NTNU – Trondheim
Norwegian University of
Science and Technology

Usability Evaluation of Windows 8 with Keyboard and Mouse

Challenges Related to Operating System
Migration in Large Organizations

Nikolas Jansen

Master of Science in Computer Science
Submission date: June 2013
Supervisor: Dag Svanæs, IDI

Norwegian University of Science and Technology
Department of Computer and Information Science

Abstract

The purpose of this study has been to evaluate the usability of Windows 8 when using keyboard and mouse. Sub goals have been to uncover the usability problems and to generate recommendations for organizations upgrading to Windows 8.

Usability testing according to ISO/IEC 25062:2006 was performed on users that had experience from Windows 7. Tests were performed on both Windows 7 and 8 for comparison purposes. Interviews with administrators involved in the operating system migration process were conducted. The interviews were performed to identify problems with previous migration processes.

This study uncovered usability problems in Windows 8. Usability tests performed gave negative results in categories of effectiveness, efficiency and satisfaction for Windows 8. The severity of the usability problems uncovered was high and affected the usability of the system. Problems were caused by fundamental Human Computer Interaction mistakes made by the design and from introducing a new mental model for navigation. Hidden functions proved to be one of the large problems for users when trying to execute their daily tasks.

The interviews of the administrators uncovered experiences made from earlier migration processes. The lack of training provided to the employees and the introduction of new IT solutions in an organization resulted in a lower satisfaction score on the yearly employee report.

Comparing results from the usability testing with feedback from the interviews have been used to suggest topics for training. The topics for training cover the elements and functions that users struggled with during the usability testing of Windows 8. Other measures for an easier transition when upgrading to a new operating system have been proposed.

Sammendrag

Hensikten med dette studiet har vært å evaluere brukbarheten av Windows 8 når brukere benytter seg av tastatur og mus. Delmål har vært å avdekke brukbarhetsproblemer og å generere anbefalinger for organisasjoner som oppgradere til Windows 8.

Brukbarhetstesting ble utført i henhold til ISO / IEC 25062: 2006 på brukere som hadde erfaring fra Windows 7. Testene ble utført på både Windows 7 og 8 for å sammenligne operativsystemene. Det ble også gjennomført intervjuer med administratorer involvert i migrasjonsprosessen for operativsystemer. Disse intervjuene ble utført for å identifisere problemer som har oppstått under tidligere migrasjonsprosesser.

Studiet avdekket brukbarhetsproblemer i Windows 8. Brukbarhetstestene som ble utført gav negative resultater for Windows 8 i kategoriene: nøyaktighet, effektivitet og tilfredshet. Alvorlighetsgraden av brukervennlighetsproblemene som ble avdekket var høy og påvirket brukbarheten av systemet. Problemer med designet var forankret i fundamentale Menneske Maskin Interaksjons feil. Innføring av en ny mental modell for navigasjon gav også problemer. Skjulte funksjoner viste seg å være en av de store vanskene for brukerne når de prøvde å utføre sine daglige oppgaver.

Intervjuene av administratorer avdekket erfaringer fra tidligere migrasjonsprosesser. Mangel på opplæring av de ansatte ved innføring av nye IT-løsninger i en organisasjon resulterte i en lavere tilfredshet på den årlige ansattrapporten.

Sammenligning av resultatene fra brukbarhetstestingen med tilbakemeldinger fra intervjuene, har blitt brukt til å foreslå temaer for opplæring. Temaene dekker elementer og funksjoner som brukerne slet med under brukertesting av Windows 8. Andre tiltak for en lettere overgang når organisasjoner oppgraderer til et nytt operativsystem er også blitt foreslått.

Acknowledgement

This thesis is a result from the work done the last semester of the Computer Science Master Study at the *Norwegian University of Science and Technology* (NTNU), *Department of Computer and Information Science* (IDI). The thesis builds on the research from the depth study performed the autumn of 2012.

I wish to thank my supervisor Dag Svanæs, Professor at IDI, for his invaluable assistance throughout the study. His expertise in the HCI field, has been an important resource in my research. I would also like to thank Terje Røsand, Senior Engineer at IDI and the Norwegian EHR Research Center (NSEP), for support with setting up the NSEP usability lab and assisting as a pilot tester.

Gratitude is also given to Aleksander A. Elvemo for assisting with the transcripts from the interviews and to Marius Q. Wollamo for assisting with processing the results from the usability testing.

I would like to thank all usability test participants that have been involved in this study. I would also like to thank the participants in the interviews for valuable feedback.

Finally, I would like to express gratitude to my parents, my fiancé and my family for their support and encourage throughout the whole process.

Nikolas Jansen

Trondheim, June 2013

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This chapter will give an introduction to the study. It will state the purpose, motivation, reasearch questions and the research method for this study

1.1 Purpose

The purpose of this study is to collect information on how the transition from Windows 7 to Windows 8 will affect workers in an organization. It will gather information on how the new interface presented in Windows 8 is received and what type of problems that comes with it. This information will then be used to analyze the roots of these problems.

The information will be gathered through usability testing performed on both Windows 7 and Windows 8, and through interviews of administrators working with migration processes in organizations. Results from the tests will then be analyzed to see what design choices and what differences between the two systems that create problems for workers. Together with the interviews, this information will be used to give advices on information presented to the users before upgrading to Windows 8.

1.2 Motivation

The sales of desktop computers have been surpassed by the sales of mobile and tablet units[19]. This has led to an evolution of the graphical user interface(GUI) on desktops inspired by touch interface designs[11, 22]. Some of these design decisions have been executed better than others and a bad design can influence the user experience and the usability of the system.

With the release of Windows 8, Microsoft has created a new GUI. This new GUI supports both desktop and touch devices. Adapting elements from both worlds and combining them into one system. Microsoft says:

> *Windows 8 delivers a fast, fluid, no-compromise experience for businesses; along with an user interface that responds equally well to touch as it does to a keyboard and mouse.[40]*

Together with Windows 8, Microsoft has released a collection of their User Experience Guidelines[28] to ensure that applications developed for their platform follows the same design principles as the rest of the Operating System.

Evaluation of usability through usability testing by following ISO/IEC 25062: 2006[13] makes it possible to identify problems and to gather information on how users experiences the system. Using a similar approach will potentially uncover problems with Windows 8 and provide a good platform for the suggestions on how to prepare an organization for the migration from earlier versions of Windows 7 to Windows 8.

By performing semi-structured interviews of administrators involved in the migration process, it will be possible to gather input on what they have experienced as problems from earlier migrations and to see if this migration will be different from them.

1.3 Research Questions

The main goal of this study will be to evaluate the usability of Windows 8 on a desktop computer with keyboard and mouse. Give recommendations to organizations on how they best can handle an upgrade to Windows 8. Structuring the goal into different research questions will help this study with the evaluation of the goal. The goal has been decomposed into these questions:

Research Question 1:

> *How is the usability*[1] *of Windows 8 when using screen, mouse and keyboard for users with experience from Windows 7?*

Research Question 2:

> *What aspects of Windows 8 are causing the observed usability problems, and what are the roots of these problems?*

Research Question 3:

> *What recommendations can be deduced from this study for organizations introducing Windows 8?*

This evaluation will be done both through usability testing and feedback from future users and by interviewing administrators of the system. The testing will give information on how well the system will be received by users, how it will influence their daily work and how much help they will need. Comparing the results from the usability testing to the information gathered from interviewing administrators, will help determining if this version will result in more problems then earlier migration processes and how to prepare the users for Windows 8.

[1] Usability defined in ISO/IEC 25062:2006[13] with efficiency, effectiveness and satisfaction.

1.4 Research Method

The usability tests will be performed on a virtual machine running Windows 7 and Windows 8. The systems will be configured the same way users from NTNU are used to, and the users testing the systems will be people from the administration at NTNU. These tests will be used to analyze and observe the problems that occurs when upgrading the operating system on a workplace from Windows 7 to Windows 8. The usability tests will be performed according to ISO/IEC 25062:2006[13], and are described more detailed in Section 5.2.

In addition to usability testing, there will be performed semi-structured interviews[30] of administrators managing and deeply involved in migration processes related to moving from one operating system to another one. This is to uncover any discrepancies from earlier migration processes and to identify how they have mitigated problems in these migrations.

Chapter 2

Background

This chapter will give a brief introduction to the history of Windows 8 and usability testing.

2.1 Windows 8

Windows 8 was released in the autumn of 2012[1]. This operating system was designed by Microsoft to bring both the desktop computers and tablets to one platform. Supporting more input options natively than earlier versions of Windows and bringing a whole new design to the operating system. Together with Windows 8, Microsoft launched the Windows Store. This store contains applications[2] for the operating system, making it easier for users to discover and download them. Windows 8 brings features such as improved security and cloud storage options for the users.

[1]Microsoft Product Lifecycle. `http://support.microsoft.com/lifecycle/search/default.aspx?sort=PN&alpha=Windows+8&Filter=FilterNO`, 2013. [Online; accessed 28-May-2013]

[2]A small software system designed to perform a specific task.

2.1.1 The Windows User Interface(UI)

Windows UI was the name of the interface that came with Windows 8. The same user interface has previously been used on Windows Phone 7 and has internally had the codename of "Metro"[25]. It brought large tiles with a flatter design striped of unnecessary visual effects making the user interface more functional.

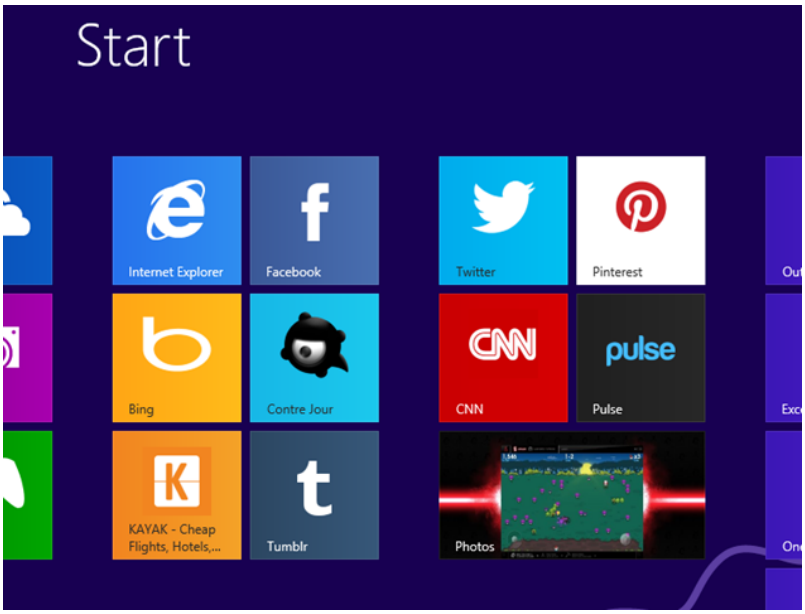


Figure 2.1: Windows UI in Windows 8.

The user interface was designed to be more playful and to support both touch and mouse input. Full screen applications launches from the start screen, and it has new features that are available through gestures or cursor movement.

2.1.2 Adaption rate

With the launch of Windows 8 it was expected to boost the PC sales, but this did not happen[18]. Windows 8 was not received as good as expected. The reception was mixed with both good and bad reviews[9, 22], where user experience expert Jakob Nielsen said it was disappointing for both novice and expert users[34].

The new user interface has been forced on the users upgrading to Windows 8. Users have therefore seen the need for third-party applications that bring their normal user interface back to their computers. Applications such as RetroUI[3] are being used by users to escape the Windows UI.

Microsoft has understood that Windows 8 has been badly received and that something has to be done[41]. They are working on an update to improve the usability of the system which will be free for users to download[26].

2.1.3 Interaction styles

There are several ways to interact with computers. Depending on the device and the intention of use, the interaction style lets humans modify the state of the computer[38]. Small phones uses touch screens that allows users to manipulate the interface with a touch of their finger, while desktop computers uses a keyboard and a mouse or a trackpad for input.

(a) Touch.

(b) Keyboard.

(c) Mouse.

(d) Trackpad.

Figure 2.2: Some of the interaction styles available.

[3] `http://retroui.com/`

Different interaction styles require a certain amount of awareness from the developer or designer creating applications for that interaction style[17]. It is easier to be accurate with a mouse, compared to using a finger to select a small object in the user interface. This is simply because the finger is bigger than the point of the cursor. Touch and trackpad allows users to use gestures such as pinch or swipe to enhance their interaction with the computer, adapting movement styles from real life.

(a) Pinch. (b) Rotate.

Figure 2.3: Gestures used with touch and trackpads.

2.2 Usability Testing

Usability testing has been used to evaluate interface designs since the start of 1980s[38]. It is an empirical research method which means it gather information through observations. This information is used to eliminate design problems and minimizing the frustration of using the product[37]. Chapter 5 will describe this research method in further details.

Chapter 3

Usability

This chapter will give a brief definition of what usability is, and how usability tests can help us improve it. It will also explain how usability testing is performed in the facilities at NTNU.

3.1 What is Usability?

Usability can be described in many ways. Dumas and Redish gives a very short and specific definition of usability[15]:

> *People who use the product can do so quickly and easily to acomplish their own tasks.*

This definition does not include much about how the user liked using the system. The International Organization for Standardization(ISO) has a more concise definition of the term usability[14]:

> *Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*

The same document defines the context of use as:

> *Users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used.*

These definition covers how the system is used, what the user thought about using it and the context of it. This can be broken further down into several metrics to rate usability, and to give a better insight on what usability consists of. These metrics are:

Effectiveness
How precise the user is able to perform a task on a system?

Efficiency
How much resources(for example time, or number of actions) are used to perform the task on a system?

Satisfaction
Does the user like using the system?

With the user in mind, also called user-centered design[1], it is possible to achieve these goals. Making an user-centered design is about getting feedback from the user during the design and development process. To always have in mind how the user solves this problem, and consolidate the users when in doubt. The user is the measure of how good the system performs and the user defines how the system scores on usability.

3.2 How to Test Usability?

There are many ways to create a good user experience. Having knowledge of expert opinions is always a good idea, and with the user in mind when designing the system, the design process is of to a good start. By performing usability testing through the project lifecycle, valuable feedback can be gathered from the users to create an interface with good usability[1]. Having an user-centered approach will help the developers to address the weakest parts of their system, and give feedback on design decisions[10].

An user-centered design can be done in many different ways during the development lifecycle[1, 38] shown in Table 3.1:

Method	**Purpose**	**Phase of the project lifecycle**
Background interviews and questionnaires	To collect data and to understand the user better	When starting the project
Focus groups	Will help developers with design issues and feedback on the system	At the start of the project
On-site observation	To both collect information of the context the system will be used in, and what the daily problems of the users are	At the start of the project
Role playing/ simulations	Will give a broader understanding of what the user expects from the system and their needs	Start to mid of the project
Automated evaluation	Gives feedback on deviations from standards or best practices. This method exclude actual users, but is based on well tested principles tested on users	Mid to end of the project
Usability testing	To measure the usability of the system and provide feedback on very specific elements that are designed bad	Abras[1] says it should be at the end of the project, while others[15, 31, 38] thinks it can be performed in iterations throughout the project
Interviews and questionnaires	Gives qualitative data on how good or bad the usability of the system was	End of the project

Table 3.1: Methods of user-centered feedback.

The purpose of this study is to gather information on how the usability of Windows 8 is for users with experience from Windows 7. Windows 8 is a finished product, and therefore this study will focus on usability testing, interviews and questionnaires. These methods are mainly used for evaluating systems at the end of the development lifecycle.

Usability testing

The purpose of usability testing is to increase the usability of a system. At the same time performing these usability tests can save the developers some time and reduce the cost of the project by removing incorrect design at an early stage[15].

The usability testing can be performed in many ways[38]. At an early stage of the project, paper prototypes are a good option to get feedback on the layout and the time between each iteration of usability testing will be shorter than performing testing on an interactive prototype. At the end of the project lifecycle, the developers can focus more on the details of the system. The different testing methods include a real potential user of the system performing tasks on the system to provide real data. Observing and recording each usability test can help the developers to analyze their system, and correct the usability flaws[15].

When performing the usability testing, developers sets goals that they want to achieve from the testing. The developer must have the reporting of the results in mind when doing the test planning[13]. This will ensure that the purpose of the test is fulfilled. The test is planed with a set of tasks that are testing daily tasks performed on the system. These tasks should allow the user to explore the system, or the parts the developer want to test of the system, and let them use some time while executing the tasks on the system[1].

The next step after making the tasks is to find participants for the test. Participants should be potential users of the system and preferably not to be overqualified for the tasks. Overqualified participants can for example be removed by asking them questions of how much computer experience they have and how they use the computer on a daily basis.

The number of users will have an impact on the number of errors found in the usability testing. From Figure 3.1, it can be seen that when the number of participants increase, the number of undetected errors becomes smaller.

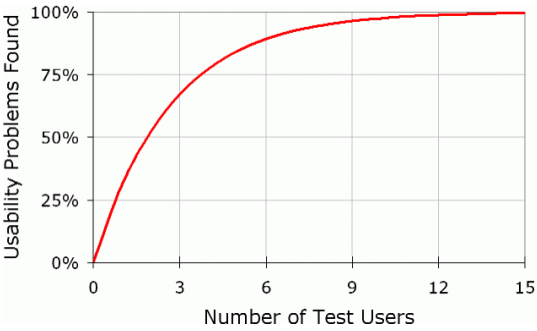


Figure 3.1: Number of users needed to find percentage of errors[32].

Usability expert Jakob Nielsen has experienced that after testing on 5 participants, 85% of the errors have been found[32]. Molich[29] says he has experienced that 6 participants are the best number. It is possible to exclude more of the usability errors from the system by testing more users. For large systems this will help to increase usability, but it requires both time and fundings to run more tests.

The next thing to consider when performing usability testing, is the testing environment. Making the test facility resemble the environment where the system will be used, allows for more valuable and correct feedback. To make the most of the tests, it is wise to perform videotaping of the process. This will help when reviewing the results from the test[1, 29]. If the tests are being recorded, remember to inform the participant of this.

Before performing tests, select a test leader to guide the participant through the process. The test leader is in charge during the testing, and acts as an interviewer to help extracting thoughts from the participant. The test is an expanded version of a qualitative interview and can be compared with a semi-structured interview as described in Chapter 6.

During the test, it is important for the test leader to inform the participant of how the testing will be performed. The test leader must explain the technique of "think aloud"[1, 38] before running the test. Be open for questions that the participant has in front of the tests, but be careful not to give away information that will affect the results from the test.

After the tasks, it is important to gather all loose ends and get answers to all the questions that might be unanswered. A System Usability Scale(SUS)[4] can be a good way to grade the usability of the system together with the

observations made during the test. It will reflect how satisfying the usability of the system is in the eyes of the users. Bangor et al.[2] have made a scale from 959 systems rated with the SUS-form where it is possible to compare the mean score of a systems SUS-score. It is possible to check if the score is on an acceptable level compared with the scores of other systems.

3.3 NSEP Usability Lab

This section describes some of the features in the facility used by NTNU to perform usability testing.

3.3.1 The facility

The NSEP Usability Lab is an usability testing facility used by NTNU. It provides different opportunities to record and observe how users interact with different types of systems. The facility contains different rooms with movable walls that can be customized into the preferred environment for different types of tests. This allows the facilitators to create an as life like situation as possible to collect valuable data from it.

The setup of the lab can be seen in Figure 3.2 and 3.3.

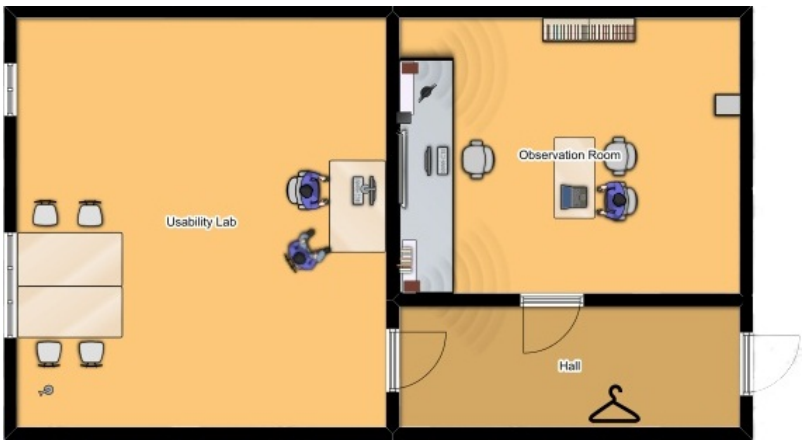


Figure 3.2: The NSEP Usability Lab in 2D[21].

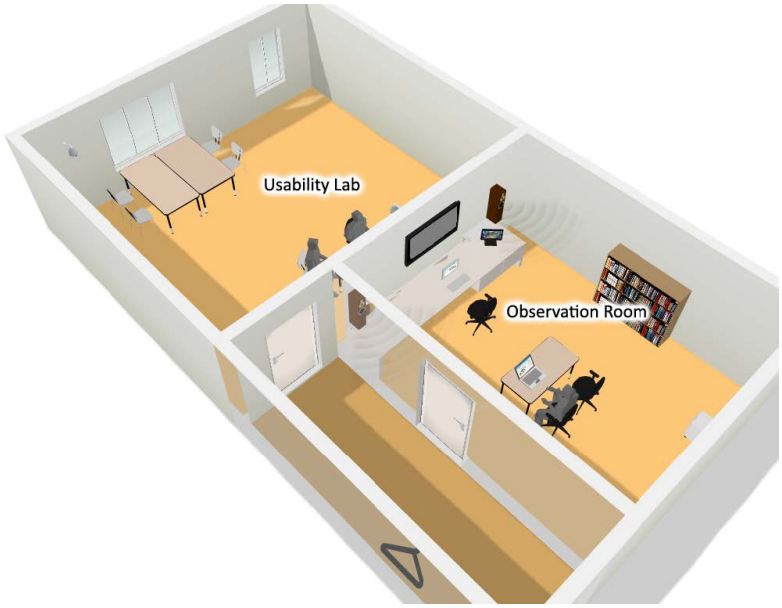


Figure 3.3: The NSEP Usability Lab in 3D[21].

There are different types of equipment that can be used to gather specific information of how users interact with the system during tasks. Some of them are more valuable than others. In this study the Tobii X60 Eye Tracker and the screen capturing software are of great importance, hence the system being based on a graphical user interface.

Tobii X60

Eye Trackers are being used more often to evaluate the usability of systems. They can gather information of what the user is looking at when using a system. What objects in the graphical user interface that attracts their attention, and how they search through the system to find the information they need.

Figure 3.4: The Tobii X60 Eye Tracker.

The Tobii X60 Eye Tracker[1] in Figure 3.4 uses infrared lights to track the movement of users eyes. The Tobii X60 Eye Tracker provides accurate readings that records the persons gaze on a screen or other types of interfaces. The Tobii X60 tracks the eye position and the eye gaze with an accuracy of less than 0.3 degree with a frame rate of 60 frames per second.

Together with the Tobii Studio version 2.3.2.0 software, the computer can record both screen and eye movement. The software makes it possible to extract gaze heatmaps and plots that can give information on what objects that attracts the users focus during execution.

[1]Tobii Technology. The tobii x60 eye tracker. `http://www.tobii.com/en/eye-tracking-research/global/products/hardware/tobii-x60x120-eye-tracker/`, 2011. [Online; accessed 28-May-2013].

Chapter 4

Design Principles and Guidelines

This chapter will give a brief definition of principles and guidelines that user experience-experts have deduced from their studies. These are principles and guidelines that have been proven to help creating good interfaces. They describe good practices and how to avoid some pitfalls.

4.1 Basic Design Principles

The following principles are some of the basic rules for designing user interfaces. Understanding how human computer interaction works from the bottom up, will help analyzing the problems occurring during usability testing and give better understanding to some of the best practices explained in Section 4.2.

4.1.1 Gestalt principles

These principles describe how the human brain percepts object relations with one another or with the environment surrounding the object[42]. The principles explains how the human brain mentally completes lines and groups of objects through continuity and closeness of objects. The same mental completion happens with objects in the same color or shape.

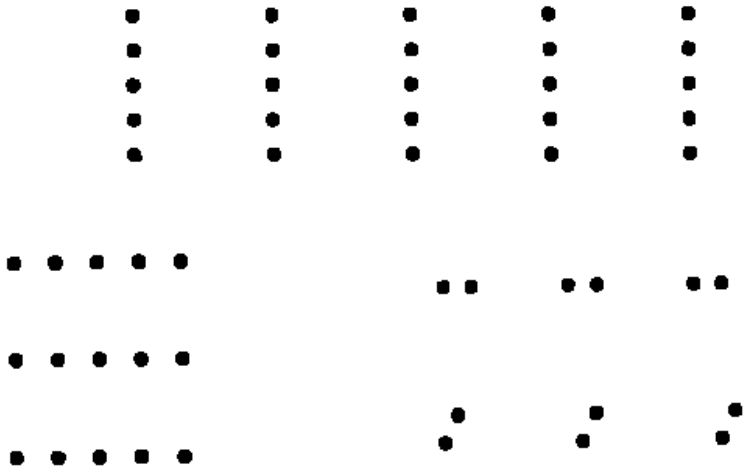


Figure 4.1: Examples of Gestalt principles where the brain mentally completes lines and grouping of objects[42].

4.1.2 Affordance

The concept of affordance in software engineering originates from *The psycology of everyday things* by D. Norman[36]. The term is used to explain how humans perceives the properties of an object, and how that object primarily can be used. The use of affordance in software engineering is still wage, but McGrenere et al. has created a framework that assesses the affordance of user interfaces[27].

Figure 4.2: A hammer affords grasping, and therefore afford a solid extension to the human arm that can be used to hit other objects with.

4.1.3 Minimalism

At an early state of software engineering, user manuals were overly complicated and Carol et al. created a way to minimize these manuals[8]. This was done by focusing on the most important tasks and processes, removing unnecessary information to reach the goals the user wanted to achieve. The results were optimized completion times of tasks when using the minimized manuals compared to the overly complicated ones. Minimalism has then been used in user interface designs to focus on the important parts of the user interface, making it easier for the user to achieve their goals by removing unnecessary features from the user interfaces[6, 7].

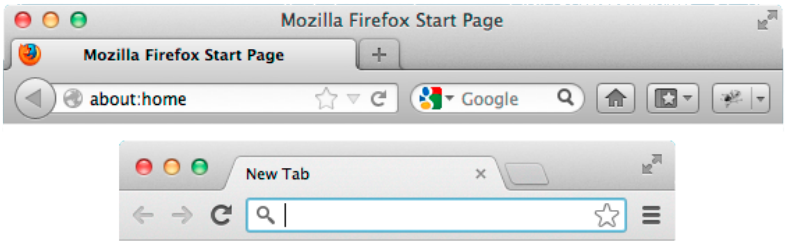


Figure 4.3: Minimalism in browsers. The top one has separate address bar and search field, while the bottom one has an unified field for both address bar and search.

The newer era of minimalism in graphical design, similar to the minimalism used in Windows 8, is inspired by the *De Stijl*[1] movement. The journal origins from 1917 and was created by Theo van Doesburg and Piet Mondrian[43]. The design uses the minimal amount of boxes and grids to compose order.

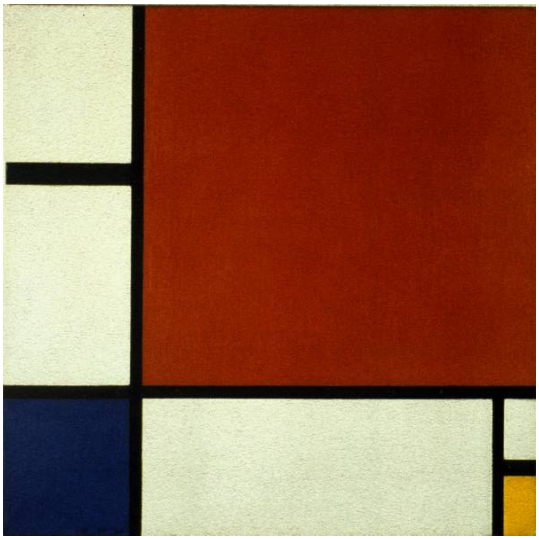


Figure 4.4: Piet Mondrian, Composition II in Red, Blue, and Yellow, 1930.

[1]Dutch for "The style".

4.1.4 Adapting design to users

When designing user interfaces the designer needs to have the user in mind. Designing for a western user group will be different from designing for an eastern user group[16]. Users from different places scan the content of the system differently. Western users will scan through the content of the system in a F-shaped pattern[33]. User groups from different ages want different design. Creating a system for children need to have a more exiting and tailored design since they have a different attention span than adults[23].

Figure 4.5: How users read in a F-shaped pattern[33].

4.2 General Guidelines for Interface Design

These rules applies on top of the principles from Section 4.1. They are general advices on how to prevent usability problems generated by the user interface. Following them will contribute towards a good user experience when designing the system.

4.2.1 Smith and Mosier

One of the first major public collections of usability guidelines was created by Smith and Mosier in 1986[39]. Smith and Mosier tried to address the problems with user interfaces, and how to design the interfaces. They focused on how elements can be designed in the interface to increase the usability of a system. Creating consistency in the user interfaces improves the usability of the system.

4.2.2 Eight golden rules

The eight golden rules by B. Shneiderman[38] are more general advices on how the designers can make more user friendly user interfaces. It describes how the user interface assists the users so that the users can complete their intended tasks. This shall be achieved through consistency, the use of universal usability and by letting the users feel they are in charge of the system. By always displaying the system status, it will help to create a feeling where the users think they are in control. The prevention of errors is highly important to create a good user experience.

4.2.3 Constraints

Creating constraints in the user interface will help guide the user to the right actions[3]. It will assist in guiding the user through the process and prevent errors. Constraints helps the user to understand different dependencies in the user interface.

4.2.4 User interfaces should provide benefits

Using tools to aid users performing their task should benefit the user instead of creating barriers that the user has to overcome. At a certain point the user would rather not use the system, because the tool creates more problems than benefits from using them[24]. Designers must have in mind that the user interface they design need to add more value then pain for the users. Making the user look for features may cause frustration, therefore the essential information on how to perform tasks must be presented in the user interface. Preventing users from having to remember all functions, and instead displaying them to the users, helps them find the solution for their problem. There is an old saying that describes this phenomenon: "Out of sight, out of mind"[35].

4.2.5 Windows 8 - user experience guidelines

The guidelines[28] created by Microsoft gives the designers and developers an introduction to the platform they are going to design an application on. It is created for desktop computers and touch capable devices. They describe how applications support different formats and how to adapt to contexts that the user might use the system in. The guidelines explains some of the design principles they want the reader to have in mind when designing their application on the platforms.

The Windows 8 guidelines cover the user interface elements commonly used in their applications. They include descriptions on the elements, appearance and guidelines for how they can be placed and used in the application. Some restrictions are included to ensure that the elements follow best practices.

Figure 4.6: Instructions from Windows 8 user experience guidelines[28].

Microsoft explains in detail how they want the navigation and layout of applications created for Windows 8 to be designed. The guidelines describes how the anatomy of navigation is supposed to be, and how the application can be decomposed into sub-pages with more details depending on what level of information the user is on. As alternatives to the tree-like structure, they describe a tab structure mainly used for sections or parts on the same level of an application.

Chapter 5

Usability Tests

This chapter describes the purpose of the tests and how they were executed. Then Section 5.4 presents the observations and results from these tests.

5.1 Purpose

The usability tests were performed to provide usability feedback on the system. These tests were created to test and discover the problems a normal user has in Windows 8 with experience from using Windows 7. Tasks the participants performed, challenged the system and were created with normal daily tasks in mind. Usability tests were performed on participants with no experience from using Windows 8 in front of these tests. This was done to gather valuable feedback on usability problems that the design and structure created, and to prevent invalid feedback from users who already knew how to perform the tasks on Windows 8.

5.2 Research Method

The execution was based on theory described in Section 3.2 and followed the checklist in ISO/IEC 25062:2006[13]. To get accurate data, the test had to be performed on potential users of the system. They were selected from the administration in several departments at NTNU, with none computer science background. The computer scientists were excluded due to the fact that they are more familiar to graphical user interfaces and the exploration of these, which provide incorrect data from performing tests with these users. The probability of computer scientists without any knowledge of the new version of Windows, was considered to be rather small. The gender distribution of the participants had to be equal to the distribution of employees at NTNU[I]

to simulate a real life scenario where users are presented a system upgrade.

The users were going to perform the same tests on both Windows 7 Professional(Build 7601) and Windows 8 Pro(Build 9200). The operating systems were installed as Virtual Machines using VirtualBox(Version 4.2.6 r82870)[1]. The tests challenged the core functionality and applications that comes preinstalled with the operating system.

Because there were two systems, the participants had to be divided into two groups. The difference between those two groups was that the groups switched up the order of the operating systems being tested. This was done to elude errors caused by sequence. Group 1 performed Windows 7 tests before Windows 8, and the opposite sequence in Group 2.

The tests had to be run on at least 5 people to ensure that it uncovered the most of the usability problems in the system[32]. This study ended up performing 12 usability tests, with 6 people in each group.

In front of performing the real usability tests, there was performed a pilot test to exclude any error sources from the test plan. This was done to prevent tests with incorrect data influencing the results.

To ensure that the participants had the background wanted for the tests, there were created two forms that they had to fill out before performing the tests. The first form, Appendix B, covers their background, and the second, Appendix C, covers their experience with computers.

The NSEP lab, as described in Section 3.3.1, was configured with one computer connected to a keyboard, mouse and screen. For this study the facility was designed as an office with a desk and a desktop computer. This created a platform for productivity and set focus on the system. This made it easier to interview the users in an office situation[30] and created situational data.

A webcamera and microphones were used to capture the users actions and what they said during the tests. The Tobii X60 Eye Tracker from Section 3.3.1 and the Tobii software were used to record the the screen and the users gaze on the screen.

[1]Oracle. Virtualbox. https://www.virtualbox.org/, 2013. [Online; accessed 07-May-2013].

A test leader was in the room during the test, and guided the testers through the process. The test leader had both the scenario and tasks described in Section 5.3.1 at hand.

The participants were put in front of a computer and given different tasks one at a time to complete. They were presented a fresh installation of the operating system with their user pre-configured with email and a local disk, similar to the computers configured in their offices. They were introduced to the "think aloud"-method, and were told to ask questions during the process. They were also told that the test leader could not answer any of these questions during the execution of the tasks. The questions were going to be discussed after they were finished with the tasks. If they felt like they were going quit the test, they had to inform the test leader so that he could provide the help that IT-support normally provides.

In front of performing the tasks on Windows 8, the users were shown the quick introduction given when Windows 8 is configured on the computer before users are able to log on. A screen capture from the introduction can be seen in Figure 5.1.

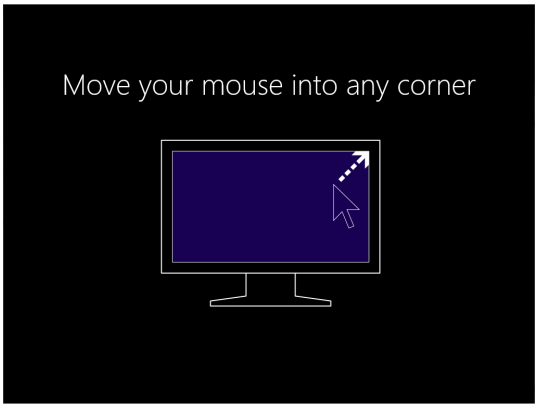


Figure 5.1: Introduction given to users before the tests.

After the tasks were finished, the participants had to answer to the forms in Appendix D and E. The test leader collected answers to unresolved questions made during the testing and questioned the participants for what they thought of the system. He tried to extract their opinions on elements that they struggled with during the test.

The results will then be used to analyze and address the usability problems with Windows 8. The errors will then be rated after their severity[15]. This will be done in four different levels: *Critical(Level 1), Significant(Level 2), Minor(Level 3) and Not-essential(Level 4).* These four levels are further explained in Section 5.4.4.

In the next sections are the answers to the forms and the results from the usability tests.

5.3 Participants

Table G.1 and G.2 in the Appendix shows the participants answers to the forms in Appendix B and C, which covers information about them and the knowledge they had with computer usage in front of performing the tests.

Table G.3 and G.4 in the Appendix contains the answers users wrote down after the test in the form from Appendix D. These questions were given to the user directly after they finished the test together with the System Usability Scale shown in Appendix E.

5.3.1 Scenario and tasks given to the users

Because the participants were from Norway, the scenario and tasks were written in Norwegian. The exact scenario and tasks handed to the participants can be seen in Appendix F. For convenience the next paragraph will give a short summary of the scenario and tasks.

The scenario: It explained that the participant came back to work one day and their computer had been upgraded. The participant was going to perform some daily tasks and had to do them on the operating system presented after the upgrade.

Task 1: Check a news web page in a browser of the participants choice.

Task 2: Find a file on the desktop.

Task 3: Send this file by e-mail.

Task 4: Find a presentation stored on the local disk.

Task 5: Turn off the computer before leaving the desk.

These tasks were chosen because they are common tasks performed in the participants daily work situation. The completion criteria for the tasks are written in Appendix F together with the tasks, but these criteria were not presented to the users.

5.4 Observations and Results

This section covers observations and results from execution of the usability tests described previously in this chapter. Described in Section 3.1, usability can be decomposed into three components. The next Sections 5.4.1, 5.4.2 and 5.4.3 will answer how well the operating system performed measured by effectiveness, efficiency and satisfaction. The most interesting observations are summarized in Section 5.4.4. Then in Section 5.5.1, there will be performed a T-test on the results from the execution times in Section 5.4.2.

5.4.1 Effectiveness

Effectiveness has been measured by how many tasks the users were able to complete during the usability testing. The completion rate is presented in Table 5.1.

Task	**1**	**2**	**3**	**4**	**5**	**Total**	**Compl. all**
Win. 7	100%	100%	100%	83%	100%	97%	83%
Win. 8	75%	92%	50%	83%	58%	72%	33%

Table 5.1: Completion rate for the two operating systems.

Most of the tasks given to the participants were harder to perform on Windows 8. While only one task on Windows 7 proved to bring some difficulties for the participants, all of the other tasks were observed to be harder to execute in Windows 8. The problem in Windows 7 occurred because the participants were used to having their local disk named with "Local disk (M:)" instead of "Local disk (C:)". The problems experienced with Windows 8 are listed in Section 5.4.4.

The column named "Total", presents the total percentage of tasks completed by the participants during the usability tests. "Compl. all" presents the percentage of users that completed all the tasks without any assistance or without failing the tasks.

5.4.2 Efficiency

Table 5.2 shows how much time the participants used to execute each task in Windows 7, while Table 5.3 shows the execution time for Windows 8. Some of the cells in those tables have been marked with an "A" or a "F". The "A" indicates if the participant was given assistance, while the "F" indicates when the participant did not complete the task. Aid given when assisting, did not explain where or how the participant could complete the task. It was only subtle hints reminding them to perform the task as they usually do, to think of how they normally performed it or to remember what the introduction demonstrated.

Table 5.4 shows the percentage of how much more time participants spent on completing tasks in Windows 8 compared to the time they used in Windows 7. Figure 5.2 shows the difference between the time used on the different tasks without the data where participants got assistance or failed completing the task.

Participant	**Task 1**	**Task 2**	**Task 3**	**Task 4**	**Task 5**	**Total**
P1	3:15	0:21	2:46	4:02(F)	0:26	10:50
P2	2:27	1:02	3:21	6:00(A)	0:24	13:14
P3	2:40	0:30	1:48	1:22	0:49	7:09
P4	2:08	0:16	2:07	1:07	0:30	6:08
P5	2:29	1:04	2:36	1:38	0:58	8:45
P6	0:46	0:13	1:40	0:46	0:27	3:52
P7	1:17	0:16	1:54	1:15	0:13	4:55
P8	1:16	0:09	1:15	0:36	0:15	3:31
P9	0:48	0:12	1:09	0:17	0:21	2:47
P10	1:41	0:14	4:38	1:07	0:17	7:57
P11	1:37	0:17	3:15	0:54	0:07	6:10
P12	0:47	0:32	1:17	0:34	0:24	3:34
Mean G1	2:17	0:34	2:23	2:29	0:36	8:20
Mean G2	1:14	0:17	2:15	0:47	0:16	4:49
Total Mean	1:46	0:26	2:19	1:38	0:26	6:34
Mean without fails	1:46	0:26	2:19	1:25	0:26	6:21
Mean without fails or assists	1:46	0:26	2:19	0:58	0:26	5:54
Confidence interval without fails or assists α **= 0.05**	0:28	0:11	0:36	0:15	0:08	-

Table 5.2: Time in minutes used on each task for Windows 7.

Participant	Task 1	Task 2	Task 3	Task 4	Task 5	Total
P1	3:34	1:18	6:32	1:31	2:41	15:36
P2	8:09(F)	0:42	7:11(F)	1:13	5:04(A)	22:19
P3	3:11	0:10	7:32	1:41	3:46	16:20
P4	11:10	6:23(F)	7:25	1:25	2:26(F)	28:49
P5	2:56	1:46	6:46(A)	1:49	4:23(F)	17:40
P6	1:32	0:52	3:11	0:44	1:00	7:19
P7	3:10	0:48	14:41(F)	4:54(F)	0:53	24:26
P8	2:19	0:29	3:57	1:03	1:38	9:26
P9	6:29(F)	0:41	7:28	4:06(F)	1:34	20:18
P10	6:28(F)	1:20	13:48(F)	4:31	4:23(F)	30:30
P11	3:00	0:17	8:14(A)	3:58	3:18(F)	18:47
P12	5:01	1:41	5:50(F)	6:20	1:02	19:54
Mean G1	5:05	1:52	6:26	1:24	3:13	18:00
Mean G2	4:25	0:53	9:00	4:09	2:08	20:33
Total Mean	4:45	1:22	7:43	2:46	2:41	19:17
Mean without fails	3:59	0:55	6:23	2:26	2:12	15:55
Mean without fails or assists	3:59	0:55	6:01	2:26	1:48	15:08
Confidence interval without fails or assists α **= 0.05**	1:52	0:19	1:33	1:09	0:47	-

Table 5.3: Time in minutes used on each task for Windows 8.

	Task 1	Task 2	Task 3	Task 4	Task 5	Total
Time usage	226%	215%	260%	253%	416%	257%

Table 5.4: Percentage of time usage when completing tasks on Windows 8 compared with time spent on Windows 7. Based on the means without fail or assists.

From these results, there are some markable differences in the dataset. Each of the groups show that the sequence from how the test was executed had some influence on the results. Merging these data gives a result that eludes the influence of sequence by some degree.

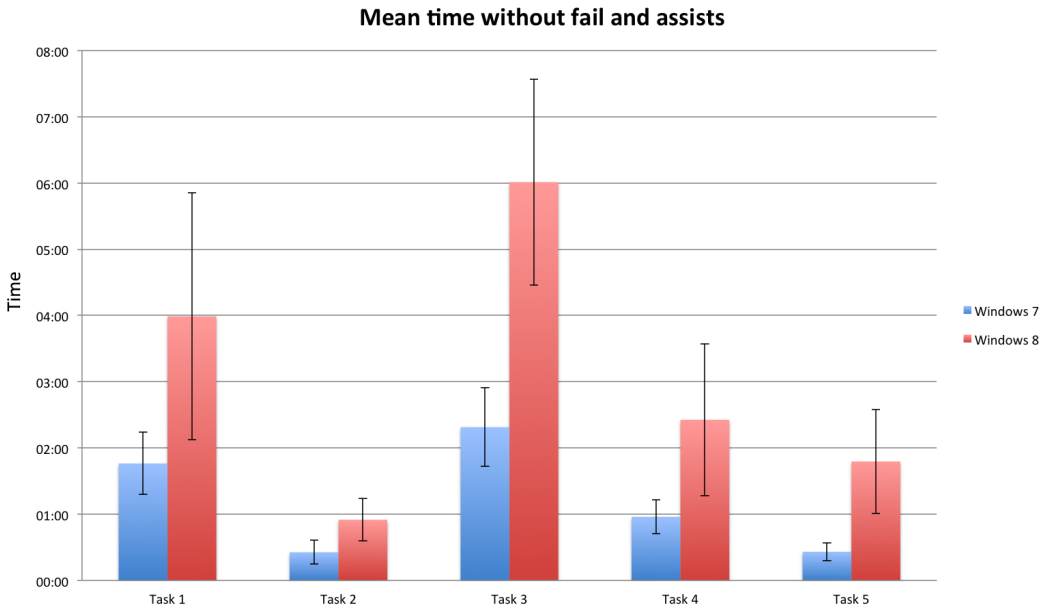


Figure 5.2: Results from each task executed on the two different operating systems based on the mean time without fail and assists.

By removing the failures and assists, the results gives a better picture of how long time it takes an user to complete the tasks. This does not remove the fact that some of the tasks were harder to execute in Windows 8, but it gives a more accurate summary of the time used to complete the task. Looking at the time usage of a participant that not completed the task, defeats the purpose of comparing completion time.

5.4.3 Satisfaction

The satisfaction of the participants when using Windows 8 was addressed by the System Usability Scale-forms[E]. Tables 5.5 and 5.6 shows the results from the participants System Usability Scale-forms after calculating the scores. The scores given by each participant are presented in Appendix H.

Participant	P1	P2	P3	P4	P5	P6	Mean
1	0	2	3	3	0	3	1.83
2	1	2	4	1	0	3	1.83
3	2	2	3	1	0	3	1.83
4	2	2	4	3	0	3	2.33
5	3	2	2	2	0	2	1.83
6	1	2	3	3	0	3	2.00
7	4	2	2	3	1	3	2.50
8	2	2	4	2	2	4	2.67
9	0	2	2	0	0	2	1.00
10	1	2	4	3	0	3	2.17
Score	40.00	50.00	77.50	52.50	7.50	72.50	50.00

Table 5.5: The system usability scale given after the test by group 1.

Participant	P7	P8	P9	P10	P11	P12	Mean
1	2	0	0	0	0	4	1.00
2	2	0	0	0	0	4	1.00
3	0	3	0	0	0	2	0.83
4	0	4	0	0	0	1	0.83
5	0	1	0	0	4	2	1.17
6	0	0	0	0	4	3	1.17
7	3	0	2	0	2	3	1.67
8	0	4	0	0	0	2	1.00
9	0	1	0	0	0	1	0.33
10	3	3	0	0	0	2	1.33
Score	25.00	40.00	5.00	0.00	25.00	60.00	25.83

Table 5.6: The system usability scale given after the test by group 2.

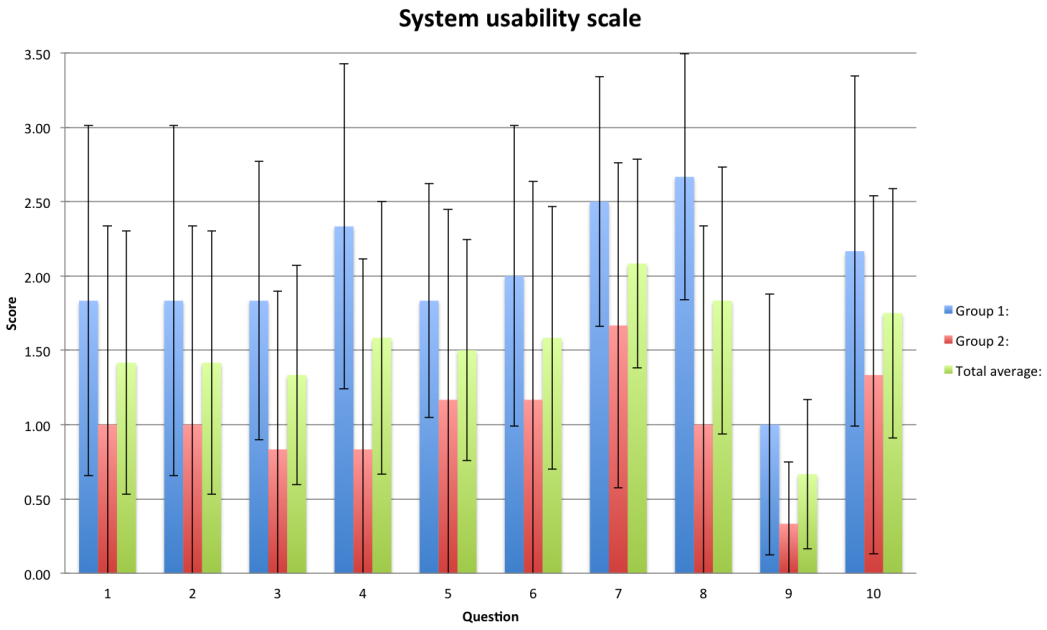


Figure 5.3: Results from the System Usability Scale.

The chart in Figure 5.3, shows that the two groups experienced the usability of the system differently and therefore scored Windows 8 differently. The confidence intervals for each of the groups are larger because of only six samples in the dataset. The confidence intervals on the total average are more accurate because of the larger sample size.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Score**
Total avg:	1.42	1.42	1.33	1.58	1.50	1.58	2.08	1.83	0.67	1.75	37.92

Table 5.7: Average from the System Usability Scale.

Combining the results from the two groups into a total average score, gives the score that shows the experience from using Windows 8 regardless of the sequence in which the usability tests were performed. The collected data shows that people who performed the tasks on Windows 8 before performing them on Windows 7, were not as satisfied with the usability as those that performed tasks on Windows 7 before Windows 8.

5.4.4 Specific problems

Table 5.8, 5.9, 5.10, 5.11, 5.12 and 5.13 document all the problems that were observed for the participants during the testing, the rate of occurrence and the severity of them. These problems have been organized together with problems affecting similar functionality of the operating system. The groups are based on where the problems physically appear in the system and are described in further detail below:

OS
: Problems that were caused by the operating system. The origin of the problems have no special placement and have therefore been put in this category. See the description of the problem for further information.

Charms/functions
: The problems in this category were related to the functions that can be found by moving the cursor to the different corners of the screen. Most of the problems in this category were related to the Charms[2].

Mail
: Problems in this category occurred in the mail application and were related to the functionality in it.

Browser
: These problems were uncovered in the browser, Internet Explorer.

[2]Charms are the functions appearing on the right of the screen. Figure 4.6 shows the location of the charms.

The problems have been marked with a grade of severity to determine how much they influenced the usability. The severity is divided into four levels[15]:

Critical(Level 1)
Prevents the participant from completing the task.

Significant(Level 2)
Generate significant problems when trying to complete the task.

Minor(Level 3)
Has minor effect on the usability of the system.

Not-essential(Level 4)
Enhancements to the system. When a participant states that "it would be nice to have".

In addition, the problems have been given a *cause-class* which is further described in Section 5.5.2.

ID	Group	Description	Number of occurrences	Severity	Cause-class
1	OS	When logging in(lock screen), information is lacking on what to do and the corners do not work as described in introduction. User end up clicking the connection icon.	10	Minor	A
2	OS	User icon on the start screen does not look like it is a button.	8	Significant	A, I
3	OS	Not able to close applications. Missing the close button in top right corner and it is therefore not easy to know if the applications are still running, or where the applications are located in the system.	7	Significant	B
4	OS	Shutdown is not in the start screen as it used to be.	7	Significant	C
5	OS	Not intuitive that users have to go to the desktop to find the file explorer when all the other applications are in the start screen.	6	Significant	D
6	OS	Not intuitive that power is located in the settings tab when using the Charms.	5	Significant	C

Table 5.8: Problems that occurred during testing, part 1.

ID	Group	Description	Number of occurrences	Severity	Cause-class
7	OS	Functions in the top right corner are not considered when looking for possible solutions. These functions are visible in the user interface, but still overlooked.	4	Significant	E
8	OS	Not able to escape applications with the "esc"-button.	4	Significant	C
9	OS	Not able to locate the desktop. Because user was stuck in full-screen application and thought that the active window had to be on top of the desktop like in Windows 7.	3	Critical	B
10	OS	Not understanding that the start screen is not a desktop.	3	Minor	B, C
11	OS	Double-clicking "tiles" in the start screen.	2	Minor	B
12	OS	Could not locate the local disk when in the file browser (same problem as some of the users had on Windows 7).	1	Critical	N
13	OS	"Too much clutter on the screens."	1	Minor	K

Table 5.9: Problems that occurred during testing, part 2.

ID	Group	Description	Number of occurrences	Severity	Cause-class
14	OS, Charms/ functions	Unable to enter the start screen because user did not understand how to use corners.	3	Critical	B, F
15	OS, Charms/ functions	"Missing a back button". Even though the upper left corner provides this functionality the user does not understand how to use it.	2	Minor	B, F, G
16	OS, Mail	Attachment file browser shows buttons that are inseparable from the text.	3	Significant	A, I
17	OS, Mail	When right-clicking, the buttons appearing at the bottom of the screen are overlooked because they are too far away from where the user did the click. The user therefore continues without seeing those buttons.	1	Critical	B, C, F

Table 5.10: Problems that occurred during testing, part 3.

ID	Group	Description	Number of occurrences	Severity	Cause-class
18	Charms/ functions	Hard to understand that the cursor has to be moved all the way into the corner for the Charms/hidden functions to appear.	5	Significant	B, F
19	Charms/ functions	The Charms-bar disappears too fast.	4	Minor	M
20	Charms/ functions	Uses the search functionality, but could not find any files on their local disk through this search. The search does not specify where the search is performed.	3	Significant	A
21	Charms/ functions	Lack description when they show up. Icons are not intuitive when they first appear.	2	Minor	A
22	Charms/ functions	Not easy to recognize the power button in the charm settings tab. It does not stand out enough from the other functions that it is grouped with.	2	Minor	H
23	Charms/ functions	User thinks the charms are a bit to slow when they appear.	1	Minor	M
24	Charms/ functions	"I do not like hidden menus."	1	Not-essential	F
25	Charms/ functions	"Do I actually have to move the cursor all the way into the corner? Why can't I just move it to the side when the menu appears from the side?"	1	Not-essential	J

Table 5.11: Problems that occurred during testing, part 4.

ID	Group	Description	Number of occurrences	Severity	Cause-class
26	Browser	Web page displayed after browser choice is confusing participants, making them think they are not in the browser. No confirmation of the browser choice.	6	Critical	L
27	Browser	Unable to locate the address bar in Internet Explorer.	4	Critical	F
28	Browser	User does not understand how to use the address bar and could not locate it.	1	Critical	M

Table 5.12: Problems that occurred during testing, part 5.

ID	Group	Description	Number of occurrences	Severity	Cause-class
29	Mail	Not intuitive to right-click when wanting to add an attachment.	7	Significant	F
30	Mail	Add content and subject in the mail is grayed out. Participants think they can not click it because it is "unavailable". Participants do not understand that they are text fields and not grayed out buttons.	5	Significant	A, I
31	Mail	Does not understand that the "+"-icon/button is used to create a new message. Lacking description.	4	Critical	A, I
32	Mail	Not able to locate the hidden buttons where they can add attachments. Information needed is hidden or unknown for the user.	3	Critical	F
33	Mail	"Send mail"-button lacking description and participants end up sending mail when they are looking for other functions.	3	Significant	I
34	Mail	Attachment file browser gives full-screen pop-ups that shows a "close"-button. This is done while the cancel button in the down right corner is still showing. This confuses the user because they try to cancel their action and not close the application.	1	Significant	D
35	Mail	"Childish mail client."	1	Not-essential	E

Table 5.13: Problems that occurred during testing, part 6.

The summary of number of errors in each severity is given in Table 5.14.

Severity	**Count**	**Percentage of total errors**
Critical	9	25.7 %
Significant	14	40 %
Minor	9	25.7 %
Not-essential	3	8.6 %

Table 5.14: Summarizing the number of errors in each severity.

The largest percentage of errors, are in the category that creates significant problems to the users, while 25.7 percent are in the category critical. Some of these errors have only occurred once and are therefor not considered as important as problems occurring often or that have high severity. By plotting the severity and number of occurrences into a chart, it become easier to see which problems that have the largest impact on usability.

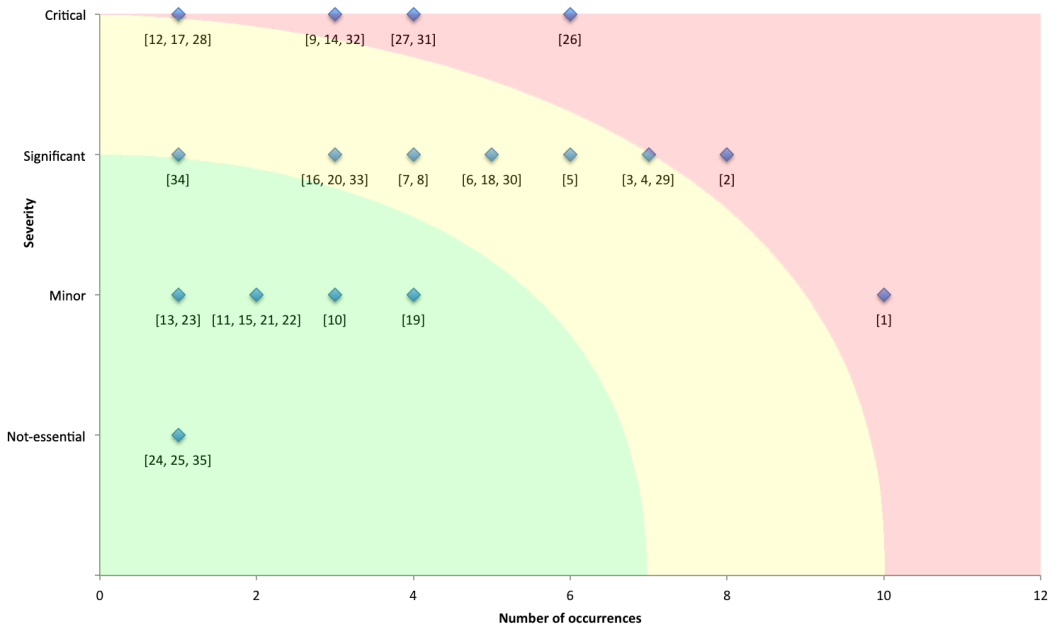


Figure 5.4: Plot of the problems. The ID of the problem is used as label for each plot.

Figure 5.4 shows three different zones. The red zone is for problems that affect usability the most, while yellow problems have reduced impact on the usability, and green problems do not cause any large effect on usability. These zones were inspired by the zones used for risk matrices[20]. The zones could have been moved in any direction, but were placed where the severity or number of occurrences greatly influenced the usability of the system.

5.5 Analyzing the Results

5.5.1 Difference in execution time

By using a T-test to check if the execution times are overlapping, it will be possible to determine the chance of error from the results given by the datasets. Using a hypothesis where the two datasets are overlapping each other.

Task	**All**	**Without fails or assists**
1	0.004	0.049
2	0.081	0.018
3	0.000	0.004
4	0.211	0.041
5	0.000	0.014
Total	0.000	0.000

Table 5.15: p-value when performing a paired, two-tailed, T-test.

With the significance level of $\alpha = 0.05$, and considering all execution times, it can be seen that task 2 and 4 are not significant. By removing execution times to tests where the participant failed, or got assistance and focusing on those that finished the task on their own, the results shows that the datasets are significantly different from each other.

5.5.2 What causes the specific problems?

Looking at the problems listed in Section 5.4.4, they can be broken down into different groups based on what design principle that failed to be satisfied.

Class	Cause	Problem ID
A	Too little information/ too minimal	1, 2, 16, 20, 21, 30, 31
B	Touch-related	3, 9, 10, 11, 14, 15, 17, 18
C	Changes to workflow	4, 6, 8, 10, 17
D	Inconsistency	5, 34
E	Adapting design to users	7, 35
F	Hidden functions	14, 15, 17, 18, 24, 27, 29, 32
G	User in charge	15
H	Gestalt principles	22
I	Affordance	2, 16, 30, 31, 33
J	Profit vs pain	25
K	Minimalism	13
L	System status	26
M	Uncategorized	19, 23, 28
N	Errors by test	12

Table 5.16: Problems grouped by their causes. The underlined problem IDs, are those in the red zone from Figure 5.4.

Table 5.16 was created by using the principles from Chapter 4. Some of the problems occur in more than one row. That is because the occurrence of the problem has roots in more than one of the causes.

Chapter 6

Interviews

This chapter will describe the interviews conducted in this study. It will briefly describe how the interviews were performed and the results from them.

6.1 Semi-Structured Interview

Semi-structured interviews are interviews where the interviewer has an incomplete script[30]. This script contains the most vital parts that the interviewer has planed to confront the interviewee with. This requires some improvisation from the interviewer and knowledge of problems and pitfalls that can affect the results.

By using a semi-structured interview, it allows for a deep and informal conversation about the subject. Semi-structured interviews allow for the interviewer and the person beeing interviewed to jump back and forth between subjects and will uncover more information than a fully scripted interview[12]. It is important to record the whole interview, because this will become necessary in the transcription process.

After the interview is complete, the interview is transcribed for further work. To maintain validity and to not loose the context, the transcription process follows the 14 stages described by P. Burnard[5].

6.2 Execution

The purpose of the interviews in this study was to gather information on earlier migration processes and to identify the problems they have encountered. The information from these interviews are then used as recommendations for organizations deciding to introduce Windows 8 at their workplace.

In this study there was performed two semi-structured interviews with administrators involved in the migration processes from one operating system to another. They were selected because of their position and involvement in those processes. They had both been involved in earlier migration processes as well as evaluating the need for Windows 8 in their organizations.

A meeting room was arranged for both of the interviews. There was no time pressure allowing for a relaxed atmosphere around the interview. A dictation device was used to record the interview. The interview was then run by one person using the script and guiding the interviewee towards the information needed for this study.

The script for the interview and the transcription of the most interesting parts can be found in Appendix J. The transcripts are in Norwegian because the interviews were performed in the native language of the participants. The results from these interviews are then described in Section 6.3. The twelve first stages of the method created by P. Burnard[5] have been followed. For convenience, a summary of how these stages were performed in this study are listed on the next page.

Stages of the analysis:

1. After the interview, notes were made for each of the topics discussed.
2. Transcripts were read through and notes were made.
3. Reading through the transcript a second time writing down categories to describe all the aspects of the content.
4. Collapsing some of the categories into broader categories. This was done to reduce the number of categories.
5. Categories that were similar or repetitious were removed.
6. One colleague was invited to independently create categories of the transcript. This was compared with the original list of categories and merged into one list to increase the validity of the categories.
7. Transcripts were read through together with the final list of categories to ensure that all the aspects of the content were included in the list.
8. Both of the transcripts were then worked through marking the content to their appropriate category.
9. Each coded section is then cut out ensuring that the context of the section is maintained.
10. The cut out sections are then put under each of the appropriate categories.
11. Transcripts, sections and categories were then sent to the interviewees to validate their authenticity.
12. All sections are filed together for direct reference[J] and recordings of the interviews have been stored.

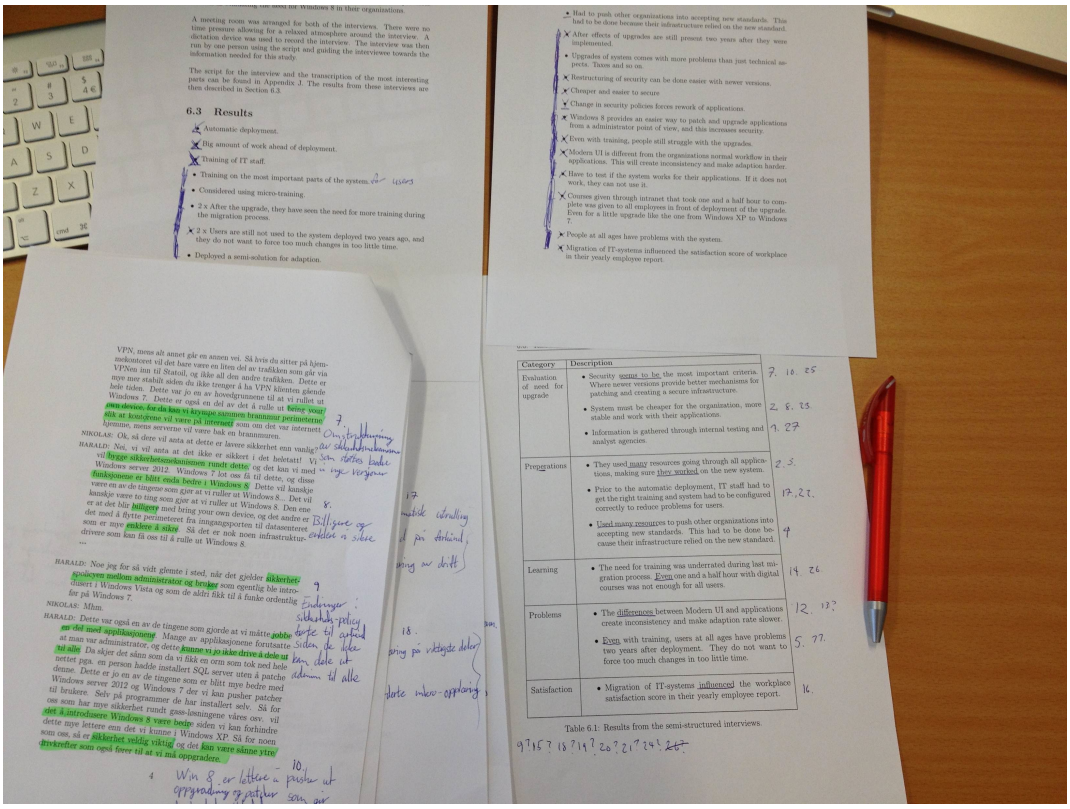


Figure 6.1: From the transcription process.

The thirteenth and fourteenth stage of the transcription method by P. Burnard[5] was not performed since they were not consistent with the research method of this study. Further evaluation of the results from the interviews were compared with results from the usability testing. Section 7.1.3 contain these evaluations.

6.3 Results

Table 6.1 has been made by following the first twelve stages in the method made by P. Burnard[5] described in Section 6.2. The results have then been grouped together in categories and the description is made from the line citations of the transcripts in Appendix J.

Category	Description
Evaluation of need for upgrade	• Security is one of the most important criteria. Where newer versions provide better mechanisms for patching and creating a secure infrastructure.[J.112, J.264, J.298] • System must be cheaper for the organization, more stable and work with their applications.[J.86, J.161, J.276] • Relevant information is gathered through internal testing and analyst agencies.[J.128, J.142]
Preparations	• All applications used in the organization had to be tested on the new operating system making sure they worked properly.[J.161, J.188] • Prior to the automatic deployment, IT staff had to get the right training and operating system had to be configured correctly to reduce problems for users.[J.18, J.75] • Put pressure on other organizations into accepting new standards. This had to be done because their infrastructure relied on the new standard.[J.195]
Training	• The need for training was underrated during last migration process. One and a half hour with digital courses were not enough for all users.[J.118, J.355]
Problems	• The Windows UI and the applications developed by the organization are designed differently and can create inconsistency between the user interfaces making the rate of adaption lower.[J.346] • There are users at all ages that completed training that still have problems two years after deployment. They do not want to introduce too many changes in too short time.[J.210, J.319]
Satisfaction	• Migration of IT-systems negatively influenced the workplace satisfaction score in their yearly employee report.[J.374]

Table 6.1: Results from the semi-structured interviews with line number of citation from Appendix J.

Chapter 7

Discussion

This chapter will go through the findings from this study and summarize the results to answer the research questions from Section 1.3.

7.1 Evaluation

This study is an evaluation of the usability of Windows 8. To address the research questions both usability testing and semi-structured interviews have been performed. The usability testing uncovered several problems that influenced the user experience, while the interviews helped by getting a better understanding of the migration process from one operating system to another. Evaluation of these results will answer the research questions.

7.1.1 Research question 1

How the usability in Windows 8 is, when using screen, mouse and keyboard for users with experience from Windows 7, has to be deduced from all the elements described in Section 5.4. The results from the usability testing conducted in this study showed a negative trend for the three different elements when testing Windows 8.

Effectiveness

The effectiveness describes how well the user could perform without having to ask other people for help. The completion rate can be used to anticipate the amount of people that need assistance when using the new operating system.

Task	**1**	**2**	**3**	**4**	**5**	**Total**	**Compl. all**
Win. 7	100%	100%	100%	83%	100%	97%	83%
Win. 8	75%	92%	50%	83%	58%	72%	33%

Table 7.1: Completion rate for the two operating systems as presented in Table 5.1.

Task 1: Check a news web page in a browser of the participants choice.

Task 2: Find a file on the desktop.

Task 3: Send this file by e-mail.

Task 4: Find a presentation stored on the local disk.

Task 5: Turn off the computer before leaving the desk.

Except for task 4, Windows 8 proved to have a lower success rate than on Windows 7. As seen in Table 7.1, only 33% of the participants completed the tasks on their own in Windows 8, while 83% completed them on Windows 7.

The two participants who did not complete all the tasks on Windows 7 had problems with recognizing their local disk because the letter was "C" instead of "M". While on Windows 8, there was only one participant who could not find the local hard drive because of the same problem. That leaves Windows 8 with one participant that could not complete the task because of problems created by the operating system. This error was not a result from the poorly configured test setup, that did not mach the exact setup at the participants workplace. As a result of this, Windows 8 performed worse on task 4 when removing the incorrect data created by the test environment.

The lowest completion rate was on task 3. Only half of the participants were able to send an email in Windows 8 on their own. Task 1 and 5 gave more problems than task 2 and 4. That is because the functions or applications from task 1 and 5 have been redesign and were unfamiliar to the user, while those on task 2 and 4 were fairly similar to how they worked on Windows 7. The problems with task 4 where the letters were changed, showed that even small changes could confuse the users of a system.

Efficiency

The efficiency can be used to anticipate the extra time it will take an user to complete tasks on the new operating system. Results from the usability tests are just an indication of what to expect after an upgrade from Windows 7 to Windows 8. The composition of an user group can be different from the one being tested in this study, and that will have effect on the impact created by the migration to Windows 8.

The participants and the execution times of tasks that they completed, without getting assistance or failing to complete the tasks, shows a large difference in Windows 7 compared to Windows 8. From the T-test performed on the dataset in Section 5.5.1, the results states that the completion times are significantly different from each other for tasks performed on each of the operating systems.

A trend that was observed during the usability testing, was that participants with knowledge of shortcuts through key combinations did have an advantage. They could find functions faster than participants that had to use the graphical user interface. This helped them to navigate through the system easier and to explore functionality faster.

Figure 5.2 displays the differences with the 95% confidence interval on each of the tasks from both Windows 7 and 8. Table 5.4 shows the differences between the mean execution times on each of the operating systems. Depending on the task, users will use more than 200% more time on Windows 8 the first time they perform it. Users adapt to new interfaces over time, but over 200% more time spent on performing it will most likely create frustration and users might have to postpone other tasks. The efficiency will be influenced by this and the users will be less productive until fully adapted to the new user interface and workflow.

Satisfaction

Satisfaction describes how good the user experience has been for the user. System Usability Scale has been used to evaluate the user experience. This is a subjective score that each of the users give after they have performed the test.

Results from System Usability Scales in this study gives variable results, ranging from 0 to 77.50. The two different groups gave different scores where group 1 was more positive than group 2. The differences can be seen in Figure 5.3. Trends from their scores show that the participants felt insecure while using the system. On the other hand, the participants were more positive when scoring how quickly people would learn to use Windows 8. The largest difference between the two groups was on question number 8. Group 1 scored the system as 2.67, while group 2 scored it to 1.00. The significant difference in these scores may be due to the three participants from group 1 that managed to completed all tasks, while on group 2 there was only one participant that completed it without failing on any tasks or without getting any assistance.

When comparing the average scores with the scale made by Bangor et al.[2], the average score of 37.92 is rated as a poor usability performance. If the average score is compared with the acceptability range, the average score from the SUS-forms rates the usability of the system as not acceptable. Using the grade scale from Bangor et al, 10 of the participants graded the system as a F, while only two participants gave the system the grade C. This is not a very impressing result for such an important system. Users should not have to deal with a bad user experience for every task they have to perform on a computer. Also the satisfaction criteria might only create frustration and postponement of work. Despite of the low scores, some of the participants experienced Windows 8 as exciting to use. These participants said they liked the new Windows User Interface with the abundant use of colors.

7.1.2 Research question 2

The aspects in Windows 8 that caused usability problems had many different roots. Table 5.16 shows the different reasons why problems occurred. Some of those problems made a greater impact than others on the usability. Figure 5.4 shows the impact of these in three different zones and can be used to analyze the problems and to provide an answer for research question 2. By going into details on some of the problems in the red zone and combining the knowledge from Chapter 4, it will provide a better understanding of what causes the usability problems in Windows 8.

ID: 1 - Cause-class: A

Figure 7.1: Lacking options and information.

The login screen presented to the users contains very little information as seen in Figure 7.1. There are no indications of what the users have to do in order to gain access to the system. Eventually all the users clicked the networking icon in the bottom left corner and thought that this was the right action in order to continue. What the users did not realize, was that they could click anywhere on the screen in order to proceed. The information explaining this to the user was lacking. Minimalism is important to remove unnecessary information as described in Section 4.1.3. In this case it looks like too much of the information has been removed, and users are not given the assistance needed to perform their task.

ID: 2 - Cause-class: A, I

(a) This is a button. (b) This is not a button.

Figure 7.2: Too little information confuses users.

When users navigated through the system they had problems distinguishing the elements that were text from the elements that were functions. Figure 7.2 shows two very similar looking elements where one is a button and the other one is just text. There is nothing that separates them from each other. When participants realized that the icon was a button they stated that they did not think it was a button at first.

Figure 7.2a do not show the affordance of a button, and is therefore overlooked as a function. This is only one of the examples in the system where there is insufficient information or the functions are lacking affordance as described in Section 4.1.3 and 4.1.2.

ID: 4 - Cause-class: C

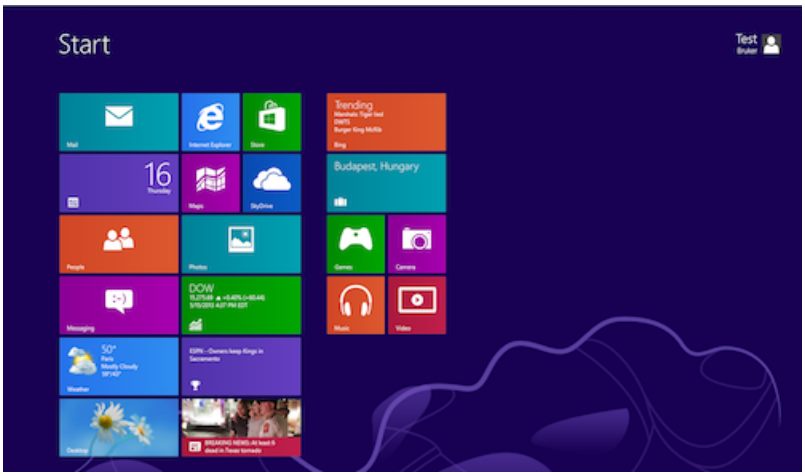
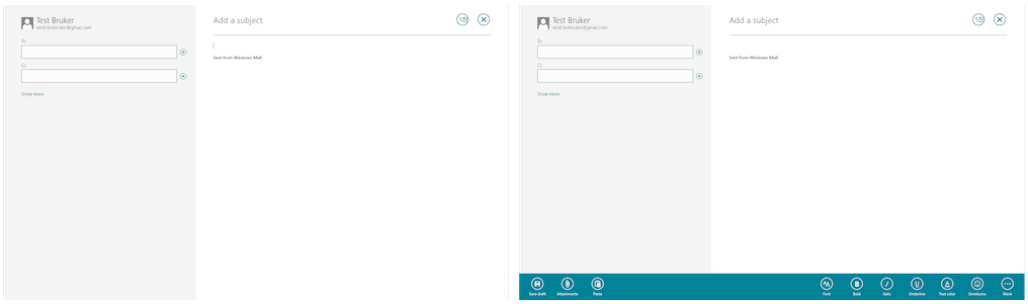


Figure 7.3: The shutdown button is not in this menu.

Users had problems when they tried to turn off the system. They were used to having the shutdown functionality in the start menu. That functionality had been moved, and that resulted in a search for the new placement of it. Most of them did not understand where to start and some of the users said that they would eventually just use the physical power-button on the computer.

Functions have been re-arranged in the new operating system and users with well established habits struggle to find the new location of functions. The only users that did not have a problem with this were the users that knew of the shortcuts through key combinations. For the rest of the users, the effort of adapting to the new system is large and the amount of effort they had to invest in order to complete their task, could not be justified by the modest advantages the system added. As described in Section 4.2.4, users that had these problems ended up being frustrated and they did not have a good user experience.

ID: 17 and 29 - Cause-class: B, C, F



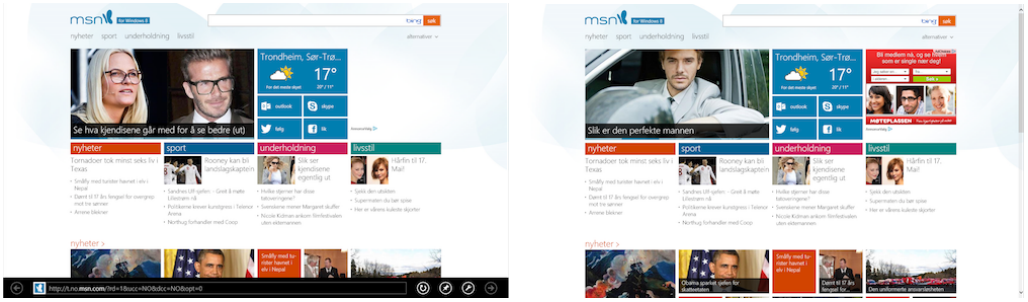
(a) Without application bar. (b) With application bar.

Figure 7.4: The application bar on the bottom of the screen is overseen.

The application bar shown in Figure 7.4b was not considered by users when trying to complete their tasks. Hiding functionality makes it hard for users to complete their tasks. Combining this with the modifications to the workflow, made one of the users oversee the application bar. The user was used to having the context-menu appearing when right-clicking, not the bar at the bottom. Therefore when the user was looking for the context-menu, the user overlooked the bar appearing on the bottom.

Described in Section 4.2.4, the operating system must not rely on the users memory. Functionality important to perform a task, need to be displayed in the user interface. This problem occurred at several occasions when users were looking for functionality and it was hidden in some of the bars around the edges. This is described in the Windows 8 user experience guidelines, and they want the developers to add functions to the hidden menus in the applications they develop[28].

ID: 27 - Cause-class: F



(a) When browsing new page. (b) After clicking or scrolling.

Figure 7.5: The address bar disappears and can not be located by user.

Participants were not able to locate the address bar in Internet Explorer. This was very similar to the problems users had with the application bar. The address bar was not at the top like it used to be, and users did not think of the hidden functions. At first when they entered the browser, the address bar was visible, but when they either clicked on something or scrolled it disappeared. Because of all the information already displayed on the web page, they did not take notice of the bar at the bottom and could not locate it when they tried to look for it.

The same effect as with problem 17 and 19 occurred, when users did not think of the hidden features. Not being able to locate the needed functions were a big problem for many of the users. In the Windows 8 user experience guidelines it is stated that functionality important for performing the task, must not be hidden in the application bar, but displayed at all time[28]. It seems like this has been forgotten in many of the applications that comes with Windows 8 by default.

ID: 26 - Cause-class: L

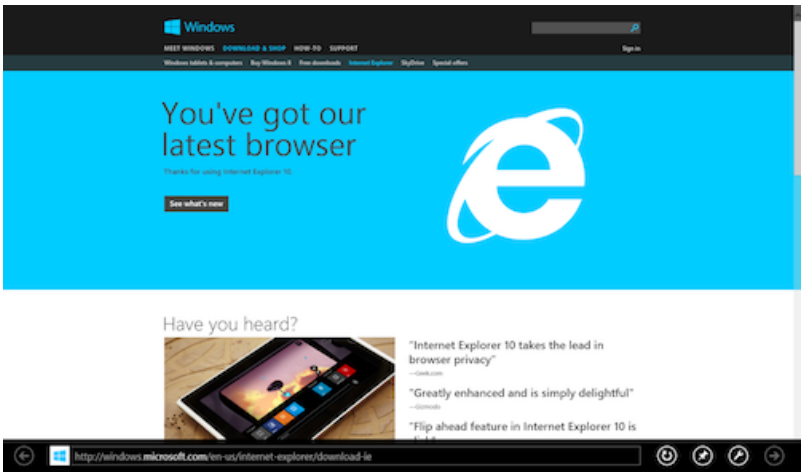
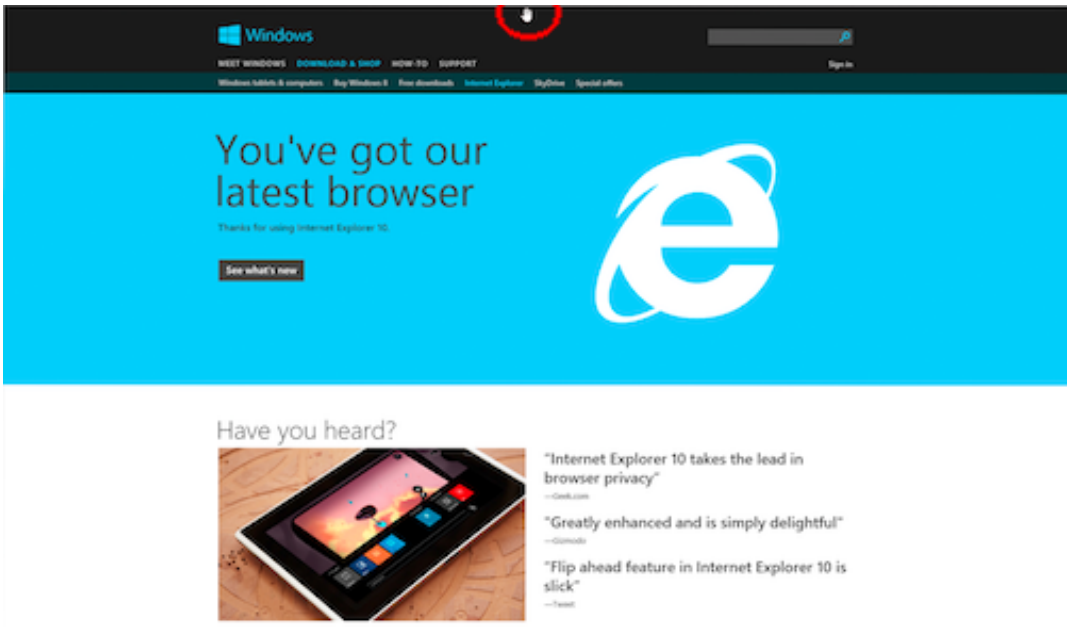


Figure 7.6: No visible indication of the system status.

This problem occurred on several different tasks, but gave the most impact when the users tried to enter Internet Explorer. The users were first presented the "browser choice"-dialog where they selected Internet Explorer. Then they were presented the page seen in Figure 7.6 saying "You've got our latest browser". What happened next was that 50% of the users said: "Okay then... Now I just need to locate Internet Explorer and open it!". The users did not realize that they were in the browser! They were lacking feedback from the system telling them what application that was running on their screen. As seen in Figure 7.6, there is nothing except from the bar at the bottom explaining the users that they are in a browser, and as seen in Figure 7.4a there is nothing in the mail application making the users comprehend what they currently are doing.

Letting the users feel like they are in control increases the usability. By giving them information about the system and what they currently have on their screen, it improves their understanding of how navigation works and users get confirmation on the choices they perform. The eight golden rules[38] described in Section 4.2.2, explains that the system provides useful feedback to the users at all time to increase the users awareness of the state of the system.

ID: 3 - Cause-class: B



(a) Cursor grabing the top of an application.

(b) Cursor has been dragged down to the bottom with the application

Figure 7.7: Closing an application.

Users had difficulty understanding how they closed an application, and they said they were missing the button to close the window in the top right corner of the window. They struggled to understand whether the applications were closed or not when they went back to the start menu, or if it was running in the background. Figure 7.7 shows one of the methods that can be used when closing an application. The user has to move the cursor to the top of the screen, then click and hold the left mouse button and drag it all the way to the bottom before releasing the button.

This is one of the examples where touch gestures have inspired the workflow in Windows 8. It works perfectly well on a tablet with a small screen, where the users can slide their hand from the top of the screen to the bottom. On a big screen with a mouse or a touch pad, it becomes a badly designed function. The user has to move the cursor to the top, click and drag, then lift the mouse or finger from the touch pad, then slide it to the bottom again until the cursor reach the bottom and they can finally release the button. For those with low cursor speed or a small touch pad, they might have to lift their finger or mouse more than once. It is a very demanding task to perform, and it is frustrating to perform this when the users were used to clicking the cross in the top right corner. The problem here is that the intended way to do this, does not work equally well on a touch device compared to a keyboard and mouse setup. The interface should have been designed differently and more intuitive and effective for the users with a keyboard and mouse setup.

ID: 9 - Cause-class: B

This problem is really not about touch gestures or the interaction with touch, it is more about the concept of how touch interfaces are designed. From earlier desktop user interfaces, the users have been used to a window-based model, where applications ran in their own windows on a desktop and they could be placed where the user wanted to. In Windows 8 the applications cover the whole screen and are run separate from each other. Figure 7.8 shows an example of how this is done. Users use the start screen(in the center) to open applications, then they need to go back to the start screen to revisit the other applications they have running. There are other ways to navigate, but those are mostly used by people with high knowledge of computers.

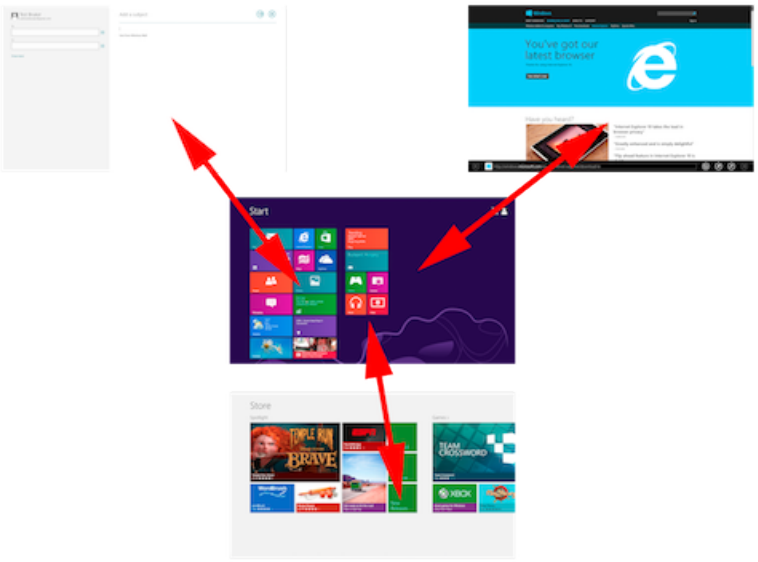


Figure 7.8: Navigation in Windows 8.

Users with little knowledge of these type of systems struggle to understand how to navigate through the system. The fundamentals of the window-based applications are gone, except from on the desktop, and they need to adapt to a new concept where they understand how this works. The lack of information does not help the users to understand this, and having the old desktop there does not contribute to the cause. The mixing of these two concepts need to be designed in such a way that the users can understand and feel like they are in control of the system[38].

7.1.3 Research question 3

Section 5.4 and 6.3 gives some pin-pointers toward recommendations for organizations. The different organizations evaluated the need for an upgrade differently and both of them did not value usability that high in their decision process. On the other hand, both organizations saw that users were not given enough training and that they were still struggling with the system deployed two years back.

One of the organizations made their employees perform digital courses that took about one and a half hour to complete, and the other one had considered performing micro-training[1] through a third party company.

Migration processes influences the satisfaction to the employees current work situation. Without proper training this will most likely create frustration and irritation. As one of the organizations stated and seen from the results in the usability tests[5.4.3], introducing Windows 8 without providing the employees with help options or training, will in general not be received positively by them. Some individuals will like the new system better than the old system, but most of them will react in a negative way.

Giving the users different options for training is the best way to handle this migration. Providing different types of courses for user groups with different qualifications. The user groups with high knowledge of computers can do digital training, while the user groups that have difficulties adapting to new IT solutions can be trained through a "hands on"-course or a workshop where they can participate in task execution. The courses should cover the mandatory topics listed in Table 7.2 as a minimum. The optional topics can be left out if the organization uses other software for these functions. A five minute course will not be enough for all users, therefore the topics have been provided with a recommendation for time used on that specific topic in a hands on situation. These times are based on observations made in the usability tests. One column is for novice users, while the other column is for experienced users.

[1] Short learning sessions that can be combined into series to educate people on different topics.

Mandatory Subjects	**Time in minutes for novice users**	**Time in minutes for expert users**
How to log in.	5	2
The mental model of the operating system. How the navigation between applications work and how this works compared to the old desktop.	15	4
Using the corners and the windows button on the keyboard.	15	3
Start menu and configuring the tiles.	10	3
Charms and all the functionality of them.	15	4
Right clicking and the hidden functions that can be located at the application bar.	15	4
How to close applications and how to switch between them.	10	3
Where to find the file explorer and how to use it.	10	2
The user icon. Users being able to change their password, lock the computer, change profile picture.	15	3
How to turn off the computer.	10	2
Total time:	120	30

Table 7.2: The important topics that creates most usability problems[5.16] and that are listed as mandatory.

Optional Subject	**Time in minutes for novice users**	**Time in minutes for expert users**
Internet explorer and how to use the browser.	15	3
Mail client and how to send email with attachments.	20	4

Table 7.3: The optional topics that creates usability problems[5.16].

Creating an information brochure that summarizes the topics in Table 7.2 and 7.3, would recall their memory when the users are back with their computers. That way there would reduce the load for the support staff, and this could help the users with completing the most basic tasks on their own.

7.2 Threats to Validity

During this study there has been performed usability testing according to ISO/IEC 25061:2006[13] and semi-structured interviews of administrators involved in the migration processes from one operating system to another. The usability tests were performed on 12 participants divided into two groups. These two groups performed the tests in different sequences. To uncover more usability problems and to get a greater sample size, the tests should have been performed on more users. More users would have provided data that can show trends more clearly. Due to time and cost limitations, it was not possible to perform more tests in this study.

The sample size in the efficiency results without fails or assists[5.4.2] was down to six samples for task three. A non-parametric method was considered, but despite of the small sample size a T-test was performed. The Mann-Whitney U test would have been performed if the sample size was lower than the six samples.

Processing of the data and interpreting the video from the usability tests were performed by the same person. This could lead to overlooking important observations discovered by this study creating invalid data. Performing more usability tests with different test leaders would give stronger and more reliable results. Different results would resolve in uncovering more usability problems and contributing in the calculation of the effect an upgrade to Windows 8 would make on efficiency, effectiveness and satisfaction.

"System Usability Scale"-forms were filled out after tasks on both systems were completed. The participants were told to rate their experience of Windows 8. Separate forms could have been given to the participants after performing tasks on each of the systems. By scoring both systems these results could have been used to compare the two operating systems and would have provided information towards what system the users preferred.

The transcripts of the interviews were evaluated by two persons. More people analyzing the data would have resulted in a higher validity of the categories and the description of them. Interviewing more administrators would have provided more data on how the migration processes are executed.

Transcripts have been written by one person. Having more people validate the transcripts with the recordings would exclude more errors from the transcripts. Transcripts were sent to interviewees to validate the citations. Following the transcription processing method by P. Burnard[5] more strictly, could have strengthened the validity of the results made from the interviews by using a well documented method.

In Section 3.1 the context of use has been described as:

> *Users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used.*

To evaluate threats that could have made an impact during the usability testing, the tests have to be seen in light of the context of use.

Users

Participants were selected based on their experience with computers, their working title and that they never had seen Windows 8 before. Participants with different computer experience could generate other results. Using other age or gender distributions could end with results different from the results in this study. Participants were selected from NTNU in the Trondheim area. Selecting participants from other organizations and from other parts of Norway would strengthen the results when comparing with the Norwegian population. The average age of the participants in this study was 47 years, which is considered as high. This study has judged the average age of participants to be a fair representation of the population in these professions based on observations from the environment at the participants workplace.

Tasks

Tasks given to the users were based on one persons observations in the participants workplaces. Gathering information to make these tasks could have been performed as a field study to collect more accurate data on the users daily tasks. Designing the tasks differently would have tested other parts and applications in Windows 8. Testing other parts of the system could have given other results. There were no observations of problems regarding the description or formulation of the tasks during the usability testing.

Equipment

The test computer was not configured exactly like the setup on the participants private computer. Some differences were discovered during testing, but all except from one of these differences was considered to be negligible and that they had no impact on the results. The one thing that resulted in incorrect data was that the local disk letter was different from what the users were used to. By doing a comparison of the test computer and their daily work computer, that error would have be eliminated before performing the usability tests. Preferably the tests could have been performed on the participants computer in their personal office to create the most accurate results, but that was not possible in this study due to security policies at NTNU.

Environment

Usability tests were performed by following the checklist in ISO/IEC 25061: 2006[13]. There was only one person performing the tests, but the checklist was followed down to every detail to assure that it was executed according to the standard. During the tests there was a good atmosphere between the participants and the test leader. The users enjoyed exploring and performing the tasks on both of the systems. Users were believed to be honest about their opinions toward Windows 8, but this can be criticized as opinions influenced by the test situation. For instance by being more positive towards the operating system due to the good atmosphere.

The usability lab simulates an office to a certain degree. It is possible to recreate the physical environment, but not the social environment. This study has tried to recreated an office, which it did, except from the individual pictures and notes that a normal employee has laying around in their office. The social environment of the lab was not the same as the participants are used to. The lab is more isolated and it does not include interrupts or noise by colleagues casually dropping by the office or the phone ringing. Without these disturbances it is less stressful to be in the lab. On the other hand the lab creates stress in another way. The focus on the tasks and the idea of being in a lab is more stressful because it is both less familiar and users might feel that they have to perform their best. All these factors influences the results generated by the usability tests. The results could also be affected by the participants being more reserved than normal in a foreign office.

7.3 Reflections on the study

After conducting this study, there are elements that should have been performed to improve the research method. A field study of an ongoing migration process from one operating system to another, would have provided valuable insight on how the process is executed. What elements users struggle with and how these are handled. It would also provide data on what preparations the users are exposed to in order to prepare the users for the migration.

An interview with a representative from Microsoft would provide feedback on why design choices were made, and what they have done to create the user interface presented in Windows 8. This information would have been valuable to understand the intentions behind the design and how they have tested the user interface on their users. A Microsoft representative would also give information on how they recommend to execute the migration to Windows 8 in an organization.

Chapter 8

Conclusions

This study uncovered usability problems that users experienced after upgrading the operating system from Windows 7 to Windows 8. Usability tests performed, gave negative results in categories of effectiveness, efficiency and satisfaction. The severity of some of the problems that were uncovered, was high and affected the usability of the system.

Usability problems that occurred in the usability testing of Windows 8, were caused by fundamental Human Computer Interaction mistakes made by the design, and from an introduction of a new mental model for navigation. Some of the participants from the usability tests experienced Windows 8 as exciting to use and liked the new Windows User Interface with the abundant use of colors. Hidden functions proved to be one of the larger problems for users when executing their daily tasks. The new functionalities were not a part of their established workflow, and were often forgotten. Touch inspired elements were poorly designed in the operating system for use with keyboard and mouse, making Windows 8 harder to use than Windows 7 for the tasks in this study.

The interviews of the administrators uncovered experiences made from earlier migration processes. The lack of training provided to the employees when upgrading to a new operating system in organizations affected the employees negatively in their daily work. The introduction of new IT solutions to the employees resulted in a lower satisfaction score on the yearly employee report. On the other hand, Windows 8 provided technical benefits such as better security and an easier way to patch software installed on the users computers.

Providing the right amount of training and support to the employees would assist the adaption process and make the transition easier for the employees. A table with topics that the users struggled with, and that the users need training in, have been proposed. This table also include an estimated time usage for the training.

8.1 Further Work

During this study there has been discovered research topics that need further research to fully understand them:

- Research on designing interfaces that both work well with touch input and with mouse and keyboard input. Could problems discovered in Section 5.4.4 been better designed to work with both input types?
- Closer examination of how to prepare employees in organizations for an upgrade of the IT solutions at their workplace. What kind of training is needed and what is the best way to perform this training to make the migration work without creating too much impact on their work situation?
- Field study of migration processes focusing on usability problems that are created when users are presented a new operating system. How can this process be improved for the users and what is the best way to provide support to the users?

Bibliography

[1] Chadia Abras, Diane Maloney-krichmar, and Jenny Preece. User-centered design. In *In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*. Publications, 2004.

[2] A. Bangor, P. Kortum, and J. Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[3] Alan Borning and Robert Duisberg. Constraint-based tools for building user interfaces. *ACM Transactions on Graphics (TOG)*, 5(4):345–374, 1986.

[4] J. Brooke. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189:194, 1996.

[5] Philip Burnard. A method of analysing interview transcripts in qualitative research. *Nurse education today*, 11(6):461–466, 1991.

[6] John M Carroll. The nurnberg funnel: designing minimalist instruction for practical computer skill. 1990.

[7] John M Carroll and Society for Technical Communicat. *Minimalism beyond the Nurnberg funnel*. The MIT Press, 1998.

[8] John M Carroll, Penny L Smith-Kerker, James R Ford, and Sandra A Mazur-Rimetz. The minimal manual. *Human-Computer Interaction*, 3(2):123–153, 1987.

[9] Loyd Case. Windows 8: The official review. `http://www.pcworld.com/article/2012830/windows-8-the-official-review.html`, 2012. [Online; accessed 26-May-2013].

[10] M.D. Corry, T.W. Frick, and L. Hansen. User-centered design and usability testing of a web site: An illustrative case study. *Educational Technology Research and Development*, 45(4):65–76, 1997.

[11] Michael deAgonia. Apple os x mountain lion review: ios-like features help unify your digital world. http://www.computerworld.com/s/article/9229339/Apple_OS_X_Mountain_Lion_review_iOS_like_features_help_unify_your_digital_world, 2012. [Online; accessed 04-May-2013].

[12] Barbara DiCicco-Bloom and Benjamin F Crabtree. The qualitative research interview. *Medical education*, 40(4):314–321, 2006.

[13] ISO DIS. Iso/iec 25062:2006 software engineering-software product quality requirements and evaluation (square)-common industry format (cif) for usability test reports. *International Organization for Standardization (ISO). Switzerland*, 2006.

[14] ISO DIS. 9241-210: 2008. ergonomics of human system interaction-part 210: Human-centred design for interactive systems (formerly known as 13407). *International Organization for Standardization (ISO). Switzerland*, 2010.

[15] J.S. Dumas and J.C. Redish. *A practical guide to usability testing*. Intellect Limited, 1999.

[16] Tony Fernandes. Global user interface design. In *Conference companion on Human factors in computing systems*, pages 363–364. ACM, 1995.

[17] Clifton Forlines, Daniel Wigdor, Chia Shen, and Ravin Balakrishnan. Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 647–656. ACM, 2007.

[18] Gartner. Gartner says declining worldwide pc shipments in fourth quarter of 2012 signal structural shift of pc market. http://www.gartner.com/newsroom/id/2301715, 2013. [Online; accessed 26-May-2013].

[19] Gartner. Gartner says worldwide pc, tablet and mobile phone combined shipments to reach 2.4 billion units in 2013. http://www.gartner.com/newsroom/id/2408515, 2013. [Online; accessed 04-May-2013].

[20] PF Garvey and Zachary F Lansdowne. Risk matrix: an approach for identifying, assessing, and ranking program risk. *Air Force journal of logistics: vol22_no1*, page 18, 2002.

[21] A.M. Gåsland. *Metodiske utfordringer med brukbarhetstesting av digitale læringsplattformer*. PhD thesis, Norwegian University of Science and Technology, 2012.

[22] Preston Gralla. Windows 8 review: Still a two-headed beast. `http://www.computerworld.com/s/article/9230287/Windows_8_review_Still_a_two_headed_beast`, 2012. [Online; accessed 04-May-2013].

[23] Libby Hanna, Kirsten Risden, and Kirsten Alexander. Guidelines for usability testing with children. *interactions*, 4(5):9–14, 1997.

[24] Scott Jenson. *The Simplicity Shift: Innovative design tactics in a corporate world*. Cambridge University Press, 2002.

[25] Charlie Kindel. Windows phone 7 series ui design and interaction guide. `http://blogs.windows.com/windows_phone/b/wpdev/archive/2010/03/18/windows-phone-7-series-ui-design-amp-interaction-guide.aspx`, 2010. [Online; accessed 26-May-2013].

[26] Brandon LeBlanc. Windows keeps getting better. `http://blogs.windows.com/windows/b/bloggingwindows/archive/2013/05/14/windows-keeps-getting-better.aspx`, 2013. [Online; accessed 26-May-2013].

[27] Joanna McGrenere and Wayne Ho. Affordances: Clarifying and evolving a concept. In *Graphics Interface*, pages 179–186, 2000.

[28] Microsoft. Windows 8 user experience guidelines, August 2012.

[29] R. Molich. *Usable web design*. Nyt Teknisk Forlag, 2008.

[30] M.D. Myers and M. Newman. The qualitative interview in is research: Examining the craft. *Information and organization*, 17(1):2–26, 2007.

[31] J. Nielsen and J.A.T. Hackos. *Usability engineering*, volume 125184069. Academic press San Diego, 1993.

[32] Jakob Nielsen. Why you only need to test with 5 users. `http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/`, 2000. [Online; accessed 28-May-2013].

[33] Jakob Nielsen. F-shaped pattern for reading web content. *Retrieved January*, 18:2007, 2006.

[34] Jakob Nielsen. Windows 8 — disappointing usability for both novice and power users. http://www.nngroup.com/articles/windows-8-disappointing-usability/, 2012. [Online; accessed 26-May-2013].

[35] Donald A Norman. Design rules based on analyses of human error. *Communications of the ACM*, 26(4):254–258, 1983.

[36] Donald A Norman. *The psychology of everyday things*. Basic Books (AZ), 1988.

[37] Jeffrey Rubin and Dana Chisnell. *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*. Wiley, 2008.

[38] S.B. Shneiderman and C. Plaisant. *Designing the user interface 5 th edition*. Pearson Addison Wesley, USA, 2009.

[39] S.L. Smith and J.N. Mosier. *Guidelines for designing user interface software*. Citeseer, 1986.

[40] Erwin Visser. What businesses will want to know about the windows 8 release preview - windows8. http://blogs.windows.com/windows/b/business/archive/2012/05/31/what-businesses-will-want-to-know-about-the-windows-8-release-preview-windows8.aspx, 2012. [Online; accessed 04-May-2013].

[41] Richard Waters. Microsoft prepares rethink on windows 8 flagship software. http://www.ft.com/intl/cms/s/2/330c8b8e-b66b-11e2-93ba-00144feabdc0.html, 2013. [Online; accessed 26-May-2013].

[42] Max Wertheimer. Laws of organization in perceptual forms. *A source book of Gestalt psychology*, pages 71–88, 1938.

[43] Michael John White and Hans Janssen. The story of de stijl. 2011.

Appendix A

Informed Consent Form

Deltakelse på brukbarhetstest av Windows 7 og 8

Samtykkeerklæring

Jeg har mottatt informasjon om studien, og fått anledning til å stille spørsmål. Jeg er klar over at det er frivillig å delta, og at jeg kan trekke meg fra studien når som helst uten å oppgi noen grunn.

Det vil bli tatt video- og lydopptak av brukbarhetstesten. Dette gjøres for at vi skal kunne analysere opptakene i etterkant og sikre at vi har forstått deres utsagn og handlinger riktig. Vi vil sørge for at materiale vil bli anonymisert slik at det ikke vil være mulig å føre opplysningene tilbake til enkeltpersonene som deltar i prosjektet. Dette innbærer at informasjon som blir formidlet til offentligheten ikke vil kunne settes i sammenheng med den enkelte. Det er kun de involverte i prosjektet som vil kunne se opptakene i ettertid.

Jeg samtykker i å delta i studien.

Trondheim,_____

_____

Underskrift

Appendix B

Questionnaire. Demographics

Venligst besvar spørsmålene under

Kjønn:
Alder:
Sivil Status:
Utdanning:
Yrke:

Appendix C

Drafts for Interview Conducted Before the Computer Task Session

Spørsmål som skal stilles før brukbarhetstesten skal gjennomføres:

Hvilken erfaring har du i bruk av - Datamskin? - Internet?	
Har du tilgang til data-maskin med Internet?	
I tilfelle, hvor ofte bruker du Internet?	
Har du smart-telefon?	
Hvor ofte bruker du sms?	
Leser du avisen på nettet?	
Bruker du nettbank?	
Er du på facebook?	

Appendix D

Drafts for Interview Conducted After the Computer Task Session

Spørsmål som skal stilles etter at brukbarhetstesten er gjennomført:

Hvordan synes du at gjennomføringen av oppgavene gikk?	
Hva var det som var vanskelig?	
Hvorfor synes du det?	
Har du brukt lignende datasystemer tidligere?	

Appendix E

Questionnaire. SUS

Noen spørsmål om systemet du har brukt.

Vennligst sett kryss i kun en rute pr. spørsmål.

	Sterkt uenig				Sterkt enig
1. Jeg kunne tenke meg å bruke dette systemet ofte.	1	2	3	4	5
2. Jeg synes systemet var unødvendig komplisert.	1	2	3	4	5
3. Jeg synes systemet var lett å bruke.	1	2	3	4	5
4. Jeg tror jeg vil måtte trenge hjelp fra en person med teknisk kunnskap for å kunne bruke dette systemet.	1	2	3	4	5
5. Jeg syntes at de forskjellige delene av systemet hang godt sammen.	1	2	3	4	5
6. Jeg syntes det var for mye inkonsistens i systemet. (Det virket "ulogisk")	1	2	3	4	5
7. Jeg vil anta at folk flest kan lære seg dette systemet veldig raskt.	1	2	3	4	5
8. Jeg synes systemet var veldig vanskelig å bruke	1	2	3	4	5
9. Jeg følte meg sikker da jeg brukte systemet.	1	2	3	4	5
10. Jeg trenger å lære meg mye før jeg kan komme i gang med å bruke dette systemet på egen hånd.	1	2	3	4	5

|

SUS
Norsk versjon ved Dag Svanæs
NTNU 2006

Appendix F

The Scenario and Tasks

Scenario:

Du kommer på jobb en morgen og ser at datamaskinen din er blitt oppgradert. Det blir ikke gitt noen opplæring, og du må fortsette arbeidsdagen som vanlig. Du har en del gjøremål som du må få gjort.

Informasjon om brukeren din i datasystemet finner du her:

Brukernavn:
Test

Epost:
win8.testbruker@gmail.com

passord:
NTNU2013

Tasks:

Oppgave 1:

Du ønsker å se hva som er skjedd i verden og åpner nettleseren din for så å gå inn på aftenposten.no.

Oppgave 2:

Du skal finne et dokument med navn: "Kontorer.txt" Sist gang lagret du det på skrivebordet(desktop) ditt.

Oppgave 3:

Du skal nå sende en mail til nikolas@stud.ntnu.no der du legger ved filen du åpnet i oppgave 2. Tittelen på mailen skal være "Her er filen med oversikt over kontorer" og innholdet er "Se vedlagt fil". Send så mailen.

Oppgave 4:

Du skal så finne frem til en presentasjon (Presentasjon av kontorene.pdf) som ligger lagret på den lokale disken under mappen "Presentasjoner". Åpne filen når du har funnet den.

Oppgave 5:

Du må nå forlate plassen din og skal derfor slå av maskinen. Slå den av.

Completion criteria:

Oppgave 1:

Aftenposten.no vises i browseren til brukeren.

Oppgave 2:

Brukeren finner filen og gir klart signal at de vet de har funnet den ved enten å si ifra til testleder eller ved å åpne filen.

Oppgave 3:

Mailen blir sendt med vedlegg til nikolas@stud.ntnu.no.

Oppgave 4:

Bruker finner og åpner "Presentasjon av kontorene.pdf".

Oppgave 5:

Datamaskinen er enten slått av eller bruker har logget helt ut.

Appendix G

The Participants of the Usability Tests

Participant	P1	P2	P3	P4	P5	P6
Gender	F	F	F	F	M	M
Age	54	60	44	45	49	40
Education	High School	Trade School	University 3 years	MSc Sociology	MSc in Economics and Business Administration	Cand. Polit
Profession	Higher executive officer	Higher executive officer	Office administration	Adviser	Adviser	Student officer
Experience with computers	Daily use. Work and private	5 hours a day	5-6 hours a day	7 hours a day	Daily	7.5 hours a day
Access to internet	Yes	Yes	Yes	Yes	Yes	Yes
Internet usage	Daily	5 hours a day	Most of the time	Daily	All the time	7.5 hours a day
Has a smartphone	Yes	Yes	Yes	Yes	Yes	Yes
SMS usage	Daily	Daily	Daily	Daily	Daily	Daily
Read newspapers online	Yes	Yes	Daily	Yes	Yes, daily	Yes
Use internet banking	Yes	Yes	Yes	Yes	Yes	Yes
Has a facebook account	Yes	Yes	Yes	Yes	Yes	Yes

Table G.1: Participants experience with IT solutions in group 1.

Participant	P7	P8	P9	P10	P11	P12
Gender	M	M	F	F	F	F
Age	41	32	56	61	49	35
Education	Master of Arts	BA - Public administration and management	College - Echonomics	Secretarial school	High school	MSc Political Science
Profession	Office manager	Senior executive officer	Senior executive officer	Senior executive officer	Adviser	Student officer
Experience with computers	4-5 hours a day	5-6 hours a day	7 hours a day	40 hours a week	10-12 hours a day	6.5 hours a day
Access to internet	Yes	Yes	Yes	Yes	Yes	Yes
Internet usage	3-4 hours a day	Daily	6 hours a day	Daily	10-12 hours a day	Daily
Has a smartphone	Yes	Yes	Yes	No	No	Yes
SMS usage	Daily	Daily	Daily	Daily	Daily	Once a week
Read newspapers online	Yes	Yes	Yes	Yes	Yes	Yes
Use internet banking	Yes	Yes	Yes	Yes	Yes	Yes
Has a facebook account	Yes	Yes	Yes	No	Yes	Yes

Table G.2: Participants experience with IT solutions in group 2.

Participant	P1	P2	P3	P4	P5	P6
How did the execution of the tasks go?	Quite okay actually. Fun.	Challenging.	Pretty much okay.	Windows 7 went great. Windows 8 did not go to well.	Okay.	It went okay.
What was difficult?	Sending an email. Finding the subject line and entering the content.	Some challenges. The old vs. the new.	Windows 8's new screens. Unusual color pallet.	I knew what to do, but could not locate where I was going to click.	Shutting down programs.	It was unfamiliar. Navigation.
Why do you think so?	Could not find them.	Different icons you have to relate to. New placings on the screen.	Big difference from Windows 7.	Buttons/icons were hidden.	It was not logical.	Did not have experience with it.
Have you used similar systems earlier?	Looks like a Mac.	No, have never used anything similar.	Looks like the screens on iphone/smartphones.	No!	No.	Yes, Windows 7.

Table G.3: Participants from group 1 evaluation after the test.

Participant	P7	P8	P9	P10	P11	P12
How did the execution of the tasks go?	Rather bad. Very frustrating.	Okay.	Much time spent searching for icons. Finding of the ”Start”-icon was random.	Time consuming, waste of time when having to search through the system.	Went bad.	Okay. Took some time. Unusual.
What was difficult?	Unfamiliar navigation and user interface.	Do not understand why it has to be so ”fancy-pancy”	Navigating from one program to the other.	To feel ”lost in space”.	Unrecognizable symbols in Windows 8.	Hard to navigate with a new appearance/mindset.
Why do you think so?	Unfamiliar and unknown.	So much unnecessary stuff.	Missing icons to quit and minimize the program-window.	Empty screen, using the cursor to look after help-icons in the corners.	Did not look like anything that is familiar to me. Little intuitive.	Big change from last version.
Have you used similar systems earlier?	No	Looks like some of the tablets.	Using some ”unusable” systems for my daily work, but they are easier to use than this one.	No	No	Looked like some of the smart-phones.

Table G.4: Participants from group 2 evaluation after the test.

Appendix H

The System Usability Scale Results

Participant	P1	P2	P3	P4	P5	P6
1	1	3	4	4	1	4
2	4	3	1	4	5	2
3	3	3	4	2	1	4
4	3	3	1	2	5	2
5	4	3	3	3	1	3
6	4	3	2	2	5	2
7	5	3	3	4	2	4
8	3	3	1	3	3	1
9	1	3	3	1	1	3
10	4	3	1	2	5	2

Table H.1: The system usability scale given after the test

Participant	P7	P8	P9	P10	P11	P12
1	3	1	1	1	1	5
2	3	5	5	5	5	1
3	1	4	1	1	1	3
4	5	1	5	5	5	4
5	1	2	1	1	5	3
6	5	5	5	5	1	2
7	4	1	3	1	3	4
8	5	1	5	5	5	3
9	1	2	1	1	1	2
10	2	2	5	5	5	3

Table H.2: The system usability scale given after the test

Appendix I

Distribution of Gender in Positions

Antall navn			**Kjønn**		
Kategori	St.kode	St.bet	K	M	Totalsum
Konto	1068	Fullmektig		1	1
	1069	Førstefullmektig	2	1	3
	1070	Sekretær	4	1	5
	1071	Kontorleder	1		1
Totalt Konto			7	3	10
LED	1004	Rektor		1	1
	1060	Avdelingsdirektør	6	18	24
	1062	Direktør	1	1	2
	1474	Dekan	1	6	7
	9305	Prorektor	2	1	3
Totalt LED			10	27	37
MLED	1054	Kontorsjef	31	21	52
	1206	Undervisningsleder	1		1
	1211	Seksjonssjef	26	35	61
Totalt MLED			58	56	114
Saksb	1063	Førstesekretær	14	1	15
	1065	Konsulent	134	20	154
	1113	Prosjektleder	8	18	26
	1182	Seniorarkitekt	1		1
	1363	Seniorkonsulent	100	40	140
	1364	Seniorrådgiver	47	49	96
	1408	Førstekonsulent	362	49	411
	1434	Rådgiver	116	56	172
Totalt Saksb			782	233	1015
Totalsum			857	319	1176

Appendix J

The Interviews

Script:

Hvordan foregår oppgraderingsprosessen for operativsystemer?

Hvordan gjennomføres dette(prosessen)?

Kriterier for utskifting?

Hvordan vektlegges brukervennlighet i prosessen?

Hva er tankene rundt windows 8?

Windows 8 sin brukervennlighet, hvordan ser de på den?

Hvor finner de informasjonen de trenger for å ta stilling til dette?

Når oppgraderinger gjennomføres, hva gjøres for å få overgangen til å gå glatt for brukerne?

Tilbyr de noen form for opplæring?

Hvordan ville de levert ut Windows 8 på brukernes maskiner? Hvilke programmer og hvordan ville de vært konfigurert?

Transcription:

Parts from interview with participant 1(P1), Infrastructure architect from the IT staff at a Norwegian University

INTERVIEWER: Du kan kanskje si litt om hva dere har gjort rundt det og hva dere har sett på?

P1: Ja, i forbindelse med Windows 8? Eller vil du vite hvordan vi gjør det med oppgraderinger fra en plattform til en annen?

INTERVIEWER: Ja, begge deler er viktig i denne sammenheng.

P1: Vet ikke hvilken ende du vil begynne i, om du da vil begynne med hvordan vi ser for oss oppgraderinger. Oppgraderinger for oss... Vi drifter grovt sett 2500 Windows PCer. Det er et greit nok antall, men vi drifter da i hovedsak for ikke-forskere. Det som da er utfordringen, eller... Det er mye utfordringer når man skal gjøre en sånn stor oppgradering, og hvordan det skal gjøres. Skal alt bli tatt på en gang osv. osv. Så vi har valgt en måte hvor vi bruker en automatisert utruller.

INTERVIEWER: Ja...

P1: Det betyr at det egentlig er mye arbeid på forhånd, fordi vi skal lage et oppsett som gjør at brukeren føler seg hjemme og får til det de skal gjøre. Så det er en vesentlig del. Derfor har mesteparten av opplæringen hos oss stort sett gått til driftsapparatet. Så opplæringen av driftsapparatet er faktisk en vesentlig del. Hvilke støtteverktøy skal man bruke for å oppgradere. Man går jo ikke ut å oppgraderer en og en PC.

INTERVIEWER: Ok, så dere gjør da tanking av PCene via nettverket?

P1: Ja

. . .

P1: Hvis man skal lære opp folk, så hadde vi det også. Vi hadde opplæring på de viktigste delene som folk bruker. Når det gjaldt overgang fra Windows XP til Windows 7, hadde du fortsatt startmeny osv, så det var applikasjonene som krevde opplæring. Så i den sammenheng vurderte vi å bruke mikro-opplæring, har du vært borti det?

INTERVIEWER: Nei, har ikke det.

P1: Vi har egentlig en avtale med et firma som heter Junglemap.no, også veldig mye brukt av Microsoft, ihvertfall i Norge. De gir sånne tips på epost der du får f.eks. fire sider som du må se igjennom. Vi vurderte mikro-læring, men tok det ikke i bruk. Vi vet det er brukt mange andre steder, men vi gjorde det ikke denne gangen... Ehm... Kunne kanskje vært lurt å gjort det fordi jeg tror folk kunne trengt denne opplæringen, spesielt ettersom vi skiftet mailplattform der vi gikk fra smtp til exchange.

INTERVIEWER: Ja.

P1: Type Office klienten... Ehm... Så vi vurderte behovet for opplæring da vi gikk fra XP til Windows 7 som lite... Windows 8, det som vi ser på for øyeblikket, er jo at vi ikke føler at Windows 7 har satt seg lenge nok til at vi innfører Windows 8 til brukermassen.

INTERVIEWER: Ok

P1: Det blir rett og slett for mye endring for fort for brukerne. Folk sliter nok egentlig med å tilpasse seg endringene ved mail og... og hovedplattformen. Så da må vi vente litt.

INTERVIEWER: Mhm...

P1: Så blir det jo litt sånn type Vista, der vi vet at annenhver versjon er grei. Så Windows 9 kanskje... Hehe... Men!... Vi har gjort en kvasi-løsning.

INTERVIEWER: Okay?

P1: Når det gjelder Windows 8. Der tror jeg faktisk Windows 8 kan ha veldig mye for seg.

INTERVIEWER: Hvordan type kvasi-løsning har dere laget?

P1: Terminal server 2012 som vi nå driver og setter opp. Hvis man setter opp Terminal server 2012 så vil man kunne logge seg på der. Der vil man bruke et Windows 8 brukergrensesnitt, en remote desktop. Der tror jeg at en innføring av Windows 8 vil virke. Man kommer med f.eks. padden sin, og logger seg på for å gjøre veldig enkle ting. Man skal starte Word, Excel og sånne ting som man ikke har på maskinen sin ellers, og man får dette med store enkle ikoner som man kan trykke på for å kjøre applikasjonene... Jeg kan vise deg det her.

. . .

INTERVIEWER: Kataloger og applikasjoner dere tilbyr via Terminal Server 2012, er dette noe dere bestemmer, eller vil brukerne kunne gjøre dette selv?

P1: Nei, dette vil være noe vi setter opp selv. Så fordelen er at de som er studenter vil få Word og Office via dette. Noe som vi har satt opp på forhånd.

. . .

INTERVIEWER: Hva er tankene deres rundt all native støtte dere går glipp av ved å gjøre det på denne måten? Dere mister jo touch og ting som eyetracking...

P1: Det har vi egentlig ikke tenkt på... Det som er utfordringen for en IT-organisasjon sånn som oss, det er at vi skal levere en solid og stabil plattform som brukerne kan leve godt med. Så f.eks. de maskinene vi får nå fra Dell... De nedgraderer vi til Windows 7 fra Windows 8.

. . .

INTERVIEWER: Du nevnte tidligere at det ikke har vært planer om å oppgradere til Windows 8, er dette noe som kan endre seg med tiden og med fremtidige oppgraderinger som blir gjennomført med Windows 8?

P1: Vi kan nok ikke se bort ifra det. Det som er, er at vi må se mer på hva konsumenten ønsker. Det er de som styrer, og hvis folk vil ha Windows 8, så vil nok vi oppgradere.

INTERVIEWER: Men pr. dags dato, så har dere ikke tatt noe analyse av Windows i større skala? Mere på privat bruk.

P1: Nei, ikke mer enn det. Det som derimot har vært ett problem, er trådløs-støtten i Windows 8.

INTERVIEWER: Ja, jeg har hørt om det.

P1: Ja, så vi må vente på at leverandørene fikser dette før det blir aktuelt med noe sånt.

. . .

INTERVIEWER: Tidligere år, har dere satt kriterier for å gjennomføre oppgraderinger? Eventuelt, hva har dere lagt til grunne for dette?

P1: Det er vel ikke så mye brukervennlighet, det er vel mer sikkerhet som oftest. Hvor lenge er det før denne plattformen går ut på dato? Må man oppgradere? Dette styrer oppgraderingen mer enn brukervennlighet og innovasjon. Så får man heller innovere på andre plattformer.

INTERVIEWER: Holdt dere da noen kurs for brukerne sist gang?
P1: Nei, vi holdt nesten ingen kurs for brukerne sist. Vi burde kanskje hatt flere kurs, men... Altså, hva er det folk har hjemme? Jo, de sitter i Windows 7. Ikke sant?
INTERVIEWER: Ja.
P1: Så de hadde sikkert lært seg å bruke Windows 7 fra før av.

. . .

INTERVIEWER: Nå har dere jo ikke gjort så mye research på Windows 8, utenom det personlig.
P1: Nei, vi har ikke det, men policyen våres er at vi pleier å støtte oss på hva de store analysebyråene sier. Vi hører en god del på hva Gartner sier... og.. ehm... Gartner er jo veldig skeptisk til Windows 8. Vi har jo ingen egen produktavdeling som går og ser på nye produkter, så da støtter vi oss på de.
INTERVIEWER: Ja, det høres jo greit ut.
P1: Vi betaler jo for de, så da hører vi jo på de i tillegg til at vi leser litt rundt, men Gartner er de viktigste.

Parts from interview with participant 2(P2), Leading Advisor Solution Architecture from the Corporate IT Staff in a Large Norwegian Oil Company(LNOC)

INTERVIEWER: Du kan få lov til å starte litt med å si hva dere har gjort av Windows 8 arbeid.

P2: Hm... Ja, altså LNOC har jo ikke rullet ut Windows 8, fordi at vi rullet ut Windows 7 for to år siden. Jeg skal ikke akkurat si at vi sliter med dette fortsatt, men det er mye ting som det enda må ryddes opp i. Så er det noe med hvor mye endringer vi kan påføre brukermassen, men når det er sagt, så har vi gjort en del eksperimenter på Windows 8. Så har vi noe på testlabben og sånt for å kjøre Windows 8 med vår infrastruktur. En ting vil jo være selve klienten, og det er jo kanskje det enkleste. For det å kjøpe en PC med Windows 8 på, det får vi jo opp ganske fort, men det å få Windows 8 til å virke i et så stort selskap som LNOC, det er en ganske stor jobb. Sånn at... Den jobben er veldig lik det vi gjorde for Windows 7. Jeg kan fortelle en del om de erfaringene fra overgangen med Windows XP til Windows 7.

INTERVIEWER: Ja, gjerne.

P2: Så vil vi jo få veldig mange av de samme erfaringene når vi ruller ut Windows 8. Det er ikke bestemt helt enda om og når dette skal gjøres, men det kan jo tenkes at det blir neste år.

INTERVIEWER: Ja, for det er interessant å høre hvordan denne prosessen var sist gang dere oppgraderte. Da vet vi litt mer hva dere ser på når dere skal oppgradere til Windows 8.

P2: Hvis vi begynner i den enkleste enden. Det er mange ting å huske på, så vi får bare begynne et sted.

INTERVIEWER: Ja.

P2: En av utfordringene som vi har i et så stort selskap, det er at vi har rundt 3000 forskjellige applikasjoner som går på Windows. Så er det sånn at noen av dem er moderne, og noen ikke fult så moderne. Noen av de er egenutviklede på utviklingsverktøy, som ikke nødvendigvis fungerer og hvor endringer i operativsystemet gjør at de ikke fungerer slik som de gjorde på en tidligere versjon av Windows. Det var en av de tingene som jeg husker vi strevet veldig med i, dette begynner å bli en stund siden, utrullingen av XP servicepack 3.

INTERVIEWER: Javel?

P2: Fordi at noen ganger innfører Microsoft nye sikkerhetsmodeller. I XP servicepack 3 innførte de data execution prevention, address space randomization layout, stack hand trace og double unlinking. Som er fire sikkerhetsmekanismer. Som brukes i operativsystemet for å sørge for at et program ikke kan ødelegge for andre programmer. De applikasjonene som vi hadde den gangen og som var utviklet tidligere, de gav oss noen utfordringer med den type ting. Dette gjør at det rett og slett er en del ting som må gjøres. Da må man rydde. Dette må gjøres med alle de 3000 applikasjonene. De skal testes, de skal verifiseres og du må jobbe med å sikre at alle applikasjonene fungerer sånn som de skal. Det betyr at hvis man tar de aller fleste applikasjonene og sier at 2500 av de applikasjonene var relativt greie. Det var da snakk om å få tak i riktig versjon av de. Adobe reader, Spotify osv, de funket greit.

INTERVIEWER: Rett og slett bare å oppgradere de?

P2: Ja, riktig! De er ikke sånn kjempe kompliserte, men de siste 500 var det slik at det gikk enormt mye resurser med til å teste, forbedre og gjøre forskjellige ting med dem. Hvis man da tar Office som et eksempel, så gjorde Microsoft noe med filformatet til Office fra 2003 til 2007.

INTERVIEWER: Ja, med docx osv?

P2: Ja, med oxml og sånt. Så vi var veldig involvert i å pushe på ISO og ECMA ettersom vi hadde mange dokumenter. Derfor samarbeidet vi med Microsoft for å utvikle oxml, og vi pushet på i ECMA og ISO for å få dette igjennom som krevet en del arbeid. Så med alle sånne oppgraderinger og grunnleggende ting må vi jo teste det, og for dette tilfellet hadde LNOC gjort mye arbeid med maler og alt mulig sånt og hadde integrert dette med forskjellige applikasjoner og sånt. Alt dette måtte gjennomgås og testes hver gang det skal gjøres en slik type teknisk oppgradering, og dette koster jo ganske mye.

INTERVIEWER: Ja.

P2: Ikke sant, så hvis du regner i kroner og øre blir dette en del.

. . .

INTERVIEWER: Dette er jo bare før utrulling har skjedd, og det vil vel trolig være en del ting som skjer etterpå?

P2: Ja, derfor kjørte vi Office 2003 en stund etterpå til vi hadde klart å dra med oss alt over på det nye. Vi brukte nesten 2 år på dette,

da vi rullet ut Windows 7. Det var jo mange applikasjoner som det måtte gjøres masse greier med for å få det til å virke på Windows 7.

. . .

INTERVIEWER: Er det slik at dere på LNOC kjører et fast operativsystem, eller kan de velge litt hva de ønsker selv? Det brukes vel en del forskjellige plattformer som byr på litt problemer?

P2: Ja, altså vi har en standard plattform om vi kan si det slik, og det er Windows 7. Så er det slik at hvis du har det vi kaller spesielle tjenstlige behov, som typisk er hvis du er lete-geolog, eller noe sånt. Så er det Linux som gjelder, og du får en Linux maskin i tillegg til din vanlige maskin. Det er jo et eget opplegg, men standard er at du får Windows 7. Så jobber vi oss mot en løsning, som er bring your own device. På mobil, og vi jobber oss dit med desktop klienter også. Uten å love noe, så er planen at man i løpet av neste år skal kunne tillate dette. Det viktigste utfordringen med utrullingen av dette, er litt den samme som med utrullingen av Windows 7. For da vi gikk over til Windows 7 ble det laptop til alle mann.

INTERVIEWER: Ja.

P2: Og det hadde masse skattemessige konsekvenser. Da måtte du plutselig begynne å skatte av arbeidsredskap... eller av fordelen med å få PC av arbeidsgiver, og dette vil nok bli enda verre for bring your own. Og dette viser jo at utrulling av denne type prosjekter har masse utfordringer med skattemyndigheter og alt mulig sånt.

. . .

P2: Går vi tilbake til det tekniske, så gjorde vi ganske mye arbeid med back-enden da vi oppgraderte fra Windows server 2003 til Windows server 2008 release 2, og siden til Windows server 2012. Så det foregår en kontinuerlig oppgradering på back-end-siden også.

INTERVIEWER: Ja.

P2: Og da får du en del fordeler rundt sikkerhet, og du får noe som heter direct access. Som er en måte å gjøre sånne level two tunneling protocol ting.

INTERVIEWER: Javel?

P2: Ja, det har med VPN å gjøre. Når du setter opp en tradisjonell VPN klient, så er det sånn at all trafikk routes gjennom den samme kanalen. Hvis du gjør en level two tunneling protocol, så blir det applikasjonsavhengig eller adresseavhengig. Så det betyr at alt som går til LNOC.no-servere går via direct access VPN, mens alt annet går en annen vei. Så hvis du sitter på hjemmekontoret vil det bare være en liten del av trafikken som går via VPNen inn til LNOC, og ikke all den andre trafikken. Dette er mye mer stabilt siden du ikke trenger å ha VPN klienten gående hele tiden. Dette var jo en av hovedgrunnene til at vi rullet ut Windows 7. Dette er også en del av det å rulle ut bring your own device, for da kan vi krympe sammen brannmur perimeterne slik at kontorene vil være på internett som om det var internett hjemme, mens serverne vil være bak en brannmuren.

INTERVIEWER: Ok, så dere vil anta at dette er lavere sikkerhet enn vanlig?

P2: Nei, vi vil anta at det ikke er sikkert i det heletatt! Vi vil bygge sikkerhetsmekanismen rundt dette, og det kan vi med Windows server 2012. Windows 7 lot oss få til dette, og disse funksjonene er blitt enda bedre i Windows 8. Dette vil kanskje være en av de tingene som gjør at vi ruller ut Windows 8... Det vil kanskje være to ting som gjør at vi ruller ut Windows 8. Den ene er at det blir billigere med bring your own device, og det andre er det med å flytte perimeteret fra inngangsporten til datasenteret som er mye enklere å sikre. Så det er nok noen infrastruktur-drivere som kan få oss til å rulle ut Windows 8.

. . .

P2: Noe jeg for så vidt glemte i sted, når det gjelder sikkerhetspolicyen mellom administrator og bruker som egentlig ble introdusert i Windows Vista og som de aldri fikk til å funke ordentlig før på Windows 7.

INTERVIEWER: Mhm.

P2: Dette var også en av de tingene som gjorde at vi måtte jobbe en del med applikasjonene. Mange av applikasjonene forutsatte at man var administrator, og dette kunne vi jo ikke drive å dele ut til alle. Da skjer det sånn som da vi fikk en orm som tok ned hele nettet pga. en person hadde installert SQL server uten å patche denne. Dette er jo en av de tingene som er blitt mye bedre med Windows server 2012 og Windows 7 der vi kan pusher patcher

til brukere. Selv på programmer de har installert selv. Så for oss som har mye sikkerhet rundt gass-løsningene våres osv. vil det å introdusere Windows 8 være bedre siden vi kan forhindre dette mye lettere enn det vi kunne i Windows XP. Så for noen som oss, så er sikkerhet veldig viktig, og det kan være sånne ytre drivkrefter som også fører til at vi må oppgradere.

. . .

INTERVIEWER: Vektlegger dere brukervennlighet i forbindelse med oppgraderingene? Jeg regner med det er en del forskjell på kunnskapsnivået til de forskjellige ansatte, så hvordan håndterer dere dette?

P2: Ja, det er riktig gjettet. En av de tingene som var... Den største overgangen til Windows 7, det var jo måten startmenyen funket på og introduksjonen av ribbons i Office og alle de produktene der.

INTERVIEWER: Altså ikoner osv. i startmenyen?

P2: Ja, og hvordan det søkes etter programmer. Tidligere var det jo slik at du tok opp startmenyen, også måtte du finne riktig gruppe, og så måtte du åpne, så var det litt sånn klikking for å komme frem til riktig program. Nå er det jo slik at du søker bare på de tre første bokstavene, så kommer programmet frem med en gang. Og det der var noe som... Vi rullet jo ut... ehm... Windows 7, Office og Communicator release 2 samtidig. Da var det masse innsats som ble lagt ned i å lage opplæringsmateriell for disse, og det ser jeg at selv to år etterpå så strever folk veldig med Windows 7. Altså jeg synes jo, at hvis du som 45-50 åring er teknologisk etterlatt, så tenker jeg som så at du har 40 år igjen å leve. Man kan jo bare tenke seg hva slags teknologisk utvikling man skal gjennom de neste 40 årene. Det er jo bare å tenke seg hvordan man ville klart seg hvis man personlig hang teknologisk igjen i 1972. Det ville ikke vært lett å klare seg slik i vår tid. Men uansett så har vi brukt mye resurser på opplæringsopplegg og å klare å gi folk informasjon. Spesielt med overgangen til ribbons. Det var en bøy for mange. Vi strevde skikkelig med å forklare hvordan dette virket, og å få folk til å forstå dette.

. . .

P2: Ser man for seg at man skal rulle ut Windows 8, så er jo det litt sånn gestures orrientert.

INTERVIEWER: Ja, det er en del gestures i det. Så er det også en overgang med tanke på Windows UI.

P2: Ja, det vil være en kjempestor overgang å gå over til den måten å tenke på. Dette er en diskusjon som vi har hatt, men som vi ikke har landet noe svar på enda. Hvor velegnet er egentlig Windows UI for den type applikasjonsbruk som LNOC har. For hvis du har en device med en skjerm, så er det greit. Men så har vi testet litt med to skjermer, eller tre, eller fire, eller fem skjermer som vi av og til har. Så blir disse Windows UI-greiene forferdelig kjipe. For det å ha mange applikasjoner oppe samtidig og det å bytte mellom disse, det er, så vidt jeg har forstått, mye vanskeligere enn i Windows 7. Dette er nok en hindring for adapsjon også. For hvis det ikke funker, så kan man ikke bruke det.

. . .

INTERVIEWER: Den opplæringen dere arrangerte, var det mere kursing osv. eller var dette tekstlig basert?

P2: Det var e-læringskurs. Veldig mye e-læringskurs. Disse kursene var på en og en halv time. Så det sier jo sitt, når vi tar ut 50000 ansatte ut i kursing når vi innførte Windows 7 for en slik liten oppdatering.

. . .

P2: Ehm... men ja... sånn uansett, så tror jeg i det øyeblikket du dytter masse endringer på folk... og det har ikke noe med alder å gjøre heller. Jeg ser jo at de 30 år gamle geologene våres kludrer like mye de altså, så det her er ikke bare forbeholdt den eldre garde.

INTERVIEWER: Nei, det er forståelig.

P2: Så de som hadde fått dette, ville plundret mye. Det betyr at vi måtte brukt masse... Altså hvis du skal rulle ut Windows 8, må du bruke masse energi på tilrettelegging, masse energi på opplæring og du får masse energitap på plunder, heftelser og irritasjon. Det så vi jo, for vi gjør jo sånne medarbeider-undersøkelser.

INTERVIEWER: Mhm.

P2: Og da er det jo blant annet spørsmål om IT-verktøy og sånt. Denne gikk jo på en smell da vi rullet ut Windows 7. Da fikk vi en dipp i satisfaction på IT-siden, fordi folk var litt frustrerte med den utrullingen. Fordi det var en stor overgang til noe nytt.

375 INTERVIEWER: Ja, siden det ikke var sånn som de var vant til?
376 P2: Ikke sant. Så noen var veldig mye mer fornøyd med dette, og noen
377 var mye mer frustrert, og man vil nok få en enda mer polarisering
378 i... Altså ingen vil være likegyldige hvis du ruller ut Windows 8.
379 Så kan man jo gjette på hva det er som blir utfallet av dette.