get to know the system. After the user uses the system, they are asked to answer some open ended
questions in order to know about their overall satisfaction about the system based on the design, User interface or speed at which the system is accessible.

Here is an example scenario of open ended usability testing: “Please take about 10 minutes of time to study the university website, based on the scenario in order to get familiar with the university and what it has to offer to you as a prospective student. After you have studied the website, you will be given a questionnaire to be filled in.”

2.3.2 Data Representation Techniques

Once usability expert collects data from the user, it has to be represented in such a way that one can analyze data and calculate the required score. There are many ways to represent the collected data. They are: [1 p35-43]

1. **Column or Bar Graphs:**

   Column graphs and bar graphs are similar with the sight difference in their orientation. These graphs are widely used in representing data while performing usability test of a system. This graphical form of representation mainly used to indicate task completion, task times, self reported data etc. These are very appropriate for presenting continues data values.

2. **Line Graphs:**

   These graphs show continuous trends of variables over time. These are the commonly used data representation formats of usability. Line graphs are very much suitable to present the values of one continuous variable as a function of another continuous variable. In this form of representation data points are more important than that of lines.

3. **Scatter Plots:**

   Scatter plots are also known as X/Y plots. In this data is represented as pairs of values. But these are not very commonly used way of representing data in usability tests. In this appropriate scales should be used, the values on the vertical axis can’t be smaller than 1.0.

4. **Pie Charts:**

   Pie charts are used to present the whole or parts of percentages of usability testing. This representation method is useful to present at any time we want to indicate the relative proportions of the parts of a whole to each other.

5. **Stacked Bar Graphs:**

   Stacked bar graphs are generally multiple pie charts represented in bar graph form. These are more suitable when we have a group of datasets, where a single part represents the whole. They are normally used in usability tests to represent different task completion states for each task.
2.3.3 Benefits of Usability Testing

Usability testing is performed throughout the project life cycle in order to minimize the errors that are being occurred. There are many advantages of usability testing, which includes [29]:

- Direct feedback from targeted audience so that project’s team members can concentrate on their requirements.
- This helps discovering the real needs and tasks of the user in the design phase early.
- Both functionality and graphic designs are balanced.
- Within the project team some indecision can be resolved by performing testing on users with a problem to view their reactions to the different options.
- System’s different issues and potential problems are highlighted before it is launched.
- Performing usability test maximizes the competitive advantage.
- This also increases user productivity.
- It decreases user’s adaptation time and errors while using the system.

When we consider about business advantages of usability testing, this can be observed only when the product is developed completely.

2.3.4 Usability Testing Issues

Apart from all benefits of usability testing, we have to think about some of the major issues of this usability testing. Here is the list of some issues that we need to take care of [25]:

- Cost
- Selecting Participants
- Testing Environment
- Time

2.4 System Usability Scale (SUS)

For performing usability testing, there are many performance metrics and scales being used. System usability scale (SUS) is one among them. SUS was initially developed by John Brooke in the year 1986 at Digital Equipment Corporation [1 p138]. SUS is a likert scale created as quick and dirty. It is a commonly used, freely distributed and reliable questionnaire consisting of 10 questions. Scoring the questionnaire yields a usability score in the range 0-100. This is a 5 point scale of agreement. Using this metric user rates the system from 1 to 5, where 1 stands for “Strongly disagree” and 5 stands for “Strongly agree”. Results attained from this can be represented graphically [17].

Here is the list of 10 questions, which are considered as SUS standard questions [1 p138].

- I think I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think I would need the support of a technical person to be able to use this system.
- I found the various functions in this system were well integrated.
- I thought this system was too inconsistent.
- I would imagine the most people would learn to use this system very quickly.
I found the system very cumbersome to use.
I felt very confident using the system.
I needed to learn a lot of things before I could get going with this system.

The standard response format for rating the system using SUS scale is:

<table>
<thead>
<tr>
<th>Strongly Disagree 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Strongly Agree 5</th>
</tr>
</thead>
</table>

Figure 1: SUS response format [18]

Rule for Scoring SUS:

- For odd items: one is to be subtracted from user response.
- For even-numbered items: subtract user response from 5.
- The range of scales are from 0 to 4, where 4 being the most positive response.
- Once all the scores are recorded, add all the converted responses for each user and multiply that total by 2.5. Doing this, scores will be converted to range of possible values from 0 to 100 instead of 0 to 40.

Properties of System Usability Scale:

- **Good SUS score:**
  
  For example, if we consider 500 studies the average SUS score is 68. An SUS score above this average score is considered as above average and anything less than this would be considered as below average score. Normalization is one of the best processes of interpreting the score to convert it to a percentile rank.

- **SUS measures Usability and Learnability:**
  
  Initially SUS was intended to measure only ease – of – use, but from the latest studies [18] it has been noticed that this scale provides a global measure of system satisfaction and sub scales of usability and learnability. From the standard questionnaire mentioned above items 4 and 10 provides learnability dimension and other 8 items provide usability dimension.

- **Reliability:**
  
  In SUS, reliability refers to how consistently a user responds to items. But it has shown to be more reliable and detect differences at smaller sample sizes than home-grown questionnaire and other commercially available ones. Sample size and reliability are unrelated. So SUS yields better results on very small sample size and still generates reliable results.
✓ **Validity:**

This property indicates how well a system can measure what it is intended to measure. This scale effectively distinguishes unusable and usable systems along with or better than proprietary questionnaires. Also, it correlates highly with all other questionnaire-based usability measurements.

✓ **SUS is not diagnostic**

SUS is not a diagnostic tool. Usually, this is used after a usability test in which all user-sessions are recorded on video tape. Low SUS scores indicate researchers to review the system to find out all problems occurred with the interface. SUS can also be used for benchmarking outside the usability test.
3. Method

3.1. Type of Study
The main objective of this study was to compare open-ended usability testing with task-based usability testing to explain which type of testing focuses more on usability issues. In this both quantitative and qualitative tests have been performed. The results obtained are represented using SUS scores graphically. In this study testing was performed on two university websites, one for open-ended testing and another for task-based testing. Usability of these two types was calculated using System Usability scale (SUS).

Both tests were performed on two different university websites as normal observation tests. In this, independent variables are open-ended and task-based usability testing and dependent variable is SUS score.

3.2. System Selection
For this study facilitator/observer has selected two university websites. It was important to consider two similar websites, because tasks designed for one website can be reused to the other website without designing them from scratch.

Selected Websites:
Selection of the system was very important and crucial for this study because, some of the measures had to be taken by the facilitator to select the system.

- The selected system should not be familiar to the targeted participants.
- Selected participants should have access to perform all the required tasks. Also, the important thing needs to be taken care of, that the selected website should be functional without any technical problems.
- For this study, selection of the participants was also important, selected users should be real participants.

Study has been conducted on two different university websites:
1. Texas State University Website (http://www.txstate.edu/)
2. Auckland University website  (http://www.auckland.ac.nz/uoac/)

Texas State University one of the oldest universities established in the year 1899, with its main campus located in San Marcos. There are about 34,087 students who come from 97 bachelors, 87 masters and 12 doctoral degree programs. There are different colleges affiliated to this university, they are Applied Arts, McCoy College of Business Administration, Education, Fine Arts and Communication, Health Professions, Liberal Arts, Science and Engineering, University College and the Graduate College. Students of this university come from all around the globe. Texas state university is a tobacco-free campus. [19]
University of Auckland is founded in the year 1883. This is one of the largest universities with more than 39000 students as of 2010 statistics located in Auckland state. This university consists of eight faculties over six campuses. Auckland University has a strong international focus and is the only New Zealand member of universities 21 and the Association of Pacific Rim universities – international consortia of research – led universities. There are about 4000 international students from 93 different countries studying in this university, while university’s 360° Auckland Abroad Student Exchange program enables New Zealand students to spend one or more semesters for their education [21].
3.3. Participant Selection
Participants play a major role in this study. So participant selection was very crucial in getting accurate results. Twenty eight users were selected for the main test. Initially, before conducting the main test, a pilot test was performed. Four people were considered for this test, they were divided into two different groups. The main aim of this study was to compare two different usability techniques, one group was considered for open-ended usability technique and the other was considered for task-based usability technique. In the same way for the main test 28 people were divided into two different groups, one group for performing one different usability technique.

3.3.1. Participants for Pilot test
It would always be a good idea to conduct the pilot testing before we conduct a main test. This test helps the testers to make sure that there are no technical failures such as computers, internet, selected systems on which we were supposed to perform actual testing. Also this would help the facilitators to know some issues that users might face during the test.

Four users have been chosen for pilot testing, in which two were assigned to perform open-ended usability testing and other two task-based usability testing. All the selected participants belonged to Linkoping University, whose ages fall between 20 and 40. Also all the measures have been taken that the selected participants are good in computers and also surfing internet. In order to make the life of our selected users easy, detailed instructions about this test was given to them. All these activities that are performed by the user have been monitored by the facilitator. Any difficulties faced by them have been solved by the concerned facilitator. All the selected users agreed voluntarily to participate in this study.

3.3.2. Participants for Real test
There were 28 real users selected to perform the main test. These participants were divided into two different groups. One group is selected to perform open-ended usability testing and other to perform task-based usability testing. The age group of all the participants range from 20 to 40 years. Selection of participants for this study was based on specific criteria that is, the selected user should have a moderate knowledge of using internet and also, it was very important that all the selected users for this study should represent as real users for the systems that have been selected. All the selected users belong to different departments of Linkoping University.

We used English as common mode of communication. This is because in this study there have been several people who have participated, they belong to different countries with their respective languages. As the selected websites for this study were in English, they have to understand the content of these selected websites and answer the given questions in order to attain accurate results. Apart from this there was an option of giving some comment on the website they used. There was no restriction of selecting even number of men and women for this study, because the gender imbalance does not affect the required results for this test.

The main aim of this study was to compare both open-ended usability testing and task-based usability testing. So, as we discussed earlier we have divided total number of participants into
two different groups. In this one group of 14 users perform open-ended usability testing and other group of equal number of users performs task-based usability tests. There are two similar websites that have been selected for this study. Each group has been given both the websites for testing, but each group would perform only one testing technique only. For example one group will perform open-ended usability test on both the given websites and other group perform the other test on two websites.

Coming to the selection of participants, it was very important for us to consider the interest level of all the selected participants. We have selected the participants only who were interested and willing to perform all the given tasks in this study. By doing this we were able to get accurate results, also we have managed to get their valuable comments on the websites and testing techniques.

3.4. Test Duration
Task-based usability tests required at most of 60 minutes of each participant’s time to complete all the tasks for both the websites. Open-ended usability tests required at most of 45 minutes of each participant’s time to complete for both the websites which is a bit low time when compared to task-based usability testing.

3.5. Task Design
As this study is about comparing both task-based usability testing and open-ended usability testing, so the set of tasks designed for these methods were different. It is also important that designing of tasks for conducting usability study is very important. So, important measures were taken for designing the tasks for this study. Because once these tasks are designed, participants would perform them and answer the SUS questionnaire by which we have calculated SUS scores.

Tasks that were designed had a proper end state, which helps the user to understand the task clearly and perform them. Generally there are two different types of tasks: [27]

Structured Tasks: This type of tasks makes user’s life very easy by guiding step by step by instructing them what to do, what not to do and where to go. Example for this type of task:
This task is to visit a website to measure its usability.

✔ Go to the home page.
✔ Type a key word to search something.
✔ Check the results.
✔ For the further search use different key words or other criteria to search.
✔ Check the results.

Uncertain Tasks: This type of task makes users bit confused and these tasks are uncertain that they are not sure of completing the tasks to get their desired results. Example for this type of task:

✔ Search the website using keywords and check the results.
✔ Search further using more specific criteria.
For this study we have used uncertain tasks to conduct task-based usability testing, where we have not guided the participant step by step. Using this type of tasks has improved the accuracy of results.

Here are the set of tasks that have been used for task-based usability testing:

**Texas State University:** (www.txstate.edu)

1. Find the information applying for Graduate aid.
2. Find the list of all the available MBA programs being offered between the years 2011-2012 in McCoy College of business administration.
3. Download the credit card authorization form.
4. Find the driving directions to reach admissions office.
5. Find the dates and deadlines for housing for the academic year 2011-2012.
6. Find the list of all colleges affiliated to Texas State University.
7. Find the Rock wall climbing center.
8. Find the undergraduate degree list from the website.

**Auckland University:** (www.auckland.ac.nz)

1. Find the page describing leisure activities in Auckland.
2. Find how to apply to the information technology course.
3. Browse the driving directions to reach Grafton Campus.
4. Find the information about the orientation programs offered to future students.
5. Find all the upcoming events in the university.
6. Find a list with movie clips from current students at the university, telling you what it’s like to study at the university.
7. Find the information about the recreation center.
8. Download the prospectus for postgraduates for the year 2012.

For the open-ended usability testing, there were no specific tasks being designed. A user was given a simple introduction about the websites that had to be used. Later user has been asked to browse the website followed by the feedback given about the given websites.

### 3.6. Test Procedure

This entire test was carried out by the facilitator in two stages. First stage was considered as pilot test and second was the main test. Pilot test was to check if the selected systems were working fine and mainly participants don’t face any unexpected errors or difficulties in the main test. If they face some difficulties, then measures were taken to overcome them in the main test. The detailed test procedure for both the tests is explained below:

#### 3.6.1. Pilot Test Procedure

- Depending on the requirement main facilitator/evaluator recruited other facilitators provided they are equally skilled or have proper knowledge about the things they have to follow to conduct usability test for this study.
- For pilot test four participants have been selected and divided two users to perform open-ended usability test and other two for task-based usability test.
To conduct the test, one participant is called at a time and he/she is asked to fill in user’s demographic details which comprise of date, age, gender and if he/she has previous experience in usability testing.

After the completion of filling participant’s demographic details, facilitator explains him/her with introduction about basic concepts of usability which is designed by the main facilitator followed by a general overview about the websites that he/she have to use for the usability test.

First participant has performed open-ended usability testing. For this initially facilitator gave a small introduction about the open-ended usability testing and afterwards the task has been assigned to the user. This has been performed for two websites selected.

After he/she told the facilitator that the browsing the first website is done, then facilitator gave System usability form to the user to fill in.

Once the user has completed answering the questionnaire for the first website, he/she has been given with the second website. The entire process is same for the second website.

After the completion of filling the form, facilitator thanked the participant for his time and support.

The other group has performed task-based usability testing. In order to conduct task-based usability test, participant has been give the introduction and what are to be followed while performing this test.

After the introduction out the test user is given the tasks designed for this test one at a time.

Once user completes all the tasks of the first website, he/she is given the system usability scale (SUS) standard questionnaire form to be filled in. In this form he/she can give some comments about the system if he/she wished to.

The same process has been followed for the second website.

This process continued for all other participants.

**3.6.2. Main Test Procedure**

- The entire testing procedure is same as the pilot test.
- In this there were a total of 28 users being selected, where 14 of them are meant for Open-ended usability test and other 14 are for Task-based usability test.
- All the participants were selected after making sure that these users fulfill all the requirements such as moderate internet knowledge, proficient verbal and written English skills.
- After the selection of participants, an appropriate location for conducting usability testing has been selected. Facilitator has selected University group rooms as the location for conducting usability tests, where the environment is very calm and users can concentrate on the work while they are performing the usability tests.
- First the users had been asked to perform open-ended usability testing by explaining all the instructions and about the tasks to be done.
- Every alternate participant has been asked to perform task-based usability testing following the entire procedure as explained in the pilot testing.
Order of performing open-ended usability testing and task-based usability testing is not fixed. These two types of testing methods were conducted randomly.

After users perform the usability tests they are given the sheet of SUS form to fill in per website. Once users filled both the SUS forms, facilitator thanked them for their patience and support for the study.

After all the tests are finished, facilitator stored the data safely, because the collected SUS scores are used to calculate the results that are required for this study.

Each type testing method’s data has been kept separately, as the main aim of this study was to compare the results of both testing methods.

At the end facilitator has tabulated all the results and represented them graphically in the ordered manner.

3.7. Dependent Variables

In this study user satisfaction has been measured using SUS which is the dependent variable of this study. In this for calculating the results of both open-ended usability testing and task-based usability testing System Usability Scale (SUS) has been used.

SUS forms are given to the users consisting of standard 10 questions with five different options. SUS score ranges from 1 to 100 in which high score represents good user experience and low score represents poor user experience with the system. Using SUS score t-test has been performed using a special function in excel. This test is used to compare the means by using t-stat and p values attained by performing t-test function in excel.
4. Results:
In this study a total of 28 participants were presented and they were divided into two groups each of 14 participants. One group has participated in Task-based Usability Testing and other group in Open-ended Usability Testing. Facilitator observed participants while performing the test and finally SUS score was calculated. SUS score is used to measure user’s overall satisfaction towards the selected systems. Obtained SUS scores were presented as each participant given scores per each website. Also t-test has been performed on each website’s SUS results.

Figure-4 shows the Texas University’s SUS score for both Task-based and Open-ended usability testing.

![Texas's SUS Score](image)

**Figure 4:** Texas University SUS score in both open-ended and Task-based testing methods

Table-1 shows the t-test results obtained from the SUS scores of Open-ended and Task-based testing methods applied on Texas University.

<table>
<thead>
<tr>
<th>t-Test: Two-Sample Assuming Equal Variances</th>
<th>Open</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>68.57143</td>
<td>48.03571</td>
</tr>
<tr>
<td>Variance</td>
<td>412.2253</td>
<td>536.7102</td>
</tr>
<tr>
<td>Observations</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Pooled Variance</td>
<td>474.4677</td>
<td></td>
</tr>
<tr>
<td>Hypothesized Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>2.494339</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.009652</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.705618</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.019303</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.055529</td>
<td></td>
</tr>
</tbody>
</table>
Figure-5 shows the Auckland University SUS scores obtained from Open-ended and Task-based usability testing.

![Auckland's SUS Score](image)

**Figure 5:** Auckland SUS score by Open-ended and Task-based Testing

Table-2 shows the t-test results obtained from the SUS scores of Open-ended and Task-based testing methods applied on Auckland University.

<table>
<thead>
<tr>
<th></th>
<th>Open</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>71.25</td>
<td>61.60714</td>
</tr>
<tr>
<td>Variance</td>
<td>453.6058</td>
<td>520.7761</td>
</tr>
<tr>
<td>Observations</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Pooled Variance</td>
<td>487.1909</td>
<td></td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>1.15586</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.129125</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.705618</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.25825</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.055529</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1:** t-test of Auckland University SUS scores obtained in Open-ended and Task-based Testing

Figure-6 shows the average SUS scores of both Texas and Auckland University obtained from both testing methods of Open-ended and Task-based.
**Figure 6:** Average SUS scores of Texas and Auckland Universities
5. Discussion
This study has been carried out to propose the relationship between open-ended and task-based usability testing in terms of measuring satisfaction of the user. Exploratory test has been performed to get some useful interesting conclusions and results in this study.

5.1. Metric Selection
There are various usability metrics which are helpful to measure the usability of the system. As the main aim of this study is to measure the relationship between task-based and open ended usability testing in terms of user satisfaction, there were many metrics available such as time on task, task success, SUS score, effectiveness, efficiency etc. This study has been performed with respect to user satisfaction but not to measure the effectiveness and efficiency for both the systems and testing techniques.

The metric used in this study was SUS score, because this study needed only the overall experience of the user towards the system. So for answering this research question SUS was the suitable metric to get the desired results.

There were many other metrics like time on task, efficiency, learnability and effectiveness could have been used for this study. But, in this study as there were two different usability testing techniques used and the main motive was not to assess if the user was able to complete the given task in specified time or how effectively user has used the system. For this reason time on task and effectiveness has not been used. Also, it was not important for the users to complete the task, if they thought that it was taking longer time than expected then they have left that particular task and went on with the remaining tasks while they were performing Task-based usability testing. For this reason task success was also not considered.

When we consider open-ended usability testing users had the flexibility to browse both the websites and express their overall satisfaction about the system by answering the standard SUS form to calculate the SUS score. So, for this reason task success and time on task has not been used.

5.2. SUS Score
SUS scores have been used for this study. In this study user’s overall experience was required. The main goal of this research was to compare open-ended and task-based usability testing techniques, measure the relation between them and propose the better testing technique in the usability expert’s point of view. Two different websites were used as mentioned in the methodology section. From the obtained results and graphs presented in the results section, Open-ended usability testing received better scores when compared to Task-based usability testing for both the websites. Considering the variance from the t-test open-ended usability testing got lower score that the task-based usability testing technique.

We will discuss about the results obtained by performing Open-ended usability testing and Task-based usability testing using both the systems considering one at a time. First we would like to consider the results attained for the “Texas State University” website for
which 14 different participants were selected for each testing technique. The average score for Texas state university from the results achieved by the participants performing open-ended usability testing was 68.57, with the highest score of 95 by 7th participant and the lowest score of 45 by the 6th participant. Coming to the task-based usability testing of the same system, the average SUS score reported was 48.035. The highest score in task-based usability testing was 87.5 by 1st and 2nd participants and the lowest score was as low as 22.5 by 3rd and 6th users. These scores indicate that Open-ended usability testing received higher scores when compared to Task-based usability testing for this website. The other system Auckland university website received the following scores: For the open-ended usability testing the highest SUS score was 95 and the lowest score was 22.5. The average score achieved was 71.25. Considering the task-based usability the highest score is 87.5 by both 1st and 2nd participants and the lowest is 22.5 by two of the users. The average SUS score for task-based testing is 61.60. From the results attained from this website open-ended usability testing attained the better scores than compared to task-based usability testing. Coming to the variance, open-ended usability testing got the value of 412.22 and the take-based testing got 536.71 for the first system Texas University website and values of variance for the second system Auckland University are as follows open-ended usability testing got 453.60 and the task-based usability testing got 520.77. The lowest variance values indicate that the values are very much closer to expected mean value.

There are some cases in which both the selected websites received higher scores for Task-based usability testing when compared to open-ended usability testing. But on the whole results that have been attained by both the websites indicate that they achieved higher scores for open-ended usability testing when compared to task-based usability testing. The reason behind this might be for open-ended testing there were no constructive tasks being designed for both the systems. This gives liberty of exploring the system by themselves without any given restrictions and also users are influenced by the design of the system, User interface etc. which is exactly opposite to the task-based usability testing. But, considering the accuracy level of testing for both the selected systems, though both the websites got high SUS scores for open-ended usability testing indicates only user satisfaction based on their overall experience by using the system without any specific restrictions. When we perform usability testing for the system using constructive tasks, this is goal based testing. Even though the task-based testing got the low SUS scores than the open-ended testing, we may not conclude that open-ended is better than that of task-based usability testing only based on SUS. Performing task-based testing would increase the accuracy level of the testing like how efficiently user is able to use the system. So it is always recommended to practice both the testing methods for accurate results.

5.3 Efficiency and effectiveness
From the obtained results, we would discuss about efficiency and effectiveness as the resultant factors of the testing methods that were studied thematically. Efficiency is related to resource consumption. For example, how much time must the user spend to accomplish a certain goal? Effectiveness is related to whether the system provides the
right functions, in other words whether or not it can help the user solve his or her problems. With task-based testing, it is fairly straightforward to measure efficiency. One can, for example, calculate the time it takes to finish a certain task, and compare it to the time it would take an expert user to finish the same task. Effectiveness can, for example, be operationalized as degree of task completion.

With open-ended testing, it is much less straightforward to measure efficiency and effectiveness. For measuring efficiency we need to compute some kind of resource consumption, but since the user has no clear goal in open-ended testing, we do not really know what the user is trying to accomplish in the system, and hence cannot know when the user has accomplished his or her goals. The same argument can be construed for effectiveness. In other words, open-ended testing can mostly give us indications of a system’s usability in terms of satisfaction.
6. Conclusion:

The main research question for this study was:

✓ What is the relationship between task-based and open ended usability testing, in terms of measuring satisfaction?

The main goal of this research was to measure the relationship between two different testing techniques in terms of user satisfaction; user satisfaction was operationalized by means of the SUS questionnaire. The SUS scores from this study indicated that, users tend to give higher SUS values when participating in an open-ended test, as compared to a task-based test. Our interpretation of this result is that users are only trying out a system without actually trying to use it for specific tasks. By this they are over-estimating the system’s usability.

The results and discussions from this study indicate that there exists the statistical difference between open-ended usability testing and task-based usability testing for both the websites.

From this study it can be concluded that, the usability testing that has been performed with both the systems got higher SUS score for open-ended usability testing compared to that of task-based usability testing. Testing of the system using open-ended usability was not goal based. Open-ended usability testing is an exploratory testing where users did not had the specific goals to be completed. So, the system was tested based on the aspects such as design, user interface etc. Open-ended testing measured the user satisfaction. Task-based usability testing performed was a goal based, where the users were given few constructive tasks to be performed. Even though this method got lower SUS scores than open-ended usability testing, we cannot get into conclusions that task-based usability testing is less preferred to measure the user satisfaction. This is because while performing task-based usability testing users had the certain limitations of what has to be performed within the system. Apart from this the task-based usability testing is fairly straight forward than open-ended testing to measure efficiency and effectiveness.

Hence, this study highly recommends the usability practitioners to use both open-ended and task-based usability testing methods for all the selected systems to get accurate results. Also, we can suggest conducting the study more in-depth by using other usability scales like CSUQ, QUIS etc. also to measure the user satisfaction from the future researchers.
7. References


4. Comparative Usability Testing
   http://teced.com/services/usability-testing-and-evaluation/benchmarking-and-comparative-testing
   Used in may 2012

5. Expert Evaluation Testing
   http://teced.com/services/usability-testing-and-evaluation/expert-evaluation
   Used in May 2012


11. Usability studies.


Glossary:

✓ **Usability**: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use (ISO 9241-11)

✓ **User Experience (UX)**: aspects of the user’s experience when interacting with the product, service, environment or facility (ISO CD 9241-210)

✓ **Usability Testing**: Method of testing used for usability.

✓ **Open-ended Testing**: One of the types of usability testing techniques.

✓ **Task-based Testing**: Also a type of usability testing techniques.

✓ **Tasks**: Work or a question given to the user to complete.

✓ **User**: Participant who performs the tasks.

✓ **Facilitator/Observer/Expert**: Person who monitors the test.

✓ **System Usability Scale**: A standard scale used to calculate the usability score.
Appendix:
A1 SUS form:

Participant ID: _____  Site: __________________________  Date: __/__/___

System Usability Scale

**Instructions:** For each of the following statements, mark one box that best describes your reactions to the website *today*.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I think that I would like to use this website frequently.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2. I found this website unnecessarily complex.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3. I thought this website was easy to use.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4. I think that I would need assistance to be able to use this website.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>5. I found the various functions in this website were well integrated.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>6. I thought there was too much inconsistency in this website.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7. I would imagine that most people would learn to use this website very quickly.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>8. I found this website very cumbersome/awkward to use.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>9. I felt very confident using this website.</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>10. I needed to learn a lot of things before I could get going with this website.</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

Please provide any comments about this website:

This questionnaire is based on the System Usability Scale (SUS), which was developed by John Brooke while working at Digital Equipment Corporation. © Digital Equipment Corporation, 1990.
A.2. Demographic form

Demographic Details

Gender (Male/Female):
How old are you:
Occupation:
Nationality:
Do you have ever participated in usability testing?

Figure 7: Demographic Details