University of Bergen

Master Thesis

Exploring design principles for self service technologies: The case of a ticket vending machine

Author:

Marius Seim

Supervisor:

Viktor Kaptelinin

in the

Department of Information Science and Media Studies

May 30, 2014

UNIVERSITY OF BERGEN

Abstract

Faculty of Social Sciences
Department of Information Science and Media Studies

Masters degree

Exploring design principles for self service technologies: The case of a ticket vending machine

by Marius Seim

The principal goal of this thesis was to research self service technologies used in public spaces, and how they can be improved upon. The main research question posed for this thesis was: *"How well do the existing design principles support the ongoing development of self-service systems and are they sufficient?"*. There were also two sub-questions posed to further explore the topic: *"How can novel design principles be used to improve the usability of self-service technologies?"* and *"Which methods are optimal for researching self-service technologies?"* In order to answer these questions a range of research methods were used. These methods included observations, focus groups, usability testing and a review of existing literature. For the purpose of having a clear focus a specific self-service technology was chosen, the Skyss ticketing system. A prototype version of the existing system was created and tested with users, leading to a novel design principle that was named integration. A second iteration was also created to demonstrate the concept of the novel principle. The conclusion of the research was that a novel design principle can potentially enhance SST. However, there is still much room for improvement in the domain of self-service technology, and more research into design principles and methods is a feasible way of doing so.

Acknowledgements

Many people have been a part of the process that has lead to the work presented in this thesis, and I would like to give them thanks for this.

Firstly, I would like to thank my supervisor Viktor Kaptelinin for his guidance, patience, wealth of knowledge and for helping to formulate an idea for my thesis.

I would like to thank my friend, Hanne Aga, that volunteered to read through the whole thesis and correct any grammar mistakes that might have snuck in.

I would also like to thank Kristian Vårdal, Nils Haukås, Snorre Davøen, Grace Kanza, Ørjan Ertkjern and Maiken Fjellanger for their positive influence, feedback, discussions and good company throughout the work on this thesis.

Lastly, I would like to thank everyone who has been a participant in either focus groups or user testing. Without you there would be no thesis to discuss.

Contents

List of Figures

List of Tables

Nomenclature

HCI	Human-Computer Interaction
ID	Interaction Design
SSTs	Self-Service Technologies
HE	Heuristic Evaluation
TVM	Ticketing Vending Machine

Chapter 1

Introduction

In recent times self-service technologies have become more and more popular in public spaces. Some of the reasons for this is the growing need for efficiency in our daily lives, reduced hardware prices and the increasing cost of manual labor. These technologies are appearing in areas such as ticket sales, bank terminals and as tour guides in information offices (Hagen and Sandnes, 2010). Previously many of these tasks involved an intermediate which assisted users in completing the process. Without this link between the user and the system, the self-service technologies will have a high demand for usability in order to gain user acceptance. One way of ensuring good usability is by employing design principles while creating the system and the interface. These principles aid developers in designing for the user experience and a set of main principles have been constructed for this purpose. While this is the case, some of the principles are converted to work well with websites and web-based applications. Therefore it would be interesting to come up with or review the principles specifically for use with self-service technologies. It will also be interesting to explore whether the design of SSTs can be guided by the introduction of specific design principles for.

In this master thesis the aim is to either formulate a new design principle or revamp the previously existing ones by studying the usability of some available self-service technologies. Based on these findings, utilizing the novel principles, and drawing on the field of Human-Computer Interaction the goal is to enhance the use of self-service technologies. For the purpose of gaining more insight into the relevance of the principles, the aim is to redesign the ticketing machines created by Skyss. The focus is specifically directed toward ticketing machines that are placed along Bybanen. These machines are meant to swiftly assist a user in purchasing a ticket for the Bybanen or refilling a traveling card. A reason for choosing this route is that as opposed to on the buses, a user does not have the possibility of buying a ticket from the driver. This means that any available assistance will have to be contained within the ticketing machine, and so improving the interface can be invaluable for the efficiency of the system. Recently Skyss has raised the prices for purchasing a ticket on the bus, meaning that more people will have to use the self-service option in order to save money. This was done to reduce the waiting time at the bus stops (Skyss, 2013). The only way this will work is if the ticketing machines are easy to understand and efficient to use. Another reason that the system needs to be efficient is that users will often not have much time to purchase a ticket before the Bybanen leaves the station. In the duration of the thesis several newspaper articles have been published, criticizing the usability of the ticketing machines, even leading to users being fined for not having a valid ticket (Kvamme, 2014; Fagervoll, 2014).

The thesis will thus involve creating novel design principles, using them and other aspects of HCI and self-service design to re-imagine the interface of the ticketing machines.

1.1 Contact with Skyss

In April of 2012, before starting the work on the master's thesis, Skyss was contacted about the possibility of working together with a member of their development team. Even though they were not able to directly collaborate valuable feedback was given on the progress by e-mail, and it was possible to gain information about why certain design choices for the ticketing machines were made.

1.2 Motivation

The motivation for research done in this thesis is based on personal experience and observations of self-service technologies. Several times during travels around the world I have encountered a machine that is severely hard to use. Everything from a train ticketing machine in France, to a self-service restaurant in Japan. The self-service restaurant had a machine without any description or image explaining how it should be used. It was quite similar to a vending machine, and so it was possible to understand the function, but without any knowledge of which meal would be received in the process. Each button had a color code, which was assumed to have something to do with the type of meat that was in the dish. Unfortunately it turned out to be a measure of how spicy the meal was, leading to a rather unpleasant experience. Such technology should be easily understood by anyone, be they tourists or permanent residents. The main goal of this thesis is to suggest a first step to improve the self-service situation, and to create more intuitive interfaces for them.

1.3 Research Questions

An important part of this thesis is to contribute some knowledge into the field of Human-Computer Interaction , and to prove the usefulness of this knowledge by developing a prototype that embraces it. The master thesis will build upon the hypothesis that: "The general design principles are insufficient for the growing changes in information systems. By revising or creating new principles for specific platforms, the resulting products will have a higher potential of covering the needs of the user."

The main research question for the thesis is:

RQ1: **How well do the existing design principles support the ongoing development of self-service systems and are they sufficient?**

In the process of exploring this question a set of sub-questions while also be considered and will be important to the study as a whole.

- **How can novel design principles be used to improve the usability of self-service technologies?**
- **Which methods are optimal for researching self-service technologies?**

To solve these questions a set of usability studies on the self-service technologies employed by Skyss will be used. The main technology will be the ticketing machines previously mentioned. A new design for the self-service ticketing machine will be developed in form of a high-fidelity prototype. The design will draw on a design principle generated in the first phase of the thesis. Conclusively a cognitive walk-through of the final prototype and data gathered will be used to measure the effect that the novel design principles might have had.

The resulting design principle and prototype, will hopefully aid in the raised satisfaction of potential users and increase the ease of use that it has. There is also hope to spark more interest in self-service technologies, and to lay a foundation for further research in the area.

Chapter 2

Research Perspective

This section presents the two main theoretical fields that forms the base of the project. Firstly there will be a brief introduction of service design and of self-service technologies. Then the focus will be on the main research field of the thesis Human-Computer Interaction (HCI). The HCI section will present the concepts of design principles, heuristics and the seven stages of action. In the end of the chapter these concepts are compared and the term accessible design and its impact on self-service will be presented.

2.1 Service Design

Service design is a human-centered approach to design, and is concerned with systematically applying design methodology and principles to the design of services. A service is often comprised of more than just an artifact. The service is produced during the process. In other words the service is mainly experienced as it is consumed or used (Holmlid, 2007).

Service has recently been used as a metaphor for many computing applications, and practitioners often discuss services rather than applications. In lieu of this, service design and HCI are beginning to converge, and researchers are trying to dig deeper in the relationship between the two practices. As services have an activity based nature, there are many things HCI could offer, such as sophisticated ways to analyze tasks and activities (Wild, 2008).

In service design the user is the co-creator of the value, meaning that he or she needs to perform some sort of activity and thus requires motivation. Therefore one should consider how the service is experienced by the user. One way of doing this is by categorizing elements of a service into satisfiers and dissatisfiers. If a system does exactly what a user expects it can be seen as neutral. Based on this neutrality, any element that does not match expectations will be treated as a dissatisfier and any element that outperforms expectations as a satisfier. If there are many dissatisfiers this can lead to a decrease in satisfaction, and as such many satisfiers can of course lead to an increase in satisfaction (Teräs and Mäkelä, 2012).

When studying users of a service a number of contextual factors apply, such as tasks, equipment and the social environment. In order to assess the customer service experience, and the customers perceived value it is important to not only look at the individual experience, but also the context in which the service is being used (Teräs and Mäkelä, 2012).

2.1.1 What is Self-Service?

Customer services are increasingly being delivered by the use of technology. A consumer will more increasingly have to interact with some form of technology in order to use or access a service. When contact with this technology is initiated and carried out by the consumer, but involves no direct or indirect contact with an employee, it is inherently self-service. In the last decades, this model of society has become more and more common. Many situations that previously required a user to leave their home can now be resolved through the Internet, and services that required many employees can be handled by a self-service machine. Benefits of such systems are that the customer can access the service at any time, from any location(in the case of applications) and that the level of service provided is consistent from place to place (James, Peter, and Glynn, 1999).

As seen in (Figure 2.1) some factors contributing to customer satisfaction in dealing with service provider have been suggested.

1. The core service being provided.
2. Various support services and systems that contribute to the delivery of the service.
3. The technical accuracy in delivering the core and support services.
4. The interaction that customers have with employees of the firm.
5. Certain affective aspects of the interaction—essentially, how the customer is made to feel.

FIGURE 2.1: "Drivers of Customer Satisfaction" From *Handbook of Services Marketing and Management* (James, Peter, and Glynn, 1999, p. 91)

James, Peter, and Glynn (1999) state that by introducing technology-based, self-service system into the interaction between customer and service provide, the firm is supposedly improving the quality of service provided to its customers. This value can be perceived through an increase in availability. A customer can access bank services at any time of day without leaving their home, or even on the bus. While these services do enable customers to access the core service that they need, such systems generally do not allow for customization of the service nor addressing special cases that may arise. This type of support is difficult to provide without human interaction, and can be one of the main challenges of self-service systems. A poor user experience can then lead to users who abandon the system, or even escalate their problem to another service channel, such as a call center (Geest, 2013). In terms of the Skyss TVM it may even lead users to enter Bybanen without paying for a ticket.

In the article "Investigating the future of self-service technology," Robertson, Szymkowiak, and Johnson (2010) say that Internet based access can be seen as a form of self-service on the part of the users, as they are in charge of the time and place of transaction. As such both a kiosk-based system that distributes film, and the streaming of media through the Internet can be seen as a self-service. Based on age, gender and technical competence the authors performed a web-based survey to indicate the likely usage of such technology in the future. In short their findings suggested a strong preference for Internet based technologies. Current technologies also confirm this trend, with services like Netflix taking up as much as one third of all bandwidth consumption in the U.S (Reed, 2013).

2.2 Human-Computer Interaction

The area of HCI aims to understand the constraints and paradigms that define how people use technology. HCI applies multiple fields of research such as cognitive science, sociology and psychology in order to predict how people react to interfaces. Some of the basic principles of HCI will have a large impact on usability, and having a proper understanding of these concepts can aid designers in solving complicated interface issues (Nielsen and Norman, 2013). The human mind is also an important factor, as people like to think that they understand themselves. The truth is that people do not always know why they do things, or why they feel the way they do. Most of human behavior is actually the result of subconscious processes (Norman, 2013). How do we then develop something for a person that does not truly know what they want? Two of the main tools that interaction designers have to effectively analyze and interpret the usability of a product are heuristic evaluation and design principles.

2.2.1 Design Principles

As mentioned, design principles are one of the concepts that are used by interaction designers to aid in their work process. Some well-known examples of such principles that are outlined by Sharp, Rogers, and Preece (2011) are:

Visibility is to ensure that the presence of important features are highly visible. The more visible a function is, the more likely it will be that a potential user will be able to perform the correct actions. To exemplify this we can look at how the controls for different operations in a car are clearly visible, e.g. indicators, headlights, a horn and warning lights that indicate a status. These functions are also placed in such a way that it makes it easy for the user to find the appropriate control for the task at hand.

If the functions are out of sight it can make them more difficult to find, and to know how to use. If one for instance changes a well known function and makes it automatic, such as with sensor-activated devices, it can create frustration in users as they will not be aware of exactly how to make them work (Sharp, Rogers, and Preece, 2011).

Feedback is related to the concept of visibility. The importance of a device giving feedback can be illustrated by using an analogy to how an activity would change without it. Imagine playing a video game where the relation between your action and the result on screen did not make sense, or the delay was too long for the player to enjoy the game. Feedback ensures that a person is given information about what action has been performed and what the result of the action has been, allowing the person to carry on with the activity. Some examples of feedback are visual, audio, tactile and combinations of these. Picking the right kind of feedback for an activity is very important, and can also contribute to the visibility of an action (Sharp, Rogers, and Preece, 2011).

Constraints refers to determining different ways of restricting what kind of user interaction can be performed at a given time. One way of doing this in terms of a graphical user interface is to shade certain menu options in gray, indicating that the functionality is deactivated. In this way the user will only be able to perform actions that are allowed at that stage in the program. Other ways to constrain a user is in physical design, for instance by designing an external slot in a computer to only allow a certain shaped cable or card (Sharp, Rogers, and Preece, 2011).

Consistency focuses on designing interfaces to use similar elements and having similar operations for achieving similar tasks. A consistent interface tries to ensure that the same operation is used to, for instance, select any graphical object in a program, such as always clicking the left mouse button. If there is inconsistency it can make it difficult for users to remember how to use an operation, and make them prone to mistakes.

A benefit of using a consistent interface is then that they are easier to learn and to use. This design principle is easy to maintain for a simple interface, where there are a small number of operations mapped. However, in an application such as Photoshop, where there are hundreds and even thousands of operations, there is simply not enough space to map each one to an individual button or operation. In this case a much more effective solution would be to create categories of commands that can be mapped into subsets of operations (Sharp, Rogers, and Preece, 2011).

Affordance as a concept has had a long history, and is thoroughly detailed by (Kaptelinin, 2013). The term had its theoretical roots as a part of an ecological approach to visual perception. It has since been adopted by many researchers within the field of HCI, and is used in relation to for instance activity theory and phenomenology. In this thesis an affordance will be used as a term to explain the attribute an object has that allows people to understand how to use it. Affordances for physical objects such as door handles are perceptually obvious and it is easy to know how to interact with it. In order to translate this to a digital setting, things such as buttons, icons and so on should be made to appear obvious. Buttons should afford pushing and scrollbars should afford the action of moving up and down.

There are two kinds of affordances: perceived and real. Physical objects are said to have real affordances, such as grasping is something that is perceptually obvious and does not have to be learned. One such object is a hammer, which has a handle that invites a user to grip it. The way it is constructed gives a very clear indication on how it is meant to be used. However a screen-based interface does not have this same type of real affordance. Meaning that it does not make sense to try and design for real affordances, unless the design is for a physical device (Sharp, Rogers, and Preece, 2011).

There are numerous more exhaustive sets of design principles that have specific examples for designing GUIs, for the web and so on. One well-known website that provides design principles is AskTog (Tognazzini, 2013) and adds principles such as autonomy, readability and anticipation to the list. As mentioned, my goal is to add to these sets with principles for use with self-service technologies.

Applying these principles in a practical setting can be problematic when trying to add more than one principle at a time. The reason for this is that some principles will contradict others, and therefore create a trade-off between them. If there for instance is a high focus on constraints in an interface the information will also become less visible. Consistency can also be tricky, as ensuring consistency with one aspect of the interface can make it inconsistent with another. Even though this is the case there can be benefits when for instance introducing inconsistency. The interface might become more difficult to learn how to use, but over time it can make it easier to use (Sharp, Rogers, and Preece, 2011). The principles are therefore subject to interpretation, and deciding which ones are most important for a particular interface is a key aspect of the design process.

2.2.2 Heuristics

A heuristic evaluation (HE) is a method of usability analysis where a number of evaluators are presented with an interface design and are asked to comment on it. This method was first proposed by Jakob Nielsen and Rolf Molich in the paper "Improving a human-computer dialogue" (Molich and Nielsen, 1990). Their later experiments showed that individual evaluators only found between 20% and 51% of the usability problems inherent in the interface that was evaluated. While evaluation from several evaluators could be aggregated to uncover many more problems, even when consisting of only three to five people. Once the number of evaluators increase past five the problems uncovered start flattening out, and as seen in (Figure 2.2) past ten, the number of problems found will normally reach its peak (Nielsen, 1990).

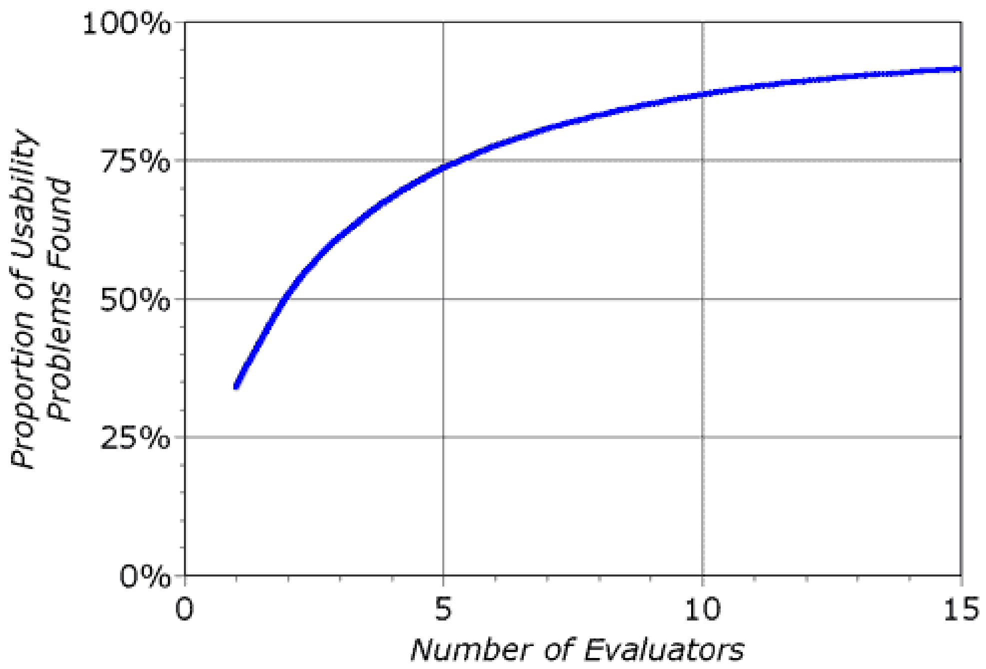


FIGURE 2.2: "Curve showing the proportions of usability problems found using heuristic evaluation with various numbers of evaluators" From *How to Conduct a Heuristic Evaluation* (Nielsen, 1995)

Before Nielsen and Molich conducted their study the collection of usability guidelines contained as much as one thousand rules to follow. It goes without saying that this would be seen as intimidating by many developers, and so they managed to cut the complexity down to nine usability heuristics which are detailed in the aforementioned paper.

1. **Simple and Natural Dialogue:** Dialogues should not contain irrelevant or rarely needed information. All information should also appear in a natural and logical order.

2. **Speak the User's Language:** The dialogue should be expressed clearly in words, phrases, and concept familiar to the user rather than in system-oriented terms.

3. **Minimize the User's Memory Load:** The user should not have to remember information from one part of the dialog to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

4. **Be Consistent:** Users should not have to wonder whether different words, situations, or actions mean the same thing.

5. **Provide Feedback:** The system should always keep the user informed about what is going on by providing him or her with appropriate feedback within reasonable time.

6. **Provide Clearly Marked Exits:** A system should never put users in a situation that has no visible escape. If a user navigates to the wrong system function there should always be a clearly marked "emergency exit".

7. **Provide Shortcuts:** Features that increase the learnability of a system are often cumbersome to an experienced user. Shortcuts that are unseen by a novice user may often be included in a system so that both inexperienced and experienced users may use it to the best of their potential.

8. **Provide Good Error Messages:** A good error message should be defensive, precise, and constructive. It should never blame the user, provide the user with an exact cause of the problem and suggest what to do next.

9. **Error Prevention:** Designing to prevent a problem from occurring in the first place is a better solution than a good error message.

As more attention is focused on the web, several different heuristics have been developed specifically with emphasis on web design issues (Sharp, Rogers, and Preece, 2011). Many of them are very similar to the existing heuristics, but may focus more on content than Nielsen's heuristics do. Some examples include to provide a clear site name, writing meaningful labels, using meaningful icons and to always highlight the current section in the navigation.

HE has been used as one of the most cost-effective and efficient forms of usability evaluation. It has also been tweaked to fit with specific domains, such as e-learning. In the study by Ssemugabi and Villiers it was seen that people who were experts in both HCI *and* the subject matter are able to uncover more usability issues. Put up against a survey conducted with 61 students, the experts uncovered 77% of the problems, while the students found 73%. Seeing as the number of experts was only comprised of 4 people, this data implies that end users are not as good at identifying usability problems.

2.2.3 The Seven Stages of Action

In his book *The Design of Everyday Things*, Norman (2013) introduces the idea of two gulfs: the Gulf of Execution, where people try to figure out how something operates and the Gulf of Evaluation, where they try to figure out what state something is in and whether the actions performed completed their goal. Bridging these gulfs is the role of the designer, for instance by use of visible elements such as a drawer handle that clearly signifies that it should be pulled. It is when operations fail that the Gulf of Execution occurs: what other operations could be done to complete the task? The Gulf of Evaluation is defined by the amount of effort it takes to interpret the physical state of a device and to determine how well the expectations and intentions have been met. The size of the gulf changes based on how well design elements are implemented. If the information about the state of the device is simple to interpret, easy to understand and matches the way a person thinks about the system then the gulf is easily bridged.

To cross these bridges a number of stages can be applied. These are the stages that Norman has coined *The Seven Stages of Action*. Each action bridge the gap between what the aim is, and the physical actions that allow the goals to be completed.

1. **Goal**(form the goal)

Once the actions that will be performed are decided upon, they must actually be set to life in the stages of execution. These are illustrated on the left side of (Figure 2.3).

2. **Plan**(the action)
3. **Specify**(an action sequence)
4. **Perform**(the action sequence)

After these stages have been completed the three stages of evaluation follow. These are illustrated on the right side of (Figure 2.3).

5. **Perceive**(the state of the world)
6. **Interpret**(the perception)
7. **Compare**(the outcome with the goal)

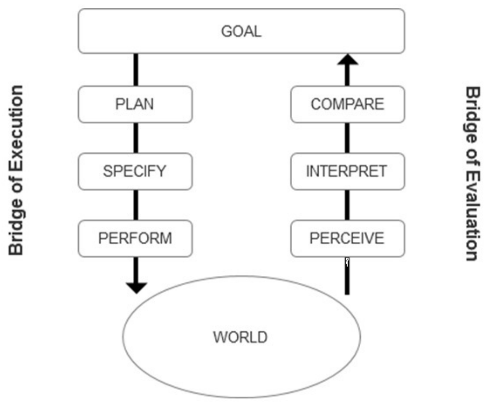


FIGURE 2.3: "The Seven Stages Action Cycle" Recreated from *The Design of Everyday Things: Revised and Expanded Edition* (Norman, 2013)

Many of the activities performed in these stages are not necessarily conscious actions. The goals tend to be, but even they may be subconscious. As such it is possible to cycle through many stages, while not being consciously aware that they are taking place. Only when a problem or a new situation that distracts from what is normal arises, do we need to give conscious attention to which stage needs to be processed. The stages are not necessarily performed in any given order, and most behavior does not require going through each stage, but can processed in several sequences. Based on this idea there are two main types of behavior, goal-driven and event-driven behavior (Norman, 2013).

The action cycle can start from the top, beginning with a new goal, meaning it is a goal-driven behavior. The cycle then starts with the goal and follows through the three stages of execution. It is also possible to start from the bottom, triggered by some event in the world, meaning it is an even-driven behavior. In this case, the cycle starts with the environment and then goes through three stages of evaluation. The act of turning on the light to be able to read is an example of an event-driven behavior. The sequence starts with the world, causing evaluation and ultimately the formulation of a goal. The difficulty reading is caused by an environmental event: lack of light, which causes a problem to occur and distracts from the original goal of reading. This leads to a subgoal of getting more light. Reading itself might also be a subgoal, as the true goal could be an even higher level goal such as learning a language, which had the goal of being able to communicate with a friend. The hierarchy of goals would then be: communicate with a friend; learn a language; read textbook; get more light.

The seven stages give developers another guideline for creating new products or services (Norman, 2013).The gulfs are good places to start looking for opportunities for product enhancement, either in the stages of execution or evaluation. Innovation can be an effect of the incremental enhancement of existing products, while radical ideas can come into existence by reconsidering the goals, and always asking what the real goal is. Insights gathered from the seven stages will also lead to seven fundamental principles of design, many of whom are also found in the earlier paragraphs.

1. **Discoverability**: It should be possible to determine what action are possible, and to ascertain the current state of the device.

2. **Feedback**: There should be full and continuous information about the results of actions made. After an action has been executed, it should also be easy to determine the new state of the product or service.

3. **Conceptual model**: The design should project all information needed to create a conceptual model that leads to understanding and a feeling of control. The model should also enhance both discoverability and evaluation of results.

4. **Affordances**: The proper affordances should exist to make the desired actions possible.

5. **Signifiers**: Effective use of signifiers ensures disoverability and that the feedback is well communicated and intelligible.

6. **Mappings**: The relationship between controls and their actions follows the principles of good mapping, enhanced as much as possible through spatial layout and temporal contiguity.

7. **Constraints**: Providing physical, logical, semantic and cultural constraints guides actions and eases interpretations.

2.2.3.1 Comparison of Heuristics, Design Principles and the Seven Stages of Action

After individually researching these three ways of evaluating and creating better usability in a product, it is clear that there is some common ground between them. While heuristics are traditionally used for evaluating existing products, and design principles for creating them, the relation between these theories are clear.

The usability heuristic of error prevention is concerned with preventing errors before they happen, rather than providing error messages when they do. It is then not a long stretch to say that the design principle of constraints is closely tied to this idea. Having constraints in place will prevent a certain kind of action, and ensure that a user cannot perform the wrong action, thus removing the need for an error message.

Affordances are mentioned in both Sharp, Rogers, and Preece (2011) and Norman (2013). In his book Norman clarifies the use of the term, and says that the industry adopted the term in a slightly different way than what was intended. An affordance of touching exists on an entire screen, but signifying *where* to touch is a different matter. It is not always enough to have an affordance in place. Therefore it is also important to have a signifier as well, so there is no room for confusion. While affordances represent the possibilities for how an agent (a person, animal, or machine) can interact with something. Signifiers are signals. Signals can be labels or signs placed in the world, such as the labels on doors instructing a user to "push", "pull" or "exit". They can also be arrows or diagrams indicating how to interact with an interface, such as in which direction to gesture. Some signifiers may also be the perceived affordances, such as the handle of a door. One thing to watch out for is misleading signifiers, that sometimes are accidental, but can also be used to try and keep people from performing an action which they are not qualified to do (Norman, 2013).

2.2.4 Accessible Design(Universal Design)

During the observation phase of the project the need to consider not only on what happens on the screen, but also what happens around it became clear. Who exactly is using this product? The answer, in terms of the Bybanen, is that it is not possible to narrow down the type of users that might wish to use it. Public transportation is something that most people will need to use. Therefore it is vital to consider users with any disability, and ensure that it is accessible to these individuals as well.

There are four main types of disability. Vision, Hearing, Motor and Cognitive. Skyss have already taken some of these disabilities into account when designing their ticketing machines, and interactive bus schedules. For instance it is possible to press a button on the interactive schedules, and it will be read out loud for a person who is blind or has poor vision. A separate, smaller ticketing machine has also been created for improved wheelchair accessibility. This solution does alleviate the problem, but also requires the company to spend a lot of money on creating two separate solutions, instead of creating a single machine that also solves the issue. An example of such a solution is the ticketing vending machine that was developed for the Austrian Federal Railway(Österreichische Bundesbahnen). Instead of having a separate machine the TVM allows a user to press anywhere on the large vertical screen, and have the interactive screen appear at the appropriate height for them. (See Figure 2.4)

FIGURE 2.4: "Accessible TVM developed for the ÖBB" (Siebenhandl et al., 2013)

Another example is a large format book, which is produced so that a user with vision defects can read it. This kind of book would be found too large for a user with good vision. Here a product such as an eReader will allow the text to be changed by the user, and will allow the same book to be read by different types of users. Some eReaders can also generate a synthesized voice, making them accessible to blind users as well.

The importance of universal design is further exemplified by the choice that the Norwegian government made to begin enforcing a set of statutory demands in the ICT sector. The law states that any new ICT-solutions that have been ordered or bought after the 1st of July must be up to the universal design standards. ICT-solutions purchased before this time do not have the same restrictions. That is, not until the 1st of January 2021 when all existing solutions must be universally accessible. The purpose of regulation is to demand that ICT-solutions are designed to be accessible to all, independent of any disabilities. All residents should have the same opportunities to take part in the information society. The regulation mainly apply to web solutions and machines, such as ticketing machines and ATMs (Difi, 2014).

This means that the Skyss TVM will have to conform to these standards by 2021. Such machines will have to follow at least ten standards. These standards cover the major categories of disability, and of functions within ICT-machines. An example is to store information regarding the users personal preferences of dialogs and interface on machines that have identity cards (Difi, 2014). The ten standards are thoroughly documented, but cost money to download and will therefore not be assessed further in this thesis.

2.2.5 Distributed Cognition

The theory of distributed cognition, as described by Hollan, Hutchins, and Kirsh (2000), seeks to understand the organization of cognitive systems. In difference from other cognitive theories, it extends the reach of what is considered cognitive. It reaches beyond the individual to include interactions between people and with resources and materials in the environment. As such a process is not seen as cognitive simply because it happens in the brain, nor is a process non-cognitive simply because it happens in the interaction among many brains. In other words not all cognitive events must happen within the head of an individual. Distributed cognition looks at a broader class of cognitive events, an example being the examination of the memory processes within an airline cockpit. A theory based on individual memory by itself is in itself insufficient to understand how the memory system works. The reason being that memory involves a rich interaction between internal processes, the manipulation of objects and the traffic in representations among the pilots. Based on these principles three types of distribution of cognitive process emerge:

1. Cognitive processes may be distributed across the members of a social group.
2. Cognitive processes may involve coordination between internal and external (material or environmental) structure.
3. Processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events.

Traditionally information processing psychology places a gulf between the inside and outside, and then bridges the gulf. These bridges are crossed with transduction processes that converts external events into internal symbolic representations. This implies that the computer and its interface are outside of cognition, and are only brought inside through symbolic transduction. As opposed to the seven stages of action (see Section 2.2.3) distributed cognition does not set a gulf between "cognitive" processes and an "external" world, so it does not attempt to bridge such a gulf.

Hollan, Hutchins, and Kirsh (2000) proposes the use of distributed cognition as a new foundation for human-computer interaction. The proposition is a framework for research that combines ethnographic observation and controlled experimentation as a basis for theoretically informed design. The framework behind distributed cognition commits to the importance of observation of human activity "in the wild" and analysis of distributions of cognitive processes.

While not explicitly used in this thesis, cognitive distribution affects the methodology used, as one of the main points of focus is an observation conducted at the beginning of the research.

2.3 Similar Research

This section looks at existing research or projects that have some of the same characteristics as this one. The section is divided into categories based on the research literature.

2.3.1 Design Principles

HCI Design Principles for eReaders: Pearson, Buchanan, and Thimbleby (2010) studied three different eReaders and discussed specific design principles for such devices. Usability issues were found through the use of heuristic evaluations, which admittedly could not uncover every single issue, but was said to be a good precursor to further work. Based on the results a set of design principles were used to think about the design, rather than using the conventional mode of HCI which is based on empirical experiments. Of the existing principles, only one was used in the study, consistency. In this case it was used to determine if the buttons of the eReaders were well labeled and always performed the same functions. An interesting concern that was uncovered during the heuristic evaluation was that of completeness.

Completeness is a principle that relates to how a physical item relates to its digital equivalent. In terms of a book there are certain actions that are not possible to re-create, such as folding, ripping and flicking a page. Although these things are difficult to implement, there are certain things that should be incorporated to not leave the product feeling incomplete. For instance it is common to add a bookmark to a physical book, and the be able to see where this bookmark has been placed while still reading. In an eReader this has not been adapted in the same way, only allowing the user to see it if they happen to be reading the page that the bookmark is on. Designers should make certain that tools and actions performed within a device should mimic the actions that can be performed on paper (or whichever physical counterpart there is), unless the solution is inherently inefficient in a digital interaction (Pearson, Buchanan, and Thimbleby, 2010).

User-Interface Design Principles for Experimental Control Software: Boring (2001) studied the use of experimental control software (ECS) and the way design principles would benefit them. There was a lack of clear usability guidelines for ECS and such poor usability could often result in a lack of statistical significance in the results. The biggest issue being if poor usability would lead to an experimental artifact, and one that would be falsely seen as a genuine effect.

Using six experimental scenarios he identified two subgroups of users, the experimental participant and the experimenter. After this process a set of usability issues were identified, and the importance of them was also recorded in terms of how many of the scenarios revealed the issue. These issues were then used to create guidelines that would increase the usability of psychology and HCI experiments.

Principles of Human-Computer Interaction in Game Design: Cai (2009) analyzed and used HCI theory as a tool in order to detail principles from the field in the design of games. Human-Computer Interaction is the core of many game elements. Although graphics, animation and sounds are integral parts of a game, it is the interaction that is the real focus. In order to improve this interaction, the deep links between HCI and game theory are used to guide the design of a game.

The principles outlined are those of simple, natural, friendly and consistent game interfaces.

- *Simple Principle* The simple principle states that the process of Human-Computer Interaction in a game should be as simple as possible. The interactions should not be too complicated, as this would make it more difficult for the player to grasp and control the game. Even though it should be simple, it should also offer a sense of entertainment experience, rather than being completely efficient. If an action is too simple to achieve it might remove the feeling of achievement when completing a goal in the game.

- *Natural Principle* The natural principle states that the process of Human-Computer Interaction should be as close to the player's life experience and cognitive habits as possible. This is to lead players into quickly linking their real life experiences with those in the virtual world.

- *Friendly Principle* The friendly principle refers to the contents and forms of information that is output from the game, and that it must help a players' understanding. This principle is reflected in multiple points:

  - *Reasonable forms of information*: Information should be logically grouped, and different arrangements and areas should show important information and secondary information.

  - *Giving automatic corrections or tips to the input, do not conform to the rules of the game*: The design of the game should have corresponding input process to player's irrational input, so as to prevent an unreasonable operation to occur.

  - *The main state and information must be given*: As a player has to make a lot of decisions according to a specific circumstance during the game, the design should give the players as much information as possible in a relevant sense. Such as visually or audibly giving the player a signal that their character is hurt, so that they do not have to pay attention to their remaining health in the game.

- *Provide a comprehensive help system*: Some designers assume that the general player will be just like them, and be able to play the game smoothly without much help. This is usually not the case, and so a form of help system should be in place. It is worth noting that the traditional way of help is quite different in the game world, as it will interrupt the play progress and so a more "invisible" way of help should be implemented.
- *The operation which can be configured and many operations for the same function*: The habits of players vary greatly, and so the way the game is operated should be open for configuration. Even if the default way of operation is in line with the players' expectation, an alternative shortcut would give them more options and also increase the depth of a game.
- *Adequate feedback*: Any legitimate operation from a player should be given feedback from the game. The feedback can for instance show the player the result of an operation and give them a sense of achievement.

- *Consistency Principle* The consistency principle states that the output of the computer and the input of the player should maintain consistency. Not only by appearance, but also logically. Let us say there are two types of doors in a game, where one can be opened, while the other may not. In this case there should be a clear difference between them, to ensure that a player does not try to open many doors that are not meant to be used.

2.3.2 Self-Service

Accessible self-service kiosks. Hagen and Sandnes (2010) developed a prototype for a universal self-service kiosk, meaning that it should be accessible to anyone, irrespective of physical and cognitive abilities. The prototype takes into account a users height, length from the screen and their accuracy level in order to tailor the experiences for each individual. Several other issues were also identified, and as the prototype was only of a low-fidelity it was concluded that more work needed to be done in order for it to be feasible.

A User-Centered Design Approach to Self-Service TVMs. Siebenhandl et al. (2013) used a UCD approach to further develop a ticket vending machine(TVM) for the Austrian Federal Railway. Their focus was on usability, and also the UX of different user groups. In the study the context of use was found by observing and interviewing users, going through literature, having accessibility workshops and regular meetings. This data was then analyzed and turned into requirements, which formed the basis for the first prototypes. The observations took place on four different occasions (weekday-weekend, urban-rural stations) and users were categorized with respect to age and genders. In all, over 250 participants were part of the entire project. The limitations of the project was in respect to integrating real currency verification, and valid tickets. Future research will then include a field test of a future fully functioning TVM.

2.4 Theoretical Framework

In summary, the theoretical framework consist of several aspects within HCI and self-service technology. The main focus centers around the creation of self-service design principles and of an interface to validate the effectiveness of these principles.

Chapter 3

Methodology and Research Strategy

This chapter presents the methodological framework that the thesis is based upon. The research project followed a multimethodological approach. The chapter also describes the research methods used and why they were chosen.

3.1 Methodologies

3.1.1 Systems Development Research Methodology

The multimethodological approach outlined by Nunamaker, Chen, and Purdin has been utilized in the thesis. A research follows the pattern of "problem, hypothesis, analysis and argument" and in this view the result of the analysis can become the argument (or evidence) of the initial hypothesis. The thesis will both contain observation and prototyping and therefore the approach was deemed to be a good choice, as the validity of the potential novel design principles should come by analyzing the resulting system/interface.

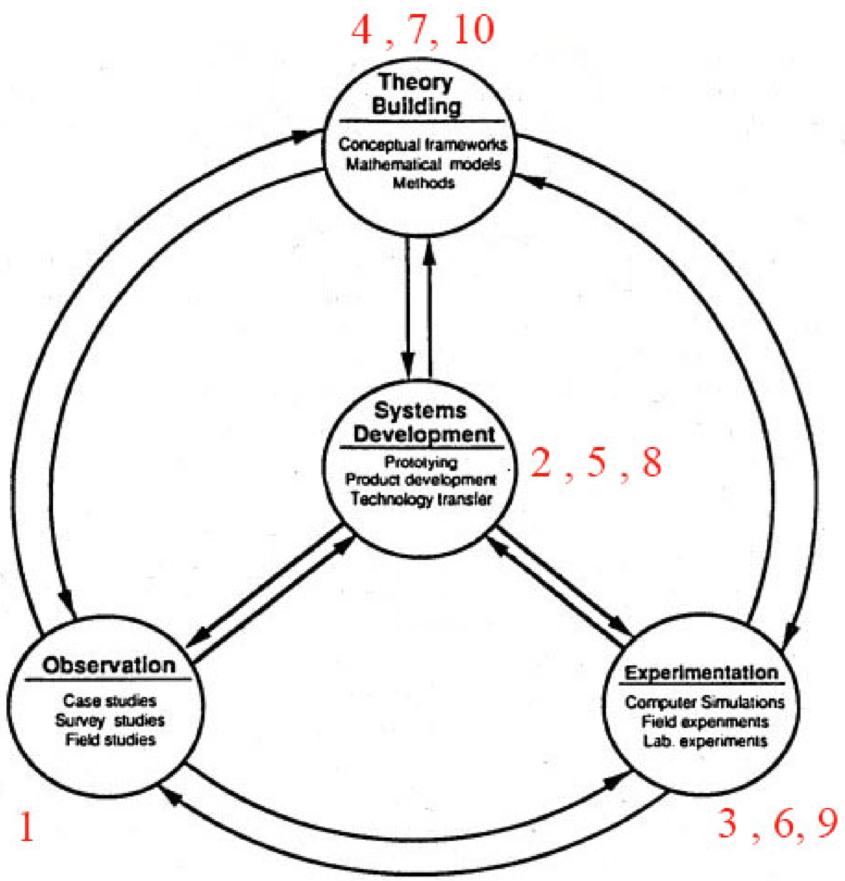


FIGURE 3.1: "A Multimethodological approach to IS Research". From "Systems development in information systems research" (Nunamaker, Chen, and Purdin, 2001)

Figure 3.1 illustrates that the approach consists of four research strategies that are tied together very closely. The numbers represent the order in which the process will be followed. Step one is to perform an observation to gain insight in the self-service technologies. Step two is to create an initial prototype based on the original system, after which step three will utilize focus groups and user testing to gather data. In step four the theory will be constructed in form of design principles. The systems development part normally consists of five stages: concept design, constructing the architecture, prototyping, product development and technology transfer (Nunamaker et al., 2001). As the thesis focuses on the interaction design of an application, the latter stages will not be considered, instead there will be more focus on the construction and evaluation of the prototypes. After creating the novel design principles the prototyping stage will be reiterated along with the other stages. The final stage will be to build theory on whether or not the novel design principles has the potential to be beneficial to self-service technologies.

3.1.2 Other Options

In the process of finding a suitable methodology for the thesis, using the design research approach was also considered. It involves the use and performance of designed artifacts to understand, explain and to improve on information systems (Vaishnavi and Kuechler, 2004). This approach could also be relevant for this thesis, but as the focus will not be on introducing a novel artifact it was believed that another methodology would be more appropriate.

3.1.3 Developing the Interface

When developing the prototype Personal Kanban (PK) will be used to structure the work. PK is a variant of Kanban that is intended for a single developer. The reason for choosing this system is that many practices in agile are meant for improving workflow in teams, and is not something that a single developer needs to consider. Even though creating a prototype is not strictly the same as system development, it is still believed that following a method such as PK will be helpful to the structure of the project.

A kanban is basically a tool that helps the visualization, organization and completion of work (Benson, 2013). There are four main steps involved in building a PK that will be used to maintain a good workflow.

Establish Your Value Stream: The Value Stream is the flow of work from the moment the work starts and until it is finished. The easiest way of doing this is by dividing tasks into a backlog, a doing list and a done list and is depicted in figure 3.2.

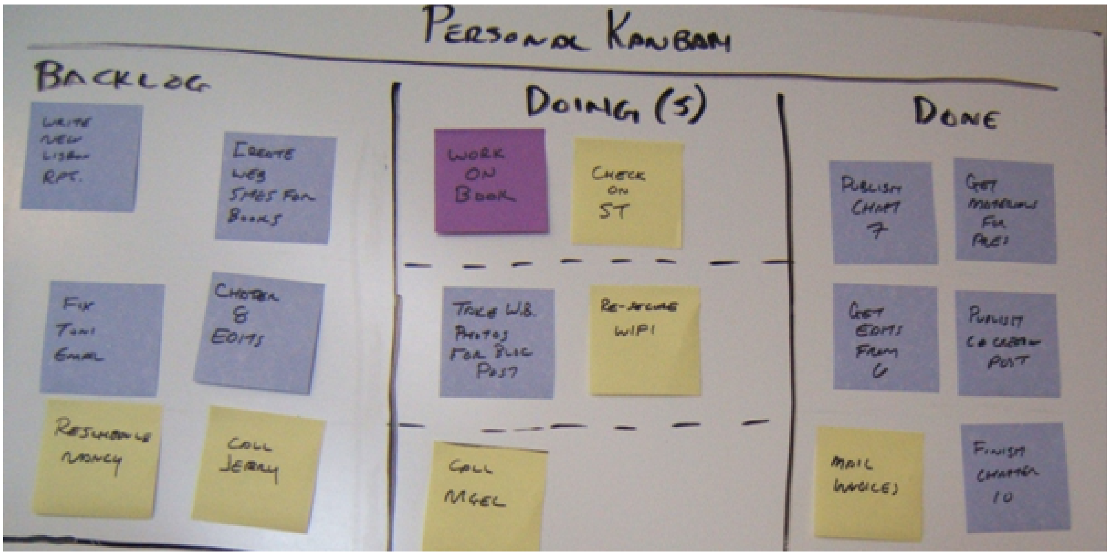


FIGURE 3.2: The Visual Flow of Kanban. From (Benson, 2013)

A white-board is often used for this purpose, but as a single developer Trello board will be utilizied in this project. A Trello board is practically a virtual white-board, and provides the ultimate flexibility so that changes can be made at any time during the process (Fog Creek Software, 2013).

Establish Your Backlog: The second step is to create a backlog for the project. This backlog contains all the work that is not yet done. It is here all the tasks that need to be completed for the project to be finalized will go (Benson, 2013).

Establish Your WIP Limit: This is the maximum amount of tasks that are classified as "Work In Progress". It is important to not have too many things on the table that are not finished as this can be very stressful in a development situations. For this reason one should add a maximum number of tasks to the doing list at the beginning of the development process. This number can then be modified later in order to find a point where the developer is doing the optimal amount of work at the optimal amount of time (Benson, 2013).

Begin To Pull: This step is basically to actually start working. In other words to start pulling tasks from the backlog and into the other columns of the board. Beyond this step the developer will focus on prioritizing current work, refining the value stream and getting things done (Benson, 2013).

These principles were also incorporated in other aspects of the project, and the Trello board was utilized when structuring the written tasks and during the analaysis phase of the project.

3.1.3.1 Fidelity

The aim is to create high-fidelity prototypes, so that it will closely resemble the look of a final product. The advantages of such a prototype is that it is fully interactive, has more functionality and more clearly defines the navigational scheme than a low-fidelity prototype would (Sharp, Rogers, and Preece, 2011). It does take longer to create this form of prototype, but considering the scope of the project it is time well spent. As seen in table 3.1 there are many more advantages and disadvantages of both variations of prototyping. The disadvantages of high-fidelity prototyping are of little concern in relation to the timespan of the thesis it is deemed to be the best choice.

Type	Advantages	Disadvantages
Low-fidelity prototype	• Lower development cost. • Evaluate multiple design concepts. • Useful communication device. • Address screen layout issues. • Useful for identifying market requirements. • Proof-of-concept.	• Limited error checking. • Poor detailed specification to code to. • Facilitator-driven. • Limited utility after requirements established. • Limited usefulness for usability tests. • Navigational and flow limitations.
High-fidelity prototype	• Complete functionality. • Fully interactive. • User-driven. • Clearly defines navigational scheme. • Use for exploration and test. • Look and feel of final product. • Serves as a living specification. • Marketing and sales tool.	• More expensive to develop. • Time-consuming to create. • Inefficient for proof-of-concept designs. • Not effective for requirements gathering.

TABLE 3.1: The Relative Effectiveness of High/Low-Fidelity Prototypes. From (Pearson, Buchanan, and Thimbleby, 2010, p. 396)

There are also some compromises one have to be aware of while selecting a particular style of prototype. They can be divided into horizontal and vertical prototyping. A horizontal prototype will provide the user with a wide range of function, but little detail. The vertical prototype will provide the user with a lot of detail, but only for a few functions (Sharp, Rogers, and Preece, 2011). Since the redesign of the ticketing machine does not have a wide variety of functions, the best choice will be a vertical prototype. In this way it should be possible to thoroughly test the potential usability of a finished design.

3.1.3.2 Prototyping Tool

To create the several iterations of prototypes the software Axure RP has been used (Axure Software Solutions, 2013), a tool that enables the creation of interactive prototypes. It can generate interactive HTML wireframes or UI mockups without the need for coding, and can also design interfaces that can be shown directly on a mobile device. This made it easier to gather valuable data about how the prototype functions, as potential users were asked to perform certain tasks that mimic a real life situation. It is also worth noting that the Axure software is used by as much as 50000 design and business professionals. The company claims that people who use their product include Disney, H&M and Nike.

3.2 The Collection of Data

In order to come up with new design principles for self-service devices it is important to figure out what a user expects from such a device. To do this a set of data should be collected and be analyzed to extract new information about the field.

A data generation method is the means by which empirical data or evidence can be produced. This data can be either quantitative or qualitative. Quantitative data mostly consists of numeric data, while qualitative data is all other types of data such as words, images and sounds (Oates, 2006).

Deciding the validity of a design principle is a subjective matter, and therefore the gathering of qualitative data has been chosen as the main focus. There are many methods to gain such data, and some methods are commonly associated with particular research strategies. Even though this is the case, the use of more than one data generation method enables us to see different sides to a phenomenon of interest. Both observation, usability testing and focus groups will be used, which is also called method triangulation and enhances the validity of the findings (Oates, 2006). In this way it is possible to corroborate the things people say they do with the things that they are observed doing.

3.2.1 Observation

Observing does not only involve looking, but can also involve other senses such as listening and smelling (Oates, 2006). In this thesis observations will be used to look at the behavior of people using the self-service ticketing machines created by Skyss. There are a wide range of approaches to observation, and as shown in table 3.2 they can be analyzed by placing them on a number of spectrums.

Highly systematic observations of pre-defined types of events	↔	Observations of anything and everything
Narrow concentration on particular type of event	↔	Broad focus
Observer takes no part in the proceedings	↔	Observer participates fully in the proceedings
Fact of observations taking place is known to all	↔	Fact of observations taking place is known to none except the researcher
No explanation, or false explanation, given for presence of observer-researcher	↔	Full explanation given for the presence of the observer–researcher
Short duration – could be as little as 5 minutes	↔	Long duration – possibly years
Record-keeping uses only simple note taking	↔	Record-keeping uses technology (e.g. audio tape, camera, stop watch, two-way mirror, computer program)
No feedback given afterwards to the observed	↔	Full feedback given afterwards to the observed

TABLE 3.2: "Different kind of Observations" From *Researching Information Systems and Computing* (Oates, 2006, p. 203)

The main distinction here is between overt and covert research. In covert research the people who are observed do not know about it, and the observer will try to not bring any attention to himself. In overt research the people know that they are being observed, and because of this they can give consent to the research being done. Both methods have their drawbacks, which is important to take in to account. A covert observation type has been chosen, as people tend to modify their behavior when they are aware of being observed. This is known as the *Hawthorne Effect* and is important to avoid in order to get the most valid data (Oates, 2006). The main drawback of a covert observation is the ethical aspect of observing people without their consent. In this case it was considered acceptable as it will be performed in a public space, and will not be obtrusive, so as to not alarm any of the people being observed.

3.2.1.1 Ethical Considerations

Since a covert form of observation will be employed, some ethical aspects come into consideration. There are four main areas of ethical principles to look into: whether there is harm to participants, whether there is a lack of informed consent, whether there is an invasion of privacy and if there is any deception involved in the research (Bryman, 2012).

If the research is likely to harm the participants it is regarded by most as unacceptable. In this study there will never be a situation in which the participants will be put in any physical danger, and the subject matter is not of a sensitive nature so the emotional impact swill be of no significant value. In other words there should not be any harm to the participants as a result of my study.

The issue with informed consent might pose a problem, as the very nature of covert observation transgresses this principle (Bryman, 2012). The users are involved in the study whether they like it or not and a proper covert study cannot be performed if the observer is know to the people being observed. Given that the observation will take place in a public space it is entirely legal to study what others do while in that space and so the covert observation technique is deemed suitable. For the interviews and potential focus group an informed consent form is used, and any participant will be required to sign it before the interview or focus group can begin.

Once again the use of covert observation can be seen as invading the privacy of others. It is therefore even more important to ensure the anonymity of the people observed. In the case of the Skyss ticketing machine it is assumed that the use of such a device is not of a sensitive nature to most users. The assumption is based on the fact that such a topic is on the surface, and not intently a personal matter.

The last point, deception, is something that occurs when a researcher represents his work as something other than what it is (Bryman, 2012). Deception is not relevant for this study and therefore the ethical aspects of it is not concerned.

3.2.2 Focus Groups

A focus group is a method of interviewing that involves more than one, and usually at least four interviewees (Bryman, 2012). A focus group will be formed in order to gather an understanding about how the prototype is perceived. This data will then be used to improve on the prototype, and several focus groups will be completed until a final version of the prototype is ready. As with semi and unstructured interviews, it is important that the moderator is not too involved, but that he may need to respond to specific points that are not being addressed (Bryman, 2012). This should not be much of an issue, as the prototype will clearly be explained to all participants before the discussion commences.

In the interview stage there were a total of 20 respondents that both evaluated the initial prototype and then answered a questionnaire based on their experience. The goal was then to bring back most of these respondents to form five focus groups. As most of the involved users have actively been using the actual system in their daily lives it is likely that they will have a lot to discuss. This means that it is most preferable with a smaller focus group, and so each group should contain no more than 3-4 people. It is also less likely that respondents will have the confidence to speak their opinion in a large group (Bryman, 2012). A good reason to not have too many focus groups is the sheer number of transcription pages they produce. For a single researcher it will both take a long time to transcribe and also to translate the information into valuable data.

3.2.3 Recording Tools

The video recording function on an SLR camera will be used to record the interviews. The main reason for this is because the interviewer has to be very alert while performing an open-ended interview. Other advantages of recording the video is that it allows more thorough examination of what people say and opens for the possibility of repeatedly going over the answers (Bryman, 2012). There are of course not only advantages, and one of the biggest drawbacks is the sheer amount of time it takes to transcribe the interviews. This time that is considered well spent, and will also make the analysis of the interviews easier. In addition a tool that supports this type of research and claims to be able to help analyze data to uncover subtle connections and justify findings will be used (QSR International, 2012).

3.2.4 Samples

The people who are observed can of course not be selected beforehand, but there is a need to sample the test subjects for the usability tests and focus groups. Most sampling in qualitative research includes purposive sampling of some kind. Purposive sampling is a non-probability form of sampling, meaning to not sample subjects on a random basis (Bryman, 2012). The different self-service systems have a very broad scope, and so the subjects could be of any age group and have very diverse backgrounds. The main focus will then be to get as varied a sample as possible.

In this thesis sample size data will be collected until theoretical saturation has been achieved. Theoretical saturation means that there is no new or relevant data seems to be emerging within the category, meaning that there is no perceived gain from doing more interviews or observations (Bryman, 2012).

In the different phases some form of personal information about the people involved will be registered, and this means that the ethical aspects of performing the studies needs to be considered as well. The participants have the right to withdraw at any time and to be anonymous, which is going to be respected.

3.2.5 Controlled Experiments

During the last phase of the thesis, the original and the improved prototype will be tested together in a type of controlled experiment. Controlled experiments is a well known approach that has been adopted from research methods in psychology. It has also become a widely used approach in the evaluation of interfaces. The question such experiments commonly answer are: does making a change in X have a significant change in Y? In this case X can be some features of the interfaces, and Y the time it takes to complete a task, or the users' subjective satisfaction of working with the interface (Blandford, Cox, and Cairns, 2008).

An assumption is made that there will be no difference between the designs, which is called the null hypothesis. It is by failing to prove this hypothesis that the evidence of a causal relationship between variables is made. In an HCI context, the changes that are made could be to interface features, interaction design, participant knowledge and so on. The variable that is changed is called an independent variable, while the variable that is measured is called a dependent variable (Blandford, Cox, and Cairns, 2008). For this thesis there will be two dependent variables. Times taken to a complete a set of tasks, and the subjective user experience for a set of users. The independent variables will be changes made to the prototype derived from the original Skyss-system.

An important aspect to be aware of while conducting a controlled experiment is the possibility of introducing confounding variables. This is a variable that is unintentionally varied between conditions of the experiment, and could affect the measured value without the experimenter realizing it (Blandford, Cox, and Cairns, 2008). If all users test two prototypes and test them in the same order, it could be that the users have learned something from the first prototype. This could then lead to the second prototype being favored as the users already understood something about the system. To minimize the effects of this confound, a within subject design approach is a good alternative. Even with this type of experiment a particular set of confounds should be taken into account, individual differences. It is of course not possible to control for all such differences, but it should be possible to control for the most likely factors such as age and gender. As such there will be a mix of male and female participants in each test group, and as far as possible an equal mix of age groups.

3.2.5.1 Within Subject Design

A within subject experiment involves each participant performing under all sets of conditions, while a between subject experiment has each participant only performing under one condition. The reason for choosing within subject design for this thesis is because the participants are required to compare differences between prototypes, and as such within subject design is essential (Blandford, Cox, and Cairns, 2008). Another reason is that the number of participants that can be recruited is limited, and would make a between subject design less useful.

The order in which the prototypes are tested is then imperative to reduce the effects of the confounds. As the participants will only be testing two different prototypes the structure of the experiment is fairly simple. One group will be testing prototype A first and then prototype B, while the second group will be testing prototype B first and then A. In this way the experiment is counterbalanced as to avoid any ordering effects that might occur (Sharp, Rogers, and Preece, 2011). In addition the experiment will be run on the same computer, with the same mouse for each test.

3.2.6 Criteria for Evaluation

In order to ensure the quality of the research it should meet the criteria for evaluation, the most prominent being **reliability**, **replication** and **validity**.

Reliability aims to see whether the results of a study a repeatable or not. In other words if the practices that have been used are consistent and can be used again to achieve similar results. Replication is very similar to reliability, and is something that happens if other researchers decide to replicate the findings from this thesis. It is not very common to do so, and is mostly valued by researchers working within a quantitative research tradition. Validity is in many ways the most important criterion and is concerned with the integrity of the conclusions that a research generates (Bryman, 2012).

As qualitative research is an important part of the study four other criteria that have been proposed for this type of research will be assessed. These being **credibility**, **transferability**, **dependability** and **confirmability**. Triangulation and the guidance of the supervisor for this thesis Viktor Kaptelinin will be followed in order to raise the credibility of this project. To ensure the transferability there will be thick descriptions of the group of people that have been a part of the study. In order to support the dependability of the project detailed records of the different phases of the research process are kept, such as field notes, interview transcriptions and so on. Lastly, to maintain confirmability, the point of view must be objective and not allow personal values to affect the research and findings.

3.3 Qualitative Data Analysis

Qualitative data analysis is not always a straightforward task. There are no clear cut frameworks on exactly how to do it, but most techniques involves abstracting the patterns and themes that are important to ones research topic from the gathered data (Oates, 2006).

3.3.1 Data Analysis

One of the most used frameworks for analyzing qualitative data is known as grounded theory, and was considered for this thesis. It focuses on doing field research, and then analyzing the data to see which theory emerges (Oates, 2006). Grounded theory has particular practices that it incorporates, which for instance determine how to select people and instances to include in the research. According to Oates this selection starts by identifying a single person (or instance) and then generating data based on just this one person. This does not coincide with the method of observing, and interviewing several people, and so grounded theory was decided against. Instead certain methods that are detailed in the book *Researching Information Systems and Computing* will be consulted(Oates, 2006).

The first step to analyzing is of course to read through all of the data to get a general impression. After getting through the data key themes that are present are identified. These could be: segments that have no relation to the research being done, and are not needed for the study. Segments that provide general information that will be needed to describe the research context, such as location and information about respondents. Then the segments that appear to be most relevant to the research questions (Oates, 2006).

Focusing on the last segment it is possible to focus on categorizing the segments and extracting further information from them. The categories will be chosen based on an inductive approach, meaning that they are obtained data, trying to have an open mind and not be influenced by previous experiences. The qualitative data will be produced during the focus group phase of the project.

3.4 Quantitative Data Analysis

Quantitative data means data, or evidence, based on numbers. As opposed to qualitative data, there is a wide range of established techniques for analyzing quantitative data.

3.4.1 Data Analysis

According to Oates the first stage of analyzing quantitative data is to establish which type of data that is being used. The reason for this is that different analysis techniques are better tailored for different kinds of data. The four main types of quantitative data are nominal, ordinal, interval and ratio data.

Nominal Data: is data which has no actual numeric value. One example is a questionnaire that asks the respondents' gender. As this gives no numerical value, the only analysis possible is one of frequency. Such data is sometimes called *categorical data.*

Ordinal Data: is data where numbers are allocated to a quantitative scale. This type of data is commonly categorized to Likert scale-based questions, where numbers are assigned to the range of responses. The responses "Disagree strongly", "Disagree", "Neither agree nor disagree", "Agree" and "Strongly agree" could for instance be coded 1, 2, 3, 4 and 5 respectively. An issue with this data is that it is hard to know by how much one response is greater or worse than another. Such data is sometimes called *ranked data.*

Interval Data: is data where measurements are made against a quantitative scale, where the differences between points are consistently the same size. For instance the difference between the years 2010 and 2014 is the same as the one between 1942 and 1946.

Ratio Data: is similar to interval data, but there is a true zero to the measurement scale. A person's age can be 0, and so could in theory their height. With such data it is possible to not only say that, 12 is the same interval from 6 as 12 is from 18, but also that 12 is twice as big as 2 and that 18 is three times as big.

The quantitative data gathered in this project will consist of nominal data from the focus groups and the observation, and of ordinal data from the user testing. The data is in turn made into charts that are easy to read, and will help organize and identify patterns in the data.

3.5 Summary

This chapter has described the research methods and explained the different data gathering and analytical elements used in the thesis. The overall structure of the work on the thesis is depicted in figure 3.3.

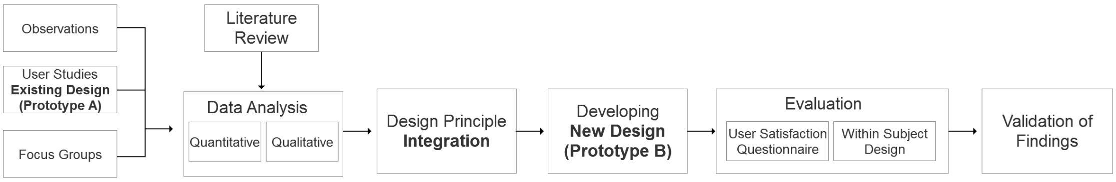


FIGURE 3.3: An overview of the project workflow

Chapter 4

Formulating a Novel Design Principle, Integration

In this chapter the process that led to the creation of a new design principle, and the further understanding of self-service technologies will be presented. The main focus has been on the Skyss self-service ticketing machines, and in order to get a wide view of its use, several techniques from the field of HCI have been utilized. This includes observations, focus groups, prototyping and task based user testing.

4.1 Observing the Users

The first phase of creating the design principle was to observe actual users of the Skyss ticketing machines. The specific type of observation performed is stated in section 3.2.1. In the following sections the method, results and analysis of the observation will be presented.

4.1.1 Method

The observations were conducted over a period of two weeks, from around 10 AM each day of observation, and took place at three different stations along the

Bybanen. Byparken, Nesttun and Danmarksplass. The reason for choosing these stops were based on both the pilot study and observations made during the actual study. Byparken is one of the final destinations of the Bybanen, and of course then the first. As it is placed in the city center it is also seen as the busiest. This meant that getting enough data would not be an issue, but that the constant rush of people would potentially be distracting. This constant stream of users was also positive in that it helped to blend in while noting down the data. Nesttun is quite close to the middle of the two final destinations, meaning that there were many travelers going both ways. This was sometimes problematic as there would not be time to get to the other side of the road in time to observe certain users. In addition it would look suspicious if someone were to start running back and forth all of a sudden. Initially Fantoft was chosen as the third location, but after an hour of observation without any travelers using the ticketing machine the location was changed to Danmarksplass.

To collect data the Evernote application was used, in which a form that contained the data points to be collected was created. These fields were gender, age group, which action was performed, which mistakes were made and how many, any design principle that could prevent such mistakes, time taken and whether or not the user was thought to be a tourist. To ascertain that someone was a tourist, language, and conversation topic would be taken into account. If uncertain they would not be included in the study. Each person observed also had an identification number. If there were any additional comments on an individual user, this was written down along with the ID number in order to easily connect the comments to the observational data. The application also directly transferred the results to the cloud, meaning that the data could be viewed on a computer screen straight after the observation was completed. In order to record how much time each user spent with the ticketing machine the timer function on a mobile phone was used. The timer was started when the user first started interacting with the screen, and stopped when the tickets were administered or a message indicating that the Skyss card was filled appeared on screen.

One of the most important parts of the observation was as mentioned in section 3.2.1 to avoid the Hawthorne Effect. This effect is commonly referred to as an increase in productivity, or even some other outcome during study which is caused by participation in the study as such (Wickström and Bendix, 2013). In this case that would mean a change in the results based on the fact that a user is aware of being observed. There were times during the observation when I was approached by people who asked for help, as I was standing nearby. I then felt obliged to assist them with completing the task, and as such the data gathered from these users were removed from the study as the data was compromised by my assistance.

4.1.1.1 Exploring the Questions

In order to get the most valuable data from the observation, a set of questions that aim to cover all of the aspects of the system were formed. The data gathered was then used to try to answer these questions and to shed some light on where the usability issues lie.

The main question for the observation is *"What are the most prominent usability issues with the Skyss ticketing machine?"* As this is a very broad question it will be useful to create a list of sub-questions that aim to discover the specific faults and successes of the machine.

- Are the main features visible and easy to use?
- Where do users make mistakes?
- Is the system efficient?
  - How much time does a general user need to complete a transaction?
  - Is there a noticeable difference between genders or age groups?
  - Are there major differences between various stops?

4.1.1.2 Validity

As mentioned validity is a very important criteria for a successful analysis. Is the observation and its outcomes valid? This is something that can't be guaranteed when it comes to observations. People do not have the same mindsets and this means that different people could get different results from the same observation (Oates, 2006). To strengthen the validity of the observation the data gathered from focus groups and from user testing has been triangulated. In this way the findings derived from the observation can be confirmed by other forms of data generation.

4.1.2 Results and Analysis

The information gathered from the observations were transferred to an Excel spreadsheet in order to easily extract quantitative data from it. As mentioned in section 3.4 this data was turned into different types of charts. In order to see if there were any major differences between the different stops a chart was defined for each of them.

In terms of demographics both genders were represented almost equally among everyone observed (see Figures B.1a, B.1b, and B.1c). This meant that it would be possible to see if there were any noticeable differences between the genders. The same held true for most of the age groups. It is believed that there would be more young users if the observation had been successful at Fantoft, as there are mostly students living near that location. These numbers can also indicate that the majority of people who use the ticketing machines are either adults or elderly. There were almost no children that used the ticketing machine by themselves (see Figures B.2a, B.2b, and B.2c).

Among all the people observed the average time to complete a transaction was 37 seconds (see Figure 4.1). Considering that the Bybanen has no option of paying on board, and will not wait at the station for very long, there is a definite room for improvement. This was further exemplified by the observation of people who would go aboard without a ticket, as the Bybanen was about to leave.

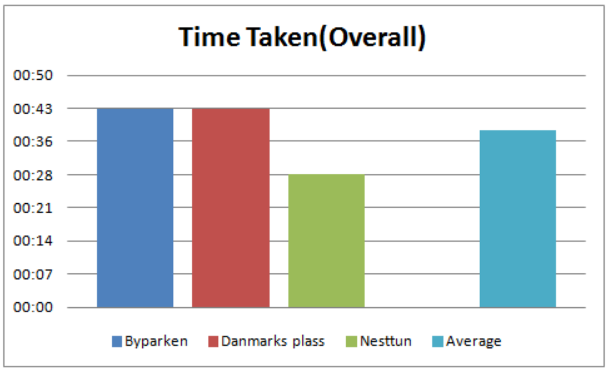
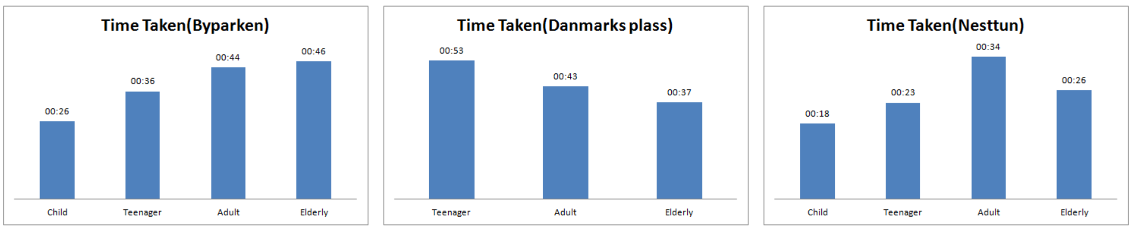


Figure 4.1: Times taken in average, overall

For each station the time differences between different age groups were varied, where elderly users spent longer time with the machine when observed at Byparken and Teenagers (see Figures 4.2a, 4.2b, and 4.2c).



(a) Times Taken: Byparken (b) Times Taken: DP (c) Times Taken: Nesttun

Figure 4.2: Times Taken: Three Stations

An assumption was made prior to the observation, that elderly users would take longer to complete a transaction than others users. This proved to be false, as the elderly users observed in this study on average completed the task slightly faster than both teenagers and adults. One potential reason for this is that most of the elderly users bought single tickets, which is inherently faster than refilling a travel card. Another possibility is that the elderly observed might have more experience with the ticketing machine or with the particular type of ticket that they were purchasing. The time differences were still only a few seconds with the exception of the child users, who had an average time of 21 seconds (see Figure 4.3). With such a small number of young users there can be no claim that they consistently perform better than other age groups.

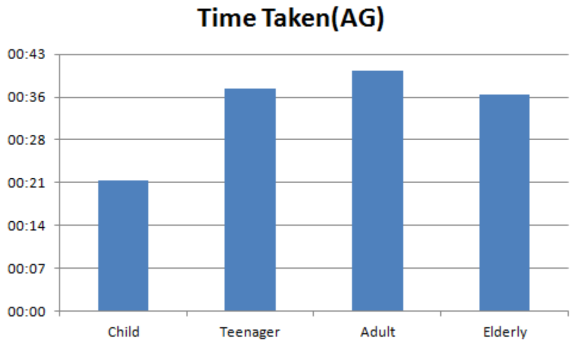


Figure 4.3: Times taken overall by different age groups

There was a small difference between genders, where male users had an average time of 42 seconds, while female users had 37 seconds. Although this is more than with the age groups, it is still not significant enough to say there is a correlation between gender and time taken (see Figure B.3). As mentioned earlier, each person observed was also assessed as to whether or not they were a tourist. On average the tourists had an average time of 1 minute and 46 seconds, while locals had an average of 36 seconds (see Figure B.4). This makes it quite clear that tourists have a harder time using the machine. Some of the reason for this is most likely that the tourists have never used the machine before, and as such has to learn how to use it. It is also then clear that the machine has potential for improvement in regards to learnability, and how easy it is to grasp the different functions.

Between the three stations there was a varying rate of errors, where the users made an average of 0,8 mistakes at Danmarksplass, around 0,7 at Byparken and only 0,35 at Nesttun (see Figure 4.4). A possible reason for the low error rate at Nesttun might be that there were not as many people traveling from that station. They did not have to worry about rushing to pay, so that the next person in line could use the ticketing machine. Removing this factor of urgency inherent in many self-service terminals might then have caused a drop in the rate of mistakes made.

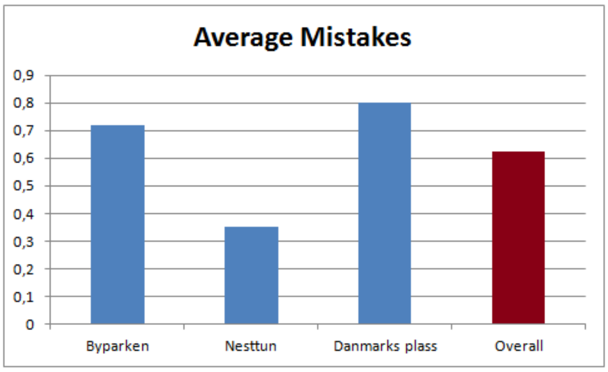


FIGURE 4.4: Average number of mistakes made overall

Once again tourists made far more mistakes than local users, with an average of 2,6 mistakes and just 0,5 for locals (see Figure B.5). Out of all the 180 people observed during this time period there were 51% that did not have any problems while using the ticketing machine. This means that almost half of all the users made one or more mistakes while interacting with the machine. If the number of mistakes made can be lowered, it will also have an effect on the time taken and raise the satisfaction of use. To know exactly how to remedy the situation, the area in which mistakes were made is important.

The types of mistakes made were fairly similar across the stations, with an exception of Danmarksplass where almost all the errors were during the payment phase (see Figures 4.5, 4.7, 4.6). The reason behind this is that the card terminal was broken, and so the machine would not register transactions by card. Some of the users decided to cross the tracks to use the machine on the other side, which was functional. One solution to this issue would be to have proper feedback if the card terminal is out of order, and ask the user to try a different machine. Some of the users also appeared to not have any coins to pay with, and had to find somewhere to exchange their bills. A few situations also arose where the terminal would not accept the coins right away, and the user had to deposit the coins several times. It is therefore clear that one of the most important issues to address is the payment phase. The solutions to these issues seem to be mostly of a mechanic nature, and as such are outside the scope of this thesis.

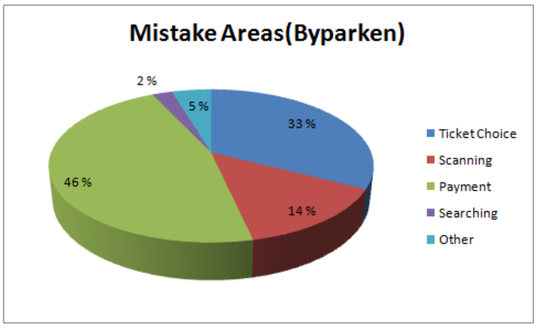


Figure 4.5: Types of mistakes made at Byparken

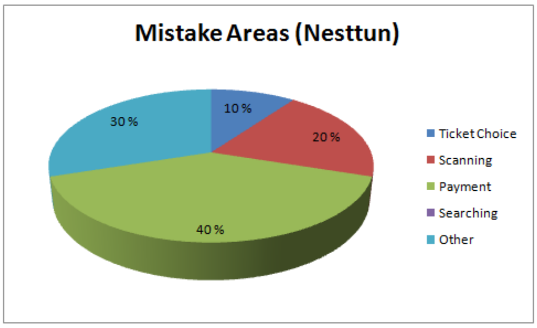


Figure 4.6: Types of mistakes made at Nesttun

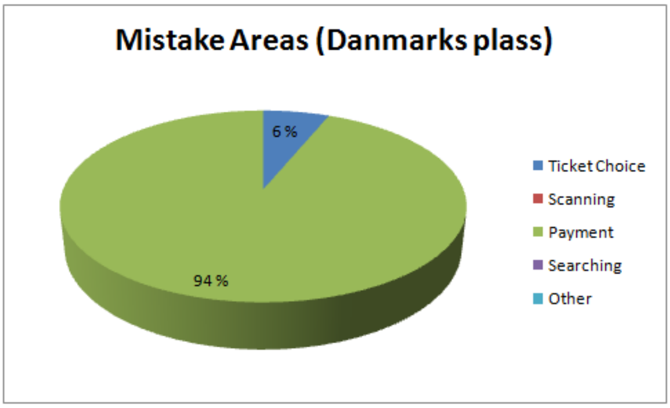


Figure 4.7: Types of mistakes made at Danmarksplass

Aside from not being able to complete the payment phase, the most common mistake areas was with ticket choice and scanning the Skyss card. Overall 23% of users were unsure of which ticket to choose. A tourist forgot to change to the English menu, and ended up choosing the student ticket, but corrected himself before paying for the wrong ticket. Several users also spent extra time purchasing two or three separate tickets, when the group ticket could be used. 12% of users made errors while scanning their card. The most common ones being to forget to scan the card twice, removing the card too quickly or the machine not reacting. 7% of the errors were with unforeseen events, such as two users who went through the whole payment process, but were then told on screen that their Skyss card was full. Only 1% of the errors made were during the search phase, where users can search for stops that are outside the main zone. As Bybanen itself only travels within the main zone, the search function is used for connecting bus travels. It can also be used to purchase tickets for bus only, as with all the other features on the ticketing machine (see Figure 4.8).

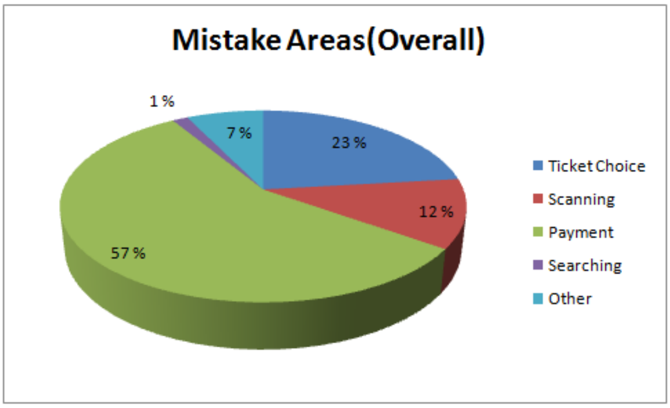


FIGURE 4.8: Types of mistakes made overall

Based on the design principles described in section 2.2.1 the principle that was most likely to prevent an error was assessed. These possible improvements also serve to see how well the principles can solve current issues, and if novel design principles could be needed. Once again the difference between the stations was not large enough to be of consequence, as the principle chosen was based on the mistakes made (see Figures 4.9, 4.10, and 4.11).

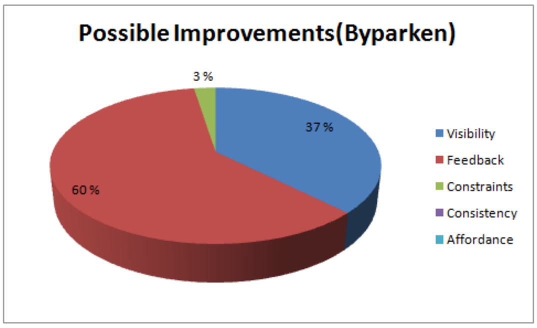


FIGURE 4.9: Potential improvements at Byparken

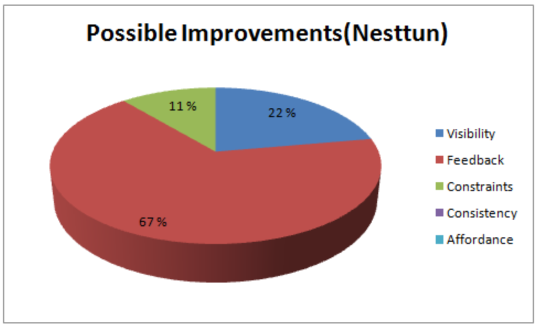


FIGURE 4.10: Potential improvements at Nesttun

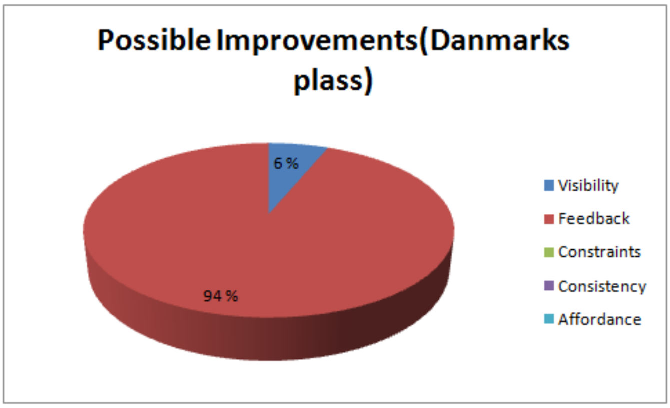


FIGURE 4.11: Potential improvements at Danmarksplass

These solutions are based on assumptions, and will require further analysis to be backed up. Overall 69% of the errors made could be prevented by the use of better feedback. Both in the case of the payment terminal not responding, and a more clear signal that the user has to scan their card twice. 28% of the errors could be prevented by better visibility. For instance the possibility of buying several tickets in one transaction should be better represented on the main page. The remaining 3% of the errors could be solved by implementing constraints, so that a user will not be able to start a transaction for refilling their card if it is full. The two design principles affordance and consistency seemed to already have been followed quite well. Buttons are shaped to afford pressing and the back/cancel buttons are always placed in the same position (see Figure 4.12).

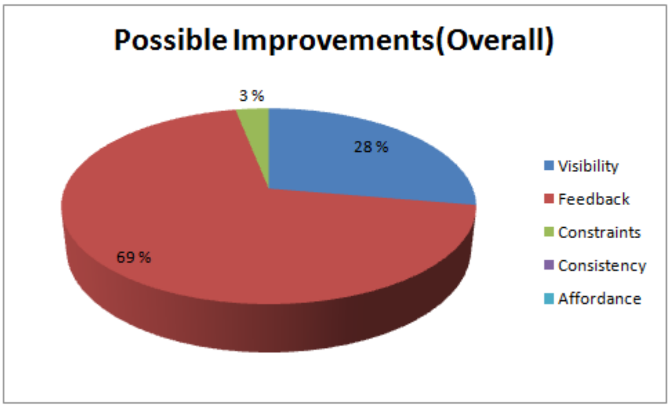


FIGURE 4.12: Potential improvements overall

The possible improvements detailed here might suggest that 100% of the potential can be solved by the use of existing principles, but this is not the case. These suggestions were based on immediate perception while observing and are only meant as possible solutions. These principles would have to be implemented and tested in order to say anything substantial.

4.1.3 Summary

Through the analysis, the questions from section 4.1.1.1 have been answered. The main features of the TVM are to purchase a single ticket for one or more people, or to refill an already existing travel card. Although many users erred at some point in their interaction, the first stage of choosing which type of ticket to purchase went fairly well. The main features were then simple to locate and use, but certain functions such as to buy several tickets at once were less obvious.

Almost half of all the users that were observed made one or more mistakes. The mistakes were made in many different areas of the system. During the payment phase, while choosing a ticket and while attempting to scan the travel card being the most prominent ones.

The efficiency of the system was found to have potential for improvement, and although existing design principles might make the system easier to use, it will not necessarily become more efficient because of it. It is here the a new design principle might benefit the most. Among the people observed the average time taken was 37 seconds, and was often cause for a traveler to be forced to wait for the next scheduled departing time. The time differences between genders was not very noticeable, with only 5 seconds separating them. As for age groups the time difference was more significant at the individual stops, but not as much while comparing the overall times. The elderly users for instance had a much higher time taken at Byparken, which could be due to the added pressure of a crowded stop, as this is one of the most busy stations.

4.2 User Testing: Existing Design

The second phase of creating the design principle was to create a working prototype of the existing Skyss ticketing system. In the following sections the method, results and analysis of the user testing and creation of the prototype will be presented.

4.2.1 Method

The initial prototype that was built was one aimed to replicate the already existing system. It was done after already having performed an observation, collected images from the pilot study and with knowledge on the general structure of the system. This, along with an efficient prototyping tool, made creating a high-fidelity simulation of the existing system quite efficient. Thus eliminating one of the disadvantages of creating a high-fidelity prototype. The prototype [1] was then uploaded to the Internet, so that users could be tested wherever they felt most comfortable. In order to not be affected by outside influences and distractions my only requirement was that the test would be performed in a room separate from other people. A drawback of such usability testing, is that the setting is not ecologically valid. Meaning that it does not take place in the same environment as with the actual device. Therefore it is even more important that other means of testing were done, as with the observation and focus groups. The test had a set of pre-conditions that users were made aware of before they were asked to perform any tasks. It would obviously not be possible to scan a travel card in the prototype, but this had been solved by an automatic transition to the next page, when scanning would otherwise be required. This was also true for the payment phase of the prototype. In order to make the test as close to the real world version as possible, the browser was minimized and placed in the middle of the screen. The users were then asked to consider this frame the only area that should be focused upon.

[1] The prototype based on the existing design is available here: http://share.axure.com/B6QCM4/

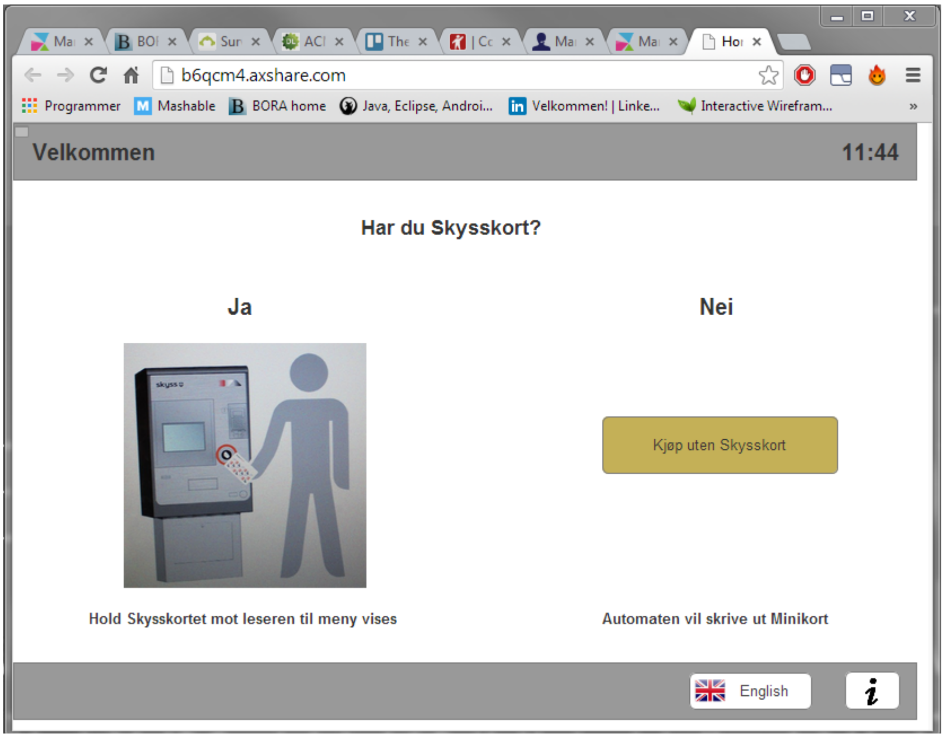


FIGURE 4.13: Resizing the prototype

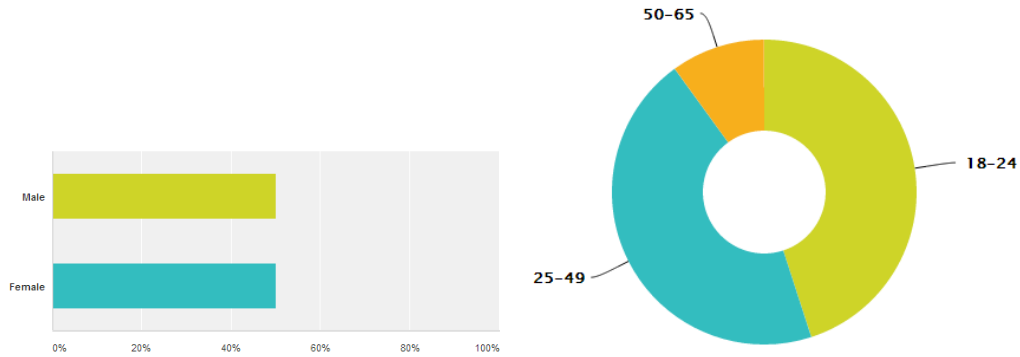
One issue that some users noticed, was the small, light gray square in the upper left corner of the prototype (see Figure 4.13). This is actually a button that opens the navigational structure of the prototype, which is something a user should not be concerned with. If this button was clicked, the test was paused for a moment and the users were explained that it was not a part of the prototype. This to avoid any further confusion throughout the later tasks.

Each task was written to test a particular part of the prototype (see Section A.2). The users were not given a time limit to complete the tasks, and after completing every task they were asked to fill out a questionnaire. In particular a user satisfaction questionnaire that was meant to elicit their opinions about the experience of using the system. With exception of questions concerning demographics the questionnaire was formatted using likert scales (see Section 3.4.1). The reason for using likert scales is that they are good for measuring opinions, attitudes and beliefs (Sharp, Rogers, and Preece, 2011). The questionnaires were also answered right after the testing phase, ensuring a fresh experience in the minds of the users.

Most of the questions were articulated in a positive way, while one question was articulated negatively. The reason behind this was to ensure that users took it seriously, and paid attention to the question at hand.

4.2.2 Results and Analysis

While the demographics for the observation were arbitrary, the test subjects could be chosen freely for the user testing. As mentioned in section 3.2.4 purposive sampling was used to gather respondents. The goal was for the samples to be similar to the ones observed using the system, for comparison. Each gender was represented equally, with ten respondents in each category. For ethical reasons there were no users under the age of 18, as this would require parental consent. In addition it proved difficult to obtain any elderly users above the age of 67. They were both contacted directly around Bybanen and through posters at retirement homes, but no one could be found that were interested in participating in the study (see Figures 4.14a, and 4.14b).



(A) Gender variation among the users (B) Age groups among the users

FIGURE 4.14: Demographics among the observed

Here follows the questionnaires analyzed in thread with the likert scales for the sake of readability.

Navigating through the system was simple and enjoyable

Almost half of the users had a neutral opinion concerning the navigation of the system. There were not many strong reactions in either direction, with four users that agreed with the statement and four who disagreed with the statement.This indicates that there is some room for improvement, but that there are no major issues with the navigation that seem to frustrate the users (see Figure 4.15).

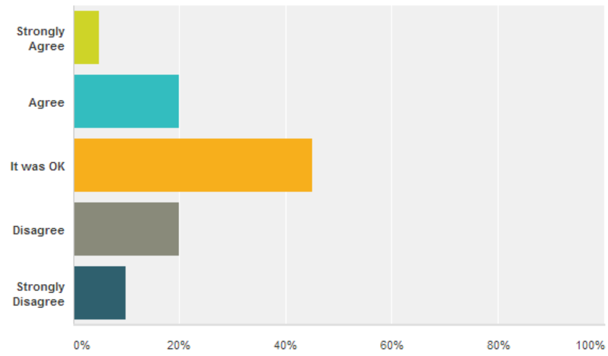


FIGURE 4.15: How users felt about the navigating through the system

I can accomplish what I want with few clicks

This statement had a similar range as with the navigation, but with slightly more users who either strongly agreed or disagreed with the statement. This again suggests that there is room for improvement, but that users feel like they can complete their tasks in relatively few steps (see Figure 4.16).

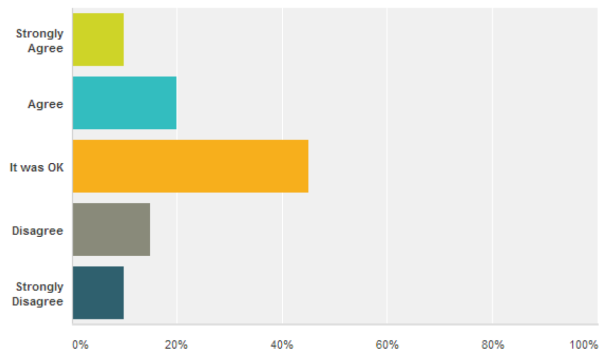


FIGURE 4.16: The users sense of efficiency with the prototype

Recovering from mistakes is quick and easy to do

The results from this statement were ambiguous, seven users agreed with the statement, and did not think it was difficult to recover if they had made a mistake. Six users were neutral and five users disagreed with the statement, feeling that it was very difficult to recover from a mistake (see Figure 4.17). This means that more than half of the participants were not entirely satisfied with this solution. If a user made an error there were two available options, one button to cancel the entire transaction and one button to return to the previous screen. It is not certain then why so many participants felt that it was difficult to recover from a mistake. With more than one way to recover from a situation, it is likely one of reasons that the rest of the participants were satisfied with this aspect of the system.

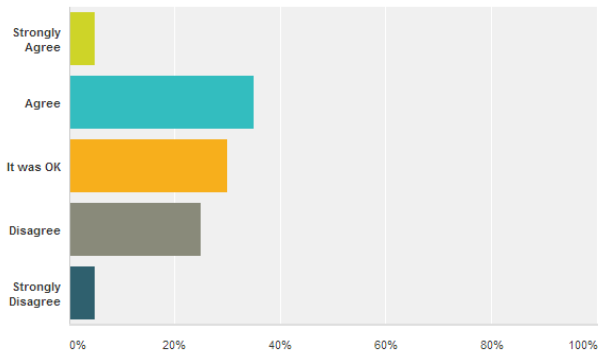


Figure 4.17: How easy users felt recovering from a mistake was

Using the system is difficult

As expected this statement had a slightly different result than the others, with more users either agreeing or strongly agreeing. The reason for this is the negative articulation, which means that almost half of the users felt that the system was difficult to handle. Seen together with the high number of errors observed during observation, it is clear that there are some elementary changes that need to be made (see Figure 4.18).

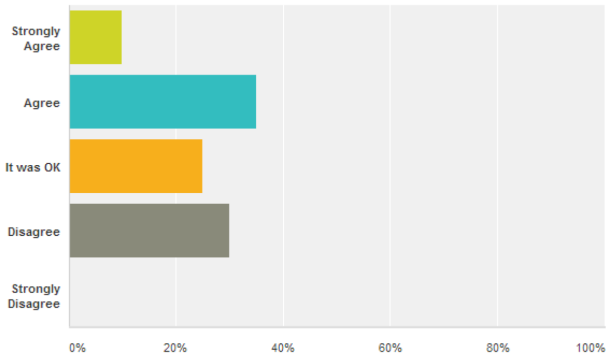


FIGURE 4.18: How difficult the users felt the system was to use

It was easy to get help from the system if I needed it

Eight users either disagreed or strongly disagreed with this statement, and eight users were neutral. Seeing such a high number of dissatisfied users suggests that this is an area where the system can use a lot of improvement. Self-service systems are as mentioned highly reliant upon the users ability to complete tasks by themselves. There is no human help available, unless of course they ask other users of the system to assist them (see Figure 4.19).

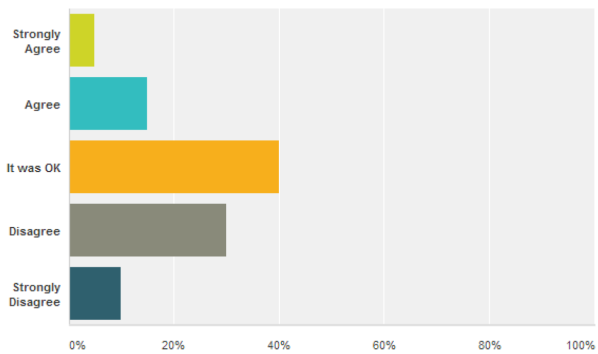


FIGURE 4.19: If users felt it was simple to get help from the system

All functions were clearly visible and easy to find

This statement is the one that most users disagreed with. 16 users either disagreed or strongly disagreed, conveying a strong need for improvement on the visibility of functions. One possible reason for this high number is the task where users were required to find a map of "Sone Bergen". Several of the users did not even manage to find the map, clearly affecting their sense of the visibility of functions (see Figure 4.20).

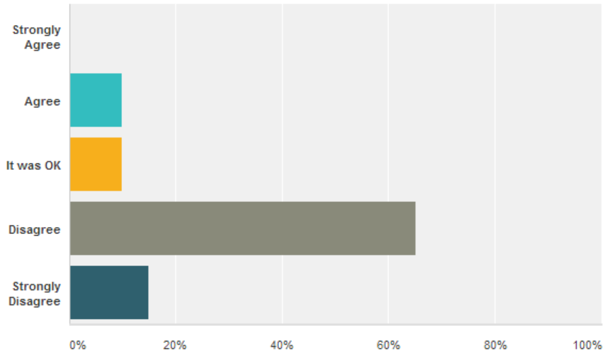


FIGURE 4.20: If users thought that functions were visible

I don't notice any inconsistencies as I use the system

A large number of users also disagreed with this statement, where 11 users either disagreed or strongly disagreed with it. While observing and using the system, no particular inconsistencies were discovered, and as such it is difficult to tell why so many users felt that the system was inconsistent. It could have something to to do with the understanding participants had with the term itself. This issue is then something that will be touched upon further in the focus groups (see Figure 4.21).

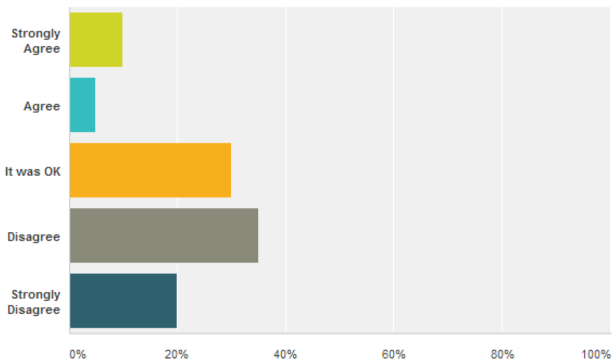


FIGURE 4.21: If users thought that functions were consistent throughout the system

Overall I am satisfied with the system

The reactions to this statement were distributed towards not agreeing with the statement, but with a large percentage of overall satisfaction being quite neutral (see Figure 4.22). User satisfaction is vital to the success of any system, and so ensuring the improvement of this statistic is one of the key goals of the new design for the prototype.

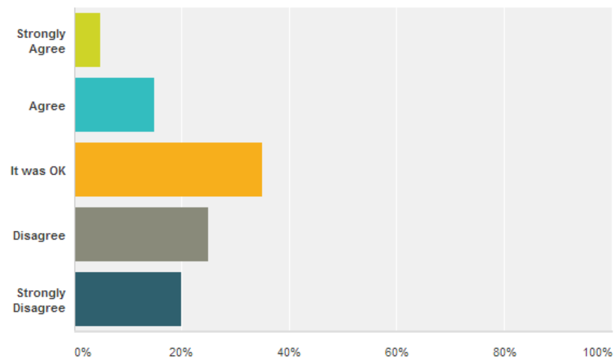


FIGURE 4.22: If users thought that the system was overall pleasurable to use

4.2.3 Summary

The analysis of the questionnaires show that the general conception of the system is not very positive. Neither extremes were represented with a majority, but almost 40% of the participants were displeased and almost 40% thought it was OK to use in terms of overall satisfaction. It is therefore clear that there is much potential for improvement to the existing design.

4.3 Focus Groups

The third and last phase of formulating the design principle was to perform several focus groups with previous participants in the user testing phase of the project. In order to determine if a focus group would suit this project the circumstances of the themes that would be discussed were considered. Would there be a potential for causing discomfort among the participants? Situations that may be the cause of such discomfort are: when intimate details of private life need to be revealed; when participants are not comfortable in each other's presence; and when participants are likely to disagree profoundly with one another (Bryman, 2012). As the subject matter is not of a particularly intimate nature it was quite certain that the questions asked would not cause any such issues. There was no guarantee that the participants would not have any major disagreements, but based on their previous participation and use of the prototype it was also deemed that a focus group would be well suited for further research. In the following sections the method, results and analysis of the focus groups will be presented.

4.3.1 Method

Out of the 20 people that completed the questionnaire, only 12 could participate further, which means that there would be 3 focus groups in total. Based on time constraints and availability this was deemed to be enough to obtain theoretical saturation. A version of the Scissor-and-Sort technique was used to analyze the data gathered from the focus groups. After an initial read-through of the transcription color-coded brackets can be used to highlight major topics and issues. The coding exercise requires several passes through the transcription to gather all the relevant data (Stewart, 2007, p. 116).

4.4 Results and Analysis

For this thesis a set of nodes were created in the transcription tool Nvivo. These nodes represented specific themes uncovered during the discussion, such as payment, expectations and zones. In turn all themes were analyzed and divided into categories of relevance. As such some themes are in multiple categories. The relevance of data from the focus groups was determined based on what would directly affect the interface, and what could be tied together with the use of design principles. The participants personal attributes were coded as follows: initials, gender and then age group.

4.4.1 Most Relevant to Research

Expectations

Self service terminals have to be clear and easy to use, and follow the user through simple steps that make sure that they can complete the action they need to. The terminal should guide you through the system, and not require the user to imagine a possible solution. The user should not need to ask questions to anyone, it should be that simple. If a situation should arise where a user needs answer to a question there should be an easily available information button present.

The process should be quick, as users are often in a rush while using the terminals. One respondent explains the process in this way : *"Jeg syns at systemet er ganske tregt, sånn at du står på en måte og venter på om, har du gjort det riktig nå eller? [...] får systemet med seg hva jeg holder på med"* (EK, Female, 25-49)? If there are technical limitations as to why the system is slow, and the developer does not have the resources to solve them, there should be a way to signal to the user that they might have to wait for a few seconds. If the user is forced to wait without knowing why there is bound to be some incoming frustration.

Information

Information buttons are a good thing, but they have to contain the appropriate information. If the user is presented with a large amount of information for each page it will be difficult to scan for the piece of information that they are looking for. Instead an idea is to have the information be specialized for the page that the user is currently looking at, and have the option to view the rest of the text if needed. The user should not have to add a lot of information themselves, the less left to the user the better.

Information icons also need to be easy to spot, and be clearly visible amongst the other relevant functions of the system. In terms of the Skyss-terminal there are two different information icons, and they point to different pages. This is not consistent, as one of the information icons also appear in different areas depending on where in the system the user is. One respondent was certain that the information icons represented the same information, and as such did not think to click both of them.

One task that almost all the respondents had trouble with was finding the map and information about the zones. Two of the respondents even had to give up looking for the information, and did not think it existed. It is actually the screen that appears when clicking the second of the information icons, and is between two large buttons while selecting which zone to travel to. In two of the focus groups the respondents agreed that a better solution would be to have a smaller icon of a map instead of the second information icon. The reason being that the only information behind the button was concerning the zones, and included an image of a map.

Feedback

Users need feedback in order to be assured that the actions they perform are the correct ones. One of the participants had paid for a ticket, and assumed that it was on their card, but when they entered the tram the scanning device showed a red light. Indicating that there was no valid ticket on the card. This happened because the participant had forgotten to scan their card a second time. Such issues could be solved with the help of more feedback, with for instance sounds or blinking lights to help the user understand where and how to perform the action.

It was also mentioned that conventions of color should not be overlooked, such as the check mark that signals a completed purchase. In the current system it is orange, but two of the respondents felt that this should be green as it is a universal sign for something that has been accomplished.

Layout

Several respondents expressed that the system was very simple to use if they just needed to buy a single ticket within "Sone Bergen". In this case they would both be able to use the quick choices and standard one, which only required a few clicks to complete. The layout itself was also deemed straightforward to use in terms of button size and the relation to other elements.

The problems first arose when they had to perform other actions than this, such as buying a "PeriodeSkyss". Two respondents mentioned that they often forgot to change the type of ticket to student, as the button to do so were not visible enough. This meant that they would enter the payment screen before they realized that the amount to pay was too high. A suggestion to solve this particular problem was to add an extra screen where the user would choose which type of "PeriodeSkyss" they wanted. Although this could ensure that a user gets the right price, it would also mean that every user has to go through one extra step in order to complete the process.

Another suggestion that got introduced by was to merge some of the steps in the purchase process. So that it would be possible to select which type of ticket, and then which zone all in the same page. The other participants agreed that this could save some time, but there could also be a risk of the layout becoming very cluttered.

One person from both first and second focus group brought up an issue that the other respondents in the group had not thought of. Why not put the quick choices to the left side of the screen instead of the right? This is the direction that most people read after all, which means that many users will not notice the quick choices since they have already found for instance the adult button that takes them further into the system. Once realizing this fact the other respondents in both groups agreed that this should definitely be a key change in a revised version of the system.

Payment

One of the things that was uncovered during the observations was that many people had difficulties when it came to paying for their tickets. In some cases because of faulty terminals, and in others because of confusion concerning ticket and payment types. In the focus groups it was also mentioned, but not as frequently as what was imagined.

One of the topics that were brought up during the second focus group was that the system timed out too fast. A respondent said:

> Når jeg skulle kjøpe billett ut til Os, når jeg var kommet helt frem til der jeg skulle betale så gikk jeg ned i lommeboken og brukte litt tid. Også plutselig så var alt vekke igjen, og så skjedde det samme om igjen at det bare forsvant og jeg fikk ikke tid til å finne frem kronene (MB, Female, 18-24).

This means that the respondent had to repeat the entire purchase process, including searching for the stop, again. Suggestions that would alleviate this issue was to add a timer on the payment page, so that the user would know how many seconds was left before the transaction was canceled. It was pointed out that this might create a more stressful situation, as you would have to figure out what the timer was for. Another respondent did not agree though, and thought it would be quite simple for a user to comfortably understand such a feature. A final idea on the matter was to also have a way to extend the timer, thus allowing the user to find the appropriate means of payment.

When a user wishes to buy a ticket with their Skyss card they have to scan the card twice. One time in order to start the transaction and once more to place the product on the card. This was confusing to many users that were observed, in all age groups. One of the respondents even claimed to have lost her ticket to another traveler. She had only scanned her card the first time, and then ran on the Bybanen as it was about to leave. Then she observed another passenger walking up to scan his card and saw the receipt pop out, confirming that he had in fact received her ticket. The other respondents reacted very negatively to this incident and did not understand why the user should have to scan their card twice. In conversation with Skyss (Nesse, 2013) two main reasons for this was presented.

1. There is a different interface presented based on which Skyss card the user has.
   (a) Registered card - The user is known, and there is no need to define which category they are in.
   (b) Unregistered card - The user is unknown, and there will be more available actions to choose their category. In addition it did not use to be possible for such users to purchase "PeriodeSkyss" or "UngdomsSkyss".
   (c) No card - The user will have limited options.
2. The transactions are stored on the Skyss card and therefore the card has to be scanned again in order to transfer the new product to the card.

In other words the two-step scanning is a conscious decision to limit the number of steps in the rest of the process, and to not present unavailable products to the different users. While this appears to be a good solution, it seemingly causes some issues and has potential for change.

4.4.2 Some Relation to Research

Expectations

A new user should be able to learn how to use the terminal quickly, and to be able to perform the most basic actions. When a user has used the terminal for a certain period of time they should be able to complete the tasks even quicker.

The terminal also needs to be consistent, which means that the design should transfer well known information about a layout from other systems that the user likely has experience with. It is also then important that all these conventions are used throughout the system, and that the design stays the same during the whole process.

The system should not force a user to go through the whole process, but have commands in place that ensure the possibility of returning to a previous action or aborting.

Information

Many respondents expressed the need for additional information in plain sight. For instance there could be an age group along with the type of ticket that they wish to purchase. That way it will be easier to avoid confusion about what age a child, a youth, an adult and an elderly ticket requires.

Language

There should be at least two languages available, the language spoken where the self service terminal is located and an international language, such as English. If a large percentage of people in the area speak a different language it would also be a good idea to include that language.

One respondent also mentions that it should be possible to change the language on all the pages, and not just the initial one. In addition the user will not be returned to the front page when clicking the language button, but will get the translation for the current page they are on.

Layout

In the upper right corner of every screen there is an icon of five colored squares that represents how far in the transaction process a user currently is. One respondent mentioned how this was very clear, and a good way to show the user how many steps are left, while another said that they did not understand what they were for. They initially thought the icons were a logo for something. Aside from this respondents felt that the general use of icons and images was very positive and made navigating the layout easier.

One respondent suggested that he would rather go back to having physical buttons, and that users are having a hard time with the touch screen devices. Physical buttons is something that the users are used to, which was his reason for believing it would make the device more manageable. While this might seem like a good idea, another user pointed out that you would then have to divert your focus away from the screen when performing an action. In particular if the system required the user to type something in. Another alternative that was then presented would be to enhance anything that is clickable. This could be done by having a specific icon that symbolized that something can be clicked.

Payment

The respondents in both the first and second focus group agreed that it should be possible to pay with bills. The issue did not come up in the third focus group. It was thought that many people, especially the elderly were very reliant on paying by cash, and so it should be possible to pay with bills and not have to walk around with tons of coins all the time. It is possible to pay with bills on the bus, but not on the Bybanen. This means that a traveler would have to risk getting a fine, even though they did have enough money to refill their card.

When asked if there could be a good reason for why it should not be possible to pay with bills the respondents were unsure. The main thought was that it would be an enlarged risk of people wanting to break into the machines, as they often are placed in very open areas. Which is also one of the reasons the person responsible for the user interface stated when faced with the same question. It was a conscious choice related to large investments, large maintenance costs and that the banknote feeders apparently did not work well in moist weather (Nesse, 2013).

Searching

In focus group two and three there were participants that had issues with the search function of the system. Both mentioned that the naming of the stops were hard to understand, as they contained codes that were unknown to them. Such as S126, and S124. The name of the stops were the same, but the codes were different. This lead to frustration, as they would then be unsure if that affected the price of the ticket, meaning they went back to click the other stops in order to find out.

Products

Throughout the focus groups there were many comments on the different products offered by the ticketing system. Some of these issues were not necessarily related to the interface, but are still interesting observations made by the participants.

The tickets are not called tickets in the system, they are named after the bus company, such as "EnkeltSkyss" and "PeriodeSkyss". One respondent found this to be hard to understand, and thought the tickets should be renamed. This issue can be related to the heuristic of speaking the users language, as to avoid confusion. As it was only mentioned by one participant, it is likely not a common issue, but should still be considered.

FamileSkyss: In all three focus groups the participants discussed the family tickets. In the second and third focus groups the participants were unsure of the difference between a group and a family ticket. Is it allowed to buy a family ticket if there is no relation between the travelers? Which tickets do you need to buy to get the discount? None of these issues were addressed within the information page.

In the first group a participant suggested that the system should calculate the discount based on which tickets that are chosen. While this method would eliminate a step in the process it could also have the potential of confusing a user. As one participant points out:

> Det er sikkert litt.. altså fordi at det liksom er noen folk som ikke ville skjønt det da. De liksom sånn der "eh.. jeg skal ikke ha en voksen og et barn, jeg skal ha en familie". For det er jo det man er vant til å kjøpe på bussen, familiebillett, men det er jo essensielt (EM, Male, 24-49).

So even though automatically registering a discount would solve one issue, it would also create a new one.

GruppeSkyss: In relation to the group ticket, several users in the second and third focus group pointed out that the icon can be misleading. It depicts three people, leading them to believe that in order to be a group there has to be at least three travelers. Instead of then using the group ticket they would go through the system twice and purchase one adult ticket for each time. This was also the case for a participant traveling with a dog. Instead of buying a group ticket she purchased one for herself and then one for the dog.

PeriodeSkyss: When buying a "PeriodeSkyss" there is no separate screen to choose which type of traveler you are, such as student or elderly. Two of the participants in the first focus group said that they always forgot to change the choice to student, meaning that they had to come back from the payment page to change. One possible reason for this was that they only had to refill their card once a month, and then forgot where to click in order to get the student discount.

A positive feedback concerning this ticket type is that there are multiple ways to renew it. You can either click the "My Card" option and renew it there or go through the "PeriodeSkyss" menu that appears after your card is scanned.

Zones

An issue that was raised by the participants in the first and second focus group was the confusion about which zone that should be used. The main agreement was that most people do not know exactly where the different zones end. Some of the reasons for this could be that the person has just moved to the city, or as one participant puts it:

> Det var mindre før, så det var liksom sånn at hvis du skulle til Bønes så måtte du en sone ekstra eller ett eller annet sånt som det der. Men nå til dags er den såpass stor at folk tenker ikke over det, men allikevel det gjør det veldig forvirrende det der med sone bergen og annen strekning og sånt som det der (EM, Male, 25-49).

This means that even if a person knew where the zones were, they could be likely to change frequently along with the new bus companies or new bus routes. As the only zone that the Bybanen travels through is the Bergen zone one participant suggested that they should have two different systems. One for the Bybanen and one for buses. This suggestion was quickly countered by another participant, stating that a traveler often wants to travel to another destination after using the Bybanen. So two different systems would mean yet another transaction that has to be made when arriving at the first destination. Another solution that was met with more enthusiasm was to be explicit on where the zones are. For instance by using actual route maps that depict which stops the buses travel to, and in which zone the stops are.

4.5 Novel Design Principle: Integration

By combining the analyzed data from the observations, user testing and focus groups it is clear that there is room for improvement. The data shows no significant differences between types of users when it comes to performance. This can indicate that changes to the system could affect users in a positive manner. As discussed in section 4.1.2 the principles of affordance and consistency are present in the Skyss TVM. Although some design principles seem to have been followed, the users are still making mistakes while operating the TVM. A common source of errors made was due to the lack of visibility of functions, or because there was not enough feedback. Therefore these design principles are vital to consider when the alternative prototype is created. In addition heuristics and the seven stages are valuable when trying to improve on the prototype. Even though many of the issues with the TVM can be solved with existing principles and guidelines, this does not mean that some specialized principles will not also be beneficial to them.

The existing principles are quite powerful, and have the potential to remedy many issues present in the TVM. In light of this, only one additional design principle will be proposed. The need for this principle is found while considering how the existing principles are used. They represent specific qualities that a design should strive to achieve. Qualities of visible buttons that afford interaction, and qualities of placing these functions consistently across the design. The new principle will aim to support such qualities, instead of trying to replace existing ideas. Further research on other types of self-service systems will be required to further extend the design principles.

The novel design principle of integration concerns the coupling of functions that go together. In order to ensure the efficient expedition of each user, the most used functions should be gathered within fewer steps. For this to be possible it has to be done without adding too much clutter to the interface. The user has to be able to understand the connection between the functions swiftly and easily. The design principle then has a main goal of increasing the efficiency of the possible tasks available in a system. It is also most useful if the self-service device has a larger number of functions, or many ways of stepping through the system.

The structure principle: Design should organize the user interface purposefully, in meaningful and useful ways based on clear, consistent models that are apparent and recognizable to users, putting related things together and separating unrelated things, differentiating dissimilar things and making similar things resemble one another. The structure principle is concerned with overall user interface architecture.

In order to implement this principle a thorough understanding of the self-service device is needed. If functions that are not necessarily related get grouped together, this would have the potential to confuse the users more than it would help them. This is also the case if the device has very few functions to begin with. Imagine a self-service checkout at an airport, the user has three options: either insert a credit or business card, enter a reference code or scan their travel document. The user is asked to confirm which passengers are traveling on one screen, which seat they want on another, and at last how much baggage they will bring. In this situation there is a specific set of actions each user is presented with, and the actions do not diverge significantly between each interaction. In this situation, where the system is already minimalistic the principle of integration would not be recommended.

The name *integration* was used for the principle as the meaning of the word closely resembles what the design principle aims to do. Integration is an act or instance of incorporating or combining something into a whole. While this could indicate that one would end up with a single feature that does two things at once, this is not necessarily the appropriate way to interpret the term. Although there could be reasons to experiment with buttons that perform two things at once, as this likely would speed up many existing processes. In the next chapter an attempt to implement this principle in a new design will be presented, as well as the following evaluation and testing of the new design.

Chapter 5

Using the Design Principle of Integration

This chapter presents the new design of the prototype created, it was based on the research done and on the novel design principle now known as integration. The main objective of the prototype was to improve the efficiency and user experience of the existing system. The specific changes made to the original are also detailed in this chapter.

5.1 Changes to the Existing Design

The new design [1] was developed with the principle of integration in mind and also on all the data gathered. Based on the previous findings a specific set of changes were made to the existing design, some based on the novel design principle of integration and some on previously existing principles. In the discussions the term existing design will be used interchangeably with Prototype A and the term new design will be used interchangeably with Prototype B.

[1] The new design of the system is available here: http://share.axure.com/UOZYWI/

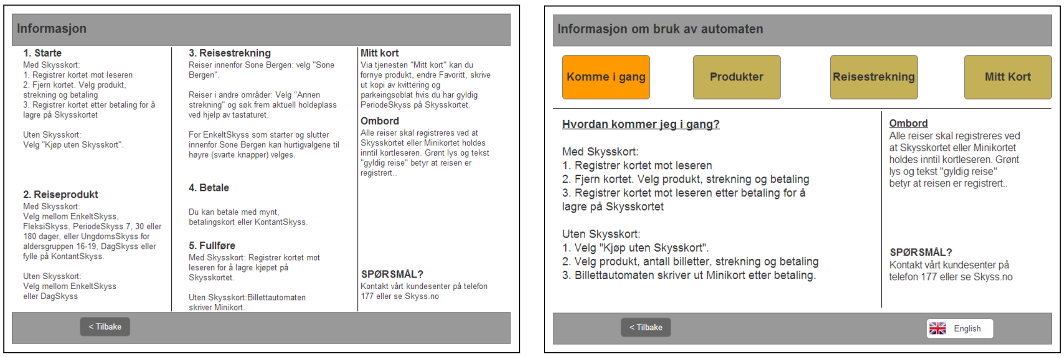
5.1.1 Changes By Use of Integration

The novel design principle of integration, detailed in section 4.5 was used while considering the ticket types and how to buy multiple tickets. Instead of being lead to a separate screen in order to get multiple tickets, it is now possible to select the number of tickets to buy, directly on the selection page. This also means that the group and family ticket types were removed, as the discount for ticket types would be calculated automatically (see Figure 5.1). When a user selects a ticket, the type of ticket appears on the right-hand side of the interface. The user will then click the ticket type again in order to purchase several of the same type. If the user removes all tickets, the button that allows them to proceed will be hidden. This enforces the design principle of constraint, ensuring that at least one ticket has been chosen before the user can proceed to checkout. The resulting design has one less screen for the user to consider, and so should make the process more efficient, especially in the case of purchasing more than one ticket.

(A) Existing Design (B) New Design

Figure 5.1: Selecting Several Tickets

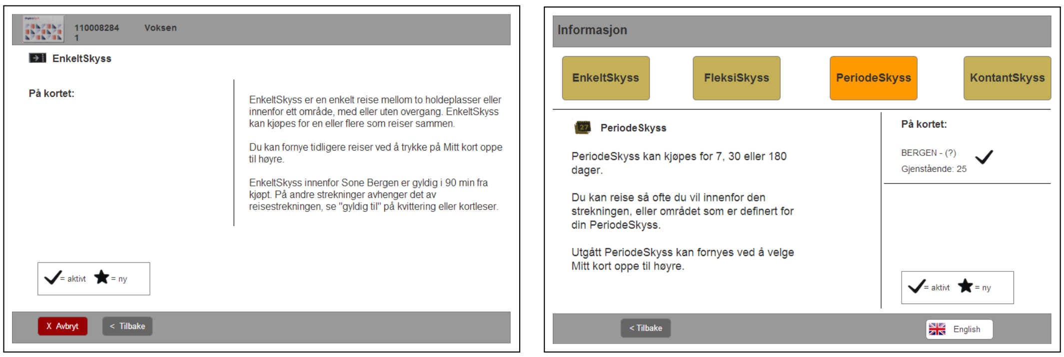
The next change was to the information buttons, and their underlying structure. How information is presented is important in deciding how easy or difficult it is to discover a specific piece of information. Subjects that are grouped into vertical categories, with meaningful accommodation, can make information easier to retrieve than if the information is not labeled and bunched up together. Information concerning the general use of the terminal was reorganized and placed under a fitting category. To navigate between these categories, several buttons that represent the category was used. The navigation itself resembles that of a standard Web page, where the current category is highlighted (see Figure 5.2).



(A) Existing Design (B) New Design

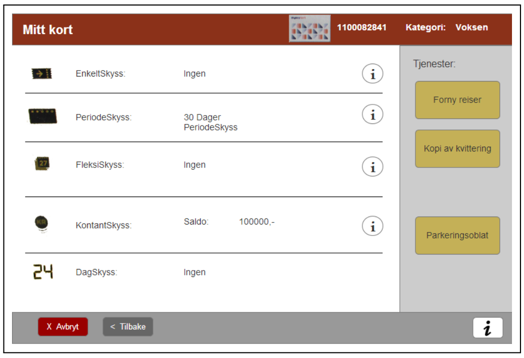
FIGURE 5.2: Main Information Screen

As with the main information page, the categories were structured as that of a webpage navigation bar (see Figure 5.3). It is then possible to navigate between the different types of purchases available, without needing to go back to the previous screen and select a different product. In addition the user would automatically be directed to the product in question if they had started to purchase, for instance, a "PeriodeSkyss". The information buttons contained within the "Mitt kort" option were also removed in the same fashion, leaving only one information button (see Figure 5.4).

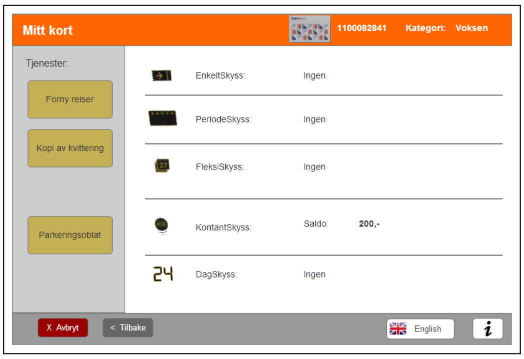


(A) Existing Design (B) New Design

FIGURE 5.3: Scanned Card Information Screen



(A) Existing Design



(B) New Design

FIGURE 5.4: My Card Information Screen

In this way several areas of information have been restructured, and instead of having four different information buttons it can all be found within one page. This is in thread with the idea of integration and the coupling of functions or information that go together.

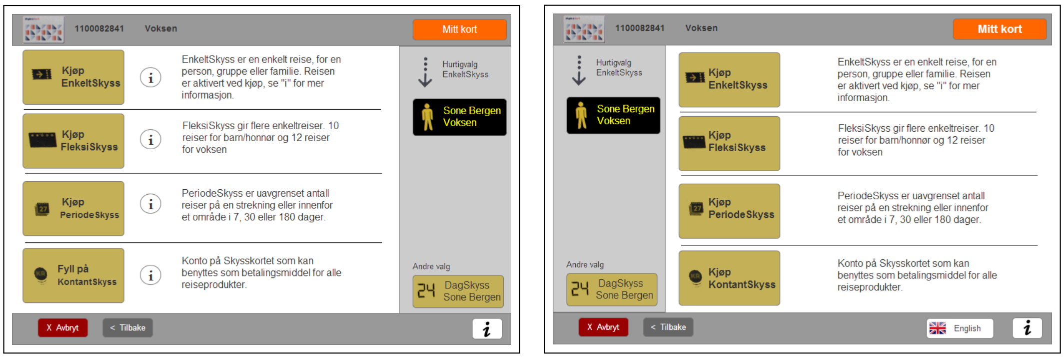
5.1.2 Other Changes

Firstly the quick choice menu of each page was moved from the right side of the interface and over to the left side (see Figure 5.5). The thought behind this is to ensure that users who are traveling alone, within Sone Bergen, will see these options before they start to interact with the main menu. The idea behind it is based on the fact that people in the western world read from the left, to the right. In an eye-tracking study by Nielsen (2006) it was seen that users often read web-pages at amazing speeds, and in an F-shaped pattern. The dominant reading pattern had three main components. Users would initially read in a horizontal movement across the upper part of the content area, then move down the page a bit and read across in a second horizontal movement. Finally users scan the contents left side, in a vertical movement, creating what resembles an F-shape in an eye-tracking heat-map. Even though the application is not a web-page, an assumption is made that the scanning pattern will begin from the left.

(A) Existing Design (B) New Design

FIGURE 5.5: From Right to Left

After scanning their card, the user is presented with a set of possible products to purchase. This page contained five information buttons, one for each type of purchase, and also the standard one that lead to the general information. To minimize the cognitive load, and gather more functionality in one place, the different types of information buttons were removed (see Figure 5.6) and the underlying information was moved to the single information button in the lower right corner of the screen.



(A) Existing Design (B) New Design

FIGURE 5.6: Information Buttons

During the initial testing phase, many users had difficulties finding the map that displayed "Sone Bergen". The small information buttons that were displayed in the zone selection screens were replaced with larger buttons that clearly states their purpose (see Figure 5.7). The same button was also placed in a logical location within the main information screen (see Figure 5.8).

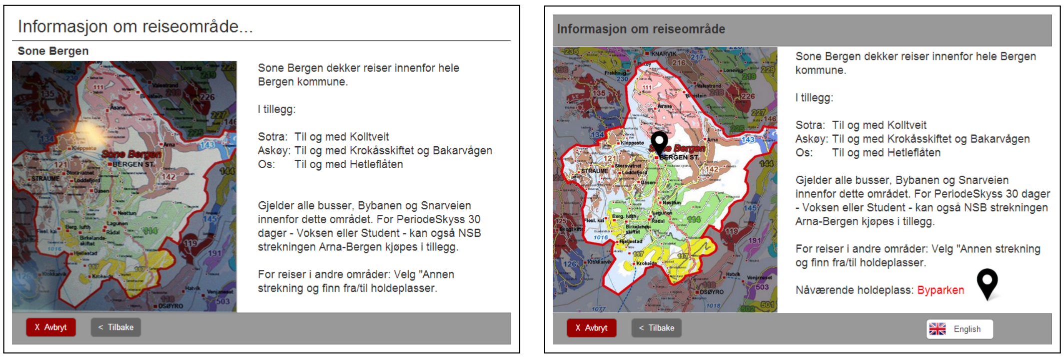
(A) Existing Design (B) New Design

(C) Existing Design (D) New Design

FIGURE 5.7: Zone Screens

FIGURE 5.8: Travel Information

A symbol indicating which terminal the user was currently located was added to the map (see Figure 5.9), making it simpler to know whether or not a different type of ticket would be needed. This symbol was also used on all the buttons that lead to the map page, consistently making the function visible to the users.



(A) Existing Design (B) New Design

Figure 5.9: Map Screens

5.2 Final Testing Session

In order to assess whether these changes to the prototype had any effect on user satisfaction and efficiency, two tests were performed simultaneously. Firstly a within subject experiment was implemented, the process of which is detailed in section 3.2.5. Secondly a user satisfaction questionnaire was answered after completing each phase of the experiment. The experiment adhered to the same ethical considerations as discussed in section 3.2.1.1.

5.2.1 Test Setup

To avoid confounding variables, and to minimize learning effects from previous participation, it was decided that none of the participants from the initial phase would be brought in. In total 20 new participants were recruited by contacting the ones that had already been part of the study. 10 were male and 10 were female. The age range went from 18-66, with the majority being in the 18-24 category.

During the test session it was important to clarify what was expected of the test subjects. It was explained that that the aim of the experiment was to evaluate user satisfaction and efficiency of two different versions of the Skyss self service ticketing system. The participants were not told which prototype was which, or in what order they would be testing them. They were informed that the task completion times would be recorded, and encouraged to complete them as fast as possible. The majority of the experiments took place in the same environment, in an empty seminar room, without any possibility for distraction. Two of the participants were not able to come to this location, and therefore were tested in their own home. The experiment was still done in a separate room, and without distraction.

As with the initial user testing, a set of tasks were made to test particular parts of the prototype. A decision was made to alter some tasks, and add a few new ones to allow for the changes to be fully explored (see Section A.3). One issue with such an approach is that it will not be possible to compare the results of the initial test with this one. While this is true, the purpose of the final testing session was not the same, and so the results from both testing phases can be used separately. The initial one to assist the creation of a design principle, and one to measure the resulting changes to a prototype.

The participants were not filmed during the test, as time constraints would not have allowed for the analysis of 20 such films to be transcribed. Each test lasted for approximately thirty minutes.

5.2.2 Test Material

Both the existing design and the new design of the prototype were used during the test. These prototypes are interactive prototypes, allowing participants to navigate the system as if it had actual functionality. Henceforth the prototype based on the existing design will be known as prototype A and the new design will be known as prototype B. The same equipment was used for all experiments, a laptop connected to a mouse and a smart-phone to measure the times taken for each task.

5.2.3 Results and Analysis

To analyze the results for each prototype, the mean value of each task was gathered and compared based on which prototype was used. The data points were normalized prior to the comparison. In the case of this study a data point was removed if it was more than twice as large as that of the mean value for the task. The sequence effect that might have been involved will be discussed in the next section.

Task 1 - Single ticket adult

In this task the participants were asked to purchase a single ticket for an adult, within the zone of Bergen and to assume that they did not have a bus card.

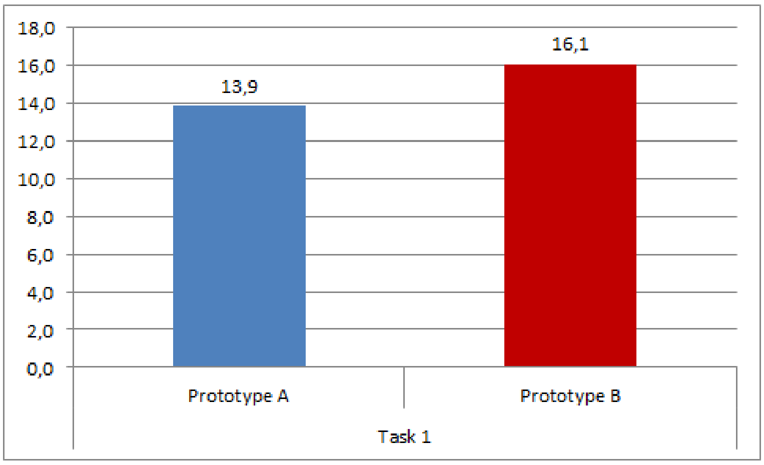


FIGURE 5.10: Times Taken for Task 1

The results show that participants in general used 2.2 seconds longer to complete the task in Prototype B (see Figure 5.10). Such a small difference does not seem to indicate a substantial change in efficiency for the new design. A possible factor that caused an increase in time taken for the new design was that some users waited a few seconds before clicking the proceed button after selecting a ticket. In a future version of the design the placement and size of the proceed button would be experimented with, in order to see if this was indeed a factor.

Task 2 - Single ticket elderly

In this task the participants were asked to purchase a single ticket for an elderly, outside the zone of Bergen and again to assume that they did not have a bus card.

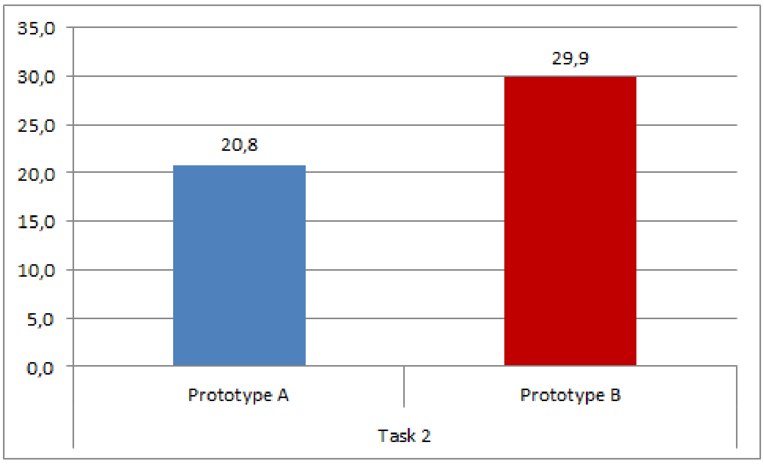


FIGURE 5.11: Times Taken for Task 2

For this task the users had to search for a specific station to travel to, which was essentially the same process in both designs. The main difference was in how to select the ticket in the first stage of the process. With almost a 10 second difference there seems to be a decrease in efficiency for the new design (see Figure 5.11). As the task essentially is the same as the previous in terms of changes made to the design, it is difficult to speculate in which factor caused such a change.

Task 3 - Single ticket child

In this task the participants was asked to purchase a single ticket for a child by using the quick choice function, within the zone of Bergen and to assume that they did not have a bus card.

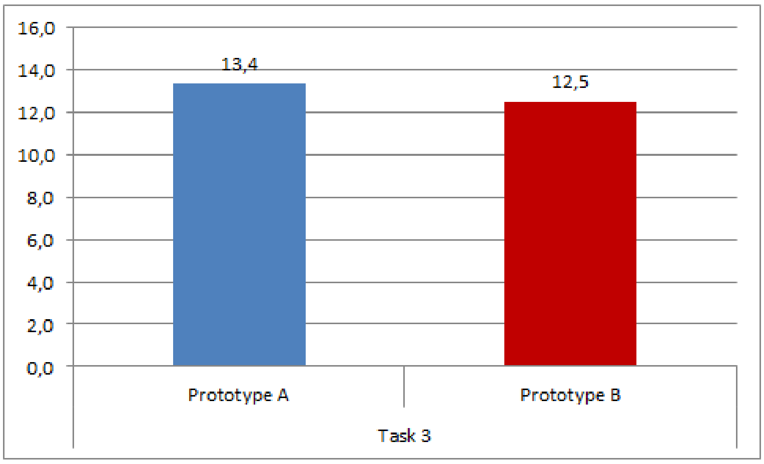


FIGURE 5.12: Times Taken for Task 3

The difference between this task for the two designs was to the placement of the quick choice function. In the existing design it is on the far right of the screen, while the new design has moved the quick choice functions to the left side. The difference in efficiency was a mere 1.1 seconds in favor of the new design (see Figure 5.12).

Task 4 - Two tickets

In this task the participants was asked to purchase two single ticket for a adults at the same time. This time also within the zone of Bergen and without a bus card.

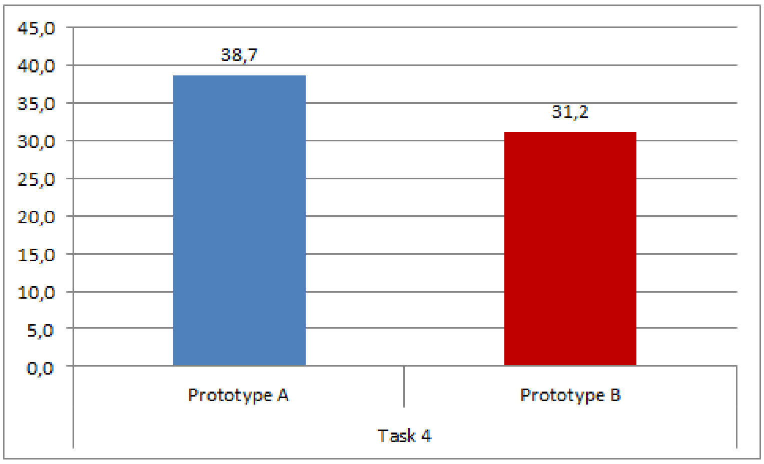


FIGURE 5.13: Times Taken for Task 4

The results show that participants in general used 7.5 seconds longer to complete the task in Prototype A (see Figure 5.13). One of the major changes to the design was in how to select multiple tickets. In the existing design the participant had to choose the correct ticket type, and then choose how many travelers would need tickets. In the existing version this could all be done by clicking, in this case, the adult ticket type two times. This is likely one of the factors that caused a positive change in efficiency toward the new design.

Task 5 - General Information

In this task the participants was asked to locate the general information about how to operate the ticketing machine.

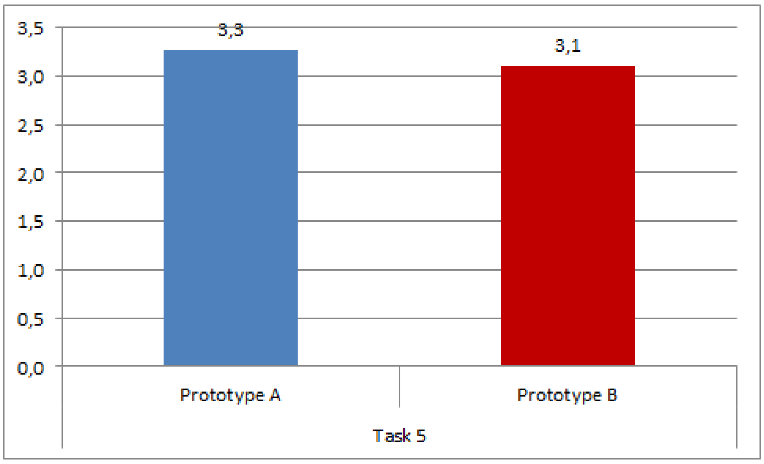


FIGURE 5.14: Times Taken for Task 5

The underlying information was changed between prototypes, but the location of the main information button stayed the same. Therefore a difference of only 0.2 seconds does not come as a surprise (see Figure 5.14).

Task 6 - Map

For this task the participants were asked to locate the map of where the zone of Bergen begins and ends.

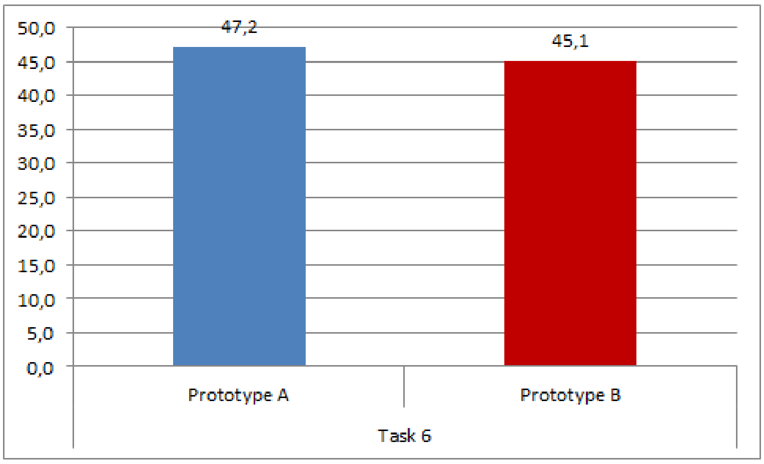


FIGURE 5.15: Times Taken for Task 6

With a difference of only 2.2 seconds there was not a substantial increase in efficiency for this task. The changes made to the design was increasing the size of the buttons used to get to the map, and also the name of the button. It was also added to the main information page. The main reason for the low increase in efficiency might be the title of the button being "where am I?". Some users expressed their confusion as they were looking for a button with a picture of a map, or the title "map" (see Figure 5.15). In conversation with Skyss the function of pointing out which station the user is currently at was not technically possible either (Nesse, 2013). In a future prototype a change could then be made to more clearly display the button as one that leads to a map of all the zones.

Task 7 - Specific Information

In task 7 the participants were asked to locate specific information about how to use the "FleksiSkyss" ticket type.

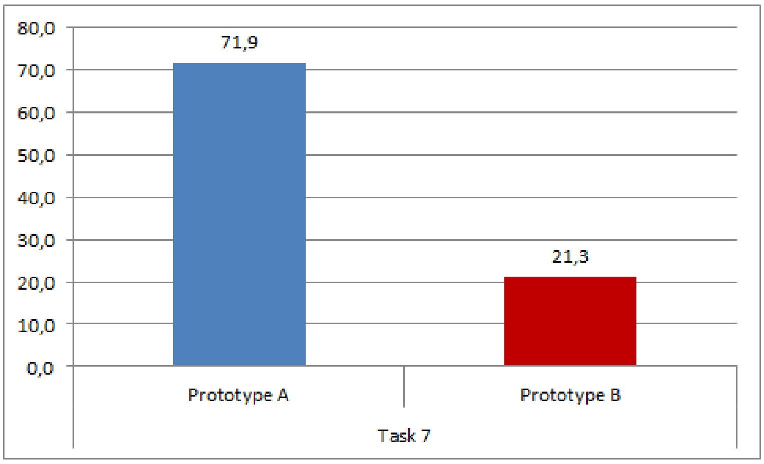


FIGURE 5.16: Times Taken for Task 7

The results show a major difference between the existing and new design. In general the participants used 61.3 seconds longer to complete the task in the existing design (see Figure 5.16). One likely reason for the large gap between designs, is the fact that there is no specific information about the ticket type under the main information button in the existing design. Thus some participants spent a lot of time reading through the information, but not discovering the actual location which was under the information button directly next to the ticket type after having scanned a bus card. It is worth noting that with a mean of 101 seconds, many data points that could have been excluded were kept as they did not breach with the rule set prior to the analysis.

Task 8 - Periode Skyss

For this task the participants were asked to buy a "PeriodeSkyss" ticket type for a duration of 30 days and to assume that they had a bus card.

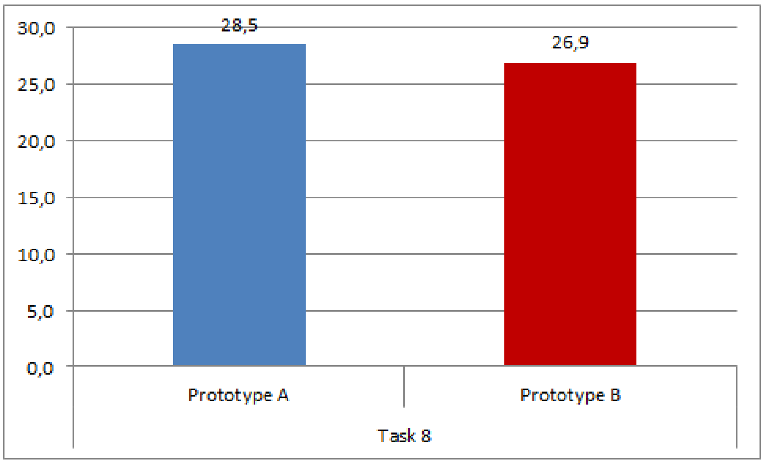


FIGURE 5.17: Times Taken for Task 8

With a difference of 1.6 seconds there was not a major difference in efficiency between the designs for this task (see Figure 5.17). The changes made from the existing design was a large font, and more focus on the possible choices for amount of days that the customer would want to buy a ticket for. There was a few seconds improvement to the new design and so this change could have had some influence on efficiency.

Task 9 - Money Remaining

For this task the participants were asked to figure out how much money that was left on the "KontantSkyss" ticket type on their fictitious bus card.

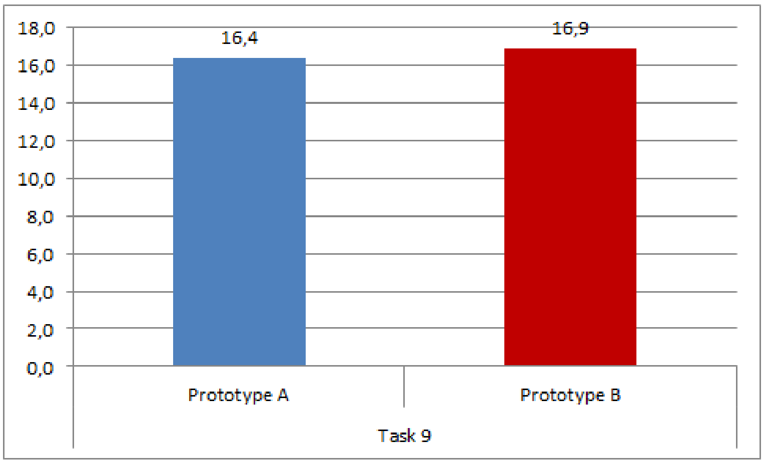


FIGURE 5.18: Times Taken for Task 9

The results show a very small difference of 0.5 seconds, where the use of the existing design was slightly faster (see Figure 5.18). The major change from the existing design was made to the content within the information screens, and to the placement of the information buttons. With such a minuscule difference in efficiency it is not possible to state whether or not these changes have improved the design.

Task 10 - Days Remaining

For this task the participants were asked to figure out how many days were left on the "PeriodeSkyss" ticket type on their fictitious bus card.

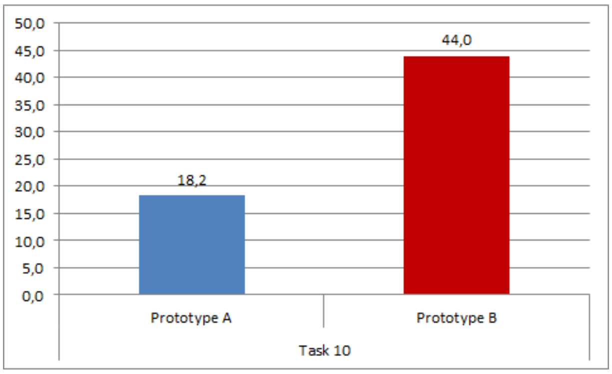


FIGURE 5.19: Times Taken for Task 10

The information about how many days were left could be found in the same location as with the money left on the card from the previous task. It is therefore odd to see a 25.8 second decrease in efficiency for the new design (see Figure 5.19).

Complete Overview

A complete overview of all the tasks show that for most tasks the level of efficiency does not differ a large amount between the two prototypes (see Figure 5.20). The existing prototype performs slightly better in three tasks, and significantly better in one. The new design performs slightly better in five tasks, and significantly better in one. Some likely reasons for the existing design still performing better in some cases will be discussed further in

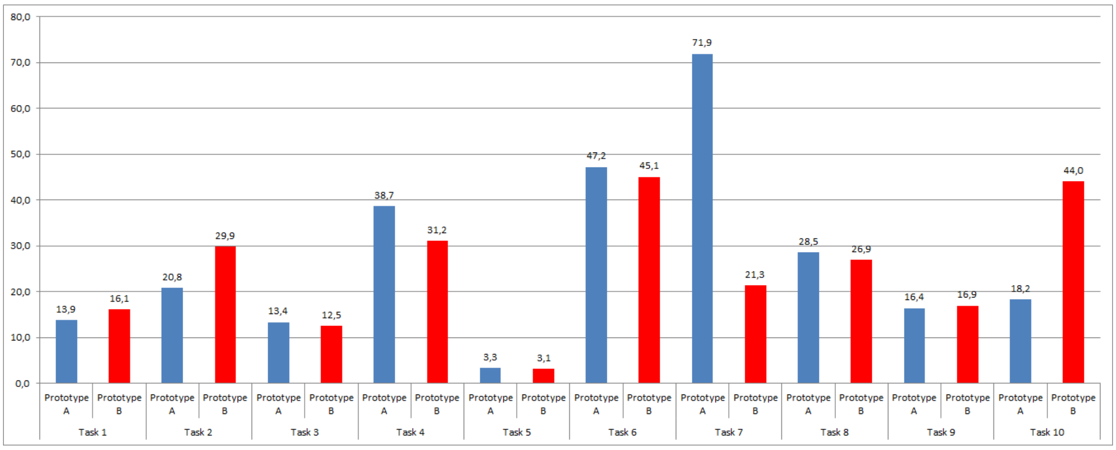


FIGURE 5.20: Times Taken an Overview

5.2.3.1 Exploring the Sequence Effect

In the within design participant experiment, a key factor was the order in which participants tested the existing and the new version of the system. This way of avoiding confounding variables may also introduce a sequence effect. This implies that participants will perform better in the second version that is tested, regardless of whether it is the existing or new version. The reason being that the participants will have learned certain aspects of the task based on the previous test, and thus will perform better. To explore this effect and the statistical significance it may have had, a Wilcoxon signed ranked test was conducted for each task.

Statistical significance is the measure of how confident we can be that the findings from our study can be generalized to the population from which the sample was selected. The level of statistical significance concerns the risk of inferring that a relationship between two variables exist, when there is in fact none. The maximum level of significance is generally noted as $p < 0.05$ where p stands for probability. Therefore a relationship is statistically significant if there is less than 5 chances in 100 that might be falsely concluding that a relationship between variables exist (Bryman, 2012).

The Wilcoxon signed rank test is used to compare two sets of scores that come from the same participants. This can occur when investigating a change in scores from one point in time to another, or when individuals are subjected to more than one condition (Laerd Statistics, 2013). In the case of this experiment the conditions are two separate prototypes. The test was run based on several variations of the data. It was run for each task for both settings, meaning participants that tested the existing design first (A-B), and then participants who tested the new design first (B-A). It was also run for the sum of all tasks combined in both settings, and for all the task individually in one data set. Many of the test results showed no statistical significance, and will therefore not be of relevance to this thesis.

Statistically Significant Results

The results concerning the statistical significance of the findings are shown first, and then the descriptive statistics. These statistics show the mean value of time taken and the standard deviation of the data set. The standard deviation is a measure of dispersion and shows the average amount of variation around the mean. The higher the deviation, the larger the spread between data points (Bryman, 2012). Two hypotheses were set as a base for the test.

H_0: There is no difference between the results for the two designs

H_1: There is a difference between the results for the two designs

For there to be a statistically significant difference between the results the Z-score has to exceed or be below a critical value of plus/minus 1.96. This critical value is found in a Z-table and is accurate for samples of 20 participants or more (StatisticsLectures, 2012). The p-value is influenced by the Z-score and as mentioned has to be smaller than 0.05 for there to be a statistically significant difference.

Wilcoxon Results for Task 6

	A→B	B→A
Z	-2,701	-2,701
p-value	0,007	0,007

Table 5.1: Wilcoxon test: Task 6

	N	Mean	Std. Dev	Minimum	Maximum
A First	10	71,120	29,1109	27,8	119,0
B Second	10	27,560	13,9752	11,0	48,8
B First	10	107,84	69,3687	15,0	210,0
A Second	10	30,510	23,1213	12,4	92,6

Table 5.2: Descriptive statistics: Task 6

This is the task where users were asked to locate a map of where the zone of Bergen begins and ends. It is the only separate task that showed statistically significant results. It is therefore possible to assume that these results would also be found in a larger group of participants as well.

The results from the Wilcoxon test for the sixth task can be seen in table 5.1, and 5.2 and show that the Z-score is -2,701 for both settings. Thus rejecting H_0, meaning that there is a difference between the results for the two designs.

In addition the p-value is below 0.05 meaning that the results are statistically significant.

Based on the descriptive statistics the participants performed 61.24% better in Prototype B when it was tested second. Participants performed 71.70% better in Prototype A when it was tested second. As such there seems to have been a sequence effect for the sixth task, as participants consistently performed better in the prototype that was tested second.

Wilcoxon Results for All Tasks Combined

	A→B	B→A
Z	-2,701	-2,4901
p-value	0,007	0,013

Table 5.3: Wilcoxon test: All Tasks Combined

	N	Mean	Std. Dev	Minimum	Maximum
A First	10	343,79	85,0186	216,4	462,2
B Second	10	224,34	68,1447	144,7	368,2
B First	10	384,32	73,7533	308,0	536,2
A Second	10	288,77	117,4255	146,1	494,7

Table 5.4: Descriptive statistics: All Tasks Combined

In this case all tasks combined means the sum of times taken for all ten tasks, per participant. So one data point in the spreadsheet would be 300 if the participant had spent 10 seconds for each task.

The results from the Wilxocon test for the sixth task can be seen in table 5.3, and 5.4 and show that the Z-score is -2,701 for the case where Prototype B is tested second and -2,49 where Prototype A is tested second. Both are outside the range for the critical Z-score thus rejecting H_0, meaning that there is a difference between the results for the two designs. In addition the p-value is below 0.05 meaning that the results are statistically significant.

Based on the descriptive statistics the participants performed 34.8% better in Prototype B when it was tested second. Participants performed 24.8% better in Prototype A when it was tested second. As such there seems to have been a sequence effect for all tasks combined, as participants consistently performed better in the prototype that was tested second.

Wilcoxon Results for All Tasks Individually

	A→B	B→A
Z	-3,327	-4,489
p-value	0,001	0,000007

TABLE 5.5: Wilcoxon test: All Tasks Individually

	N	Mean	Std. Dev	Minimum	Maximum
A First	100	34,379	39,6842	1,3	216,0
B Second	100	22,434	19,4160	1,0	144,0
B First	100	38,206	40,6633	1,9	210,0
A Second	100	28,877	39,6281	2,0	254,0

TABLE 5.6: Descriptive statistics: All Tasks Individually

In this case all tasks individually means that all data from each task was placed in the same spreadsheet and analyzed as a whole. An assumption was made based on the results from the all tasks combined test. This assumption was that the results would be somewhat similar in terms of statistical significance and performance based on times taken.

The results from the Wilxocon test for all tasks individually can be seen in table 5.5, and 5.6 and show that the Z-score is -3,327 for the case where Prototype

B is tested second and -4,489 where Prototype A is tested second. Both are outside the range for the critical Z-score thus rejecting H_0, meaning that there is a difference between the results for the two designs. In addition the p-value is below 0.05 meaning that the results are statistically significant.

Based on the descriptive statistics the participants performed 35.2% better in Prototype B when it was tested second. Participants performed 26.31% better in Prototype A when it was tested second. As assumed the results for this test were similar to the ones for all the sum of all tasks combined. The main difference is that the level of significance was even higher in this test. It is then possible to be fairly certain that this result is something that could be found in a larger sample size as well.

5.2.3.2 Summary

In this section the goal has been to derive statistically significant data based on the amount of time each user needed in order to complete specific tasks. The results show that in most tasks the participants perform better with the second prototype that is tested, regardless of whether it is prototype A or prototype B. It is therefore clear that the participants have experienced a learning process, which affected the results of the second test in a positive way.

Some tasks were not represented in the previous section, the reason being that they were not statistically significant and therefore not relevant for an actual conclusion. In task 1 through 3 there was improvement in efficiency for both settings, but the statistical significance was only found in the case where prototype B was tested first. In other words it is not possible to state that prototype B had any profound effect on the users in regards to these tasks. In task 4 and 5 there is no statistical significance in either case, as with the previous tasks it is then not possible to determine a causal relationship between the changes made and the time it took to solve a task. For task 7 there was a large improvement from setting A→B, but a large decrease in efficiency from setting B→A. Even though these are positive results in favor of the new design only setting A→B showed significance, meaning that it is not possible to derive any certain conclusions from them. For task 8 and 9 there was improvement for each setting, but no statistical significance for both cases. The results for task 10 were in favor of the existing design, showing better times for prototype A in both settings. This time only setting B$\rightarrow$ had significance, again leaving the findings not valid.

This means than only one task, when tested separately proved any statistical significance. For task 6 there is a statistical significance in both cases, where results from the test where users tested prototype B first had a slightly higher increase in performance. In this case both versions of the prototype performed much better when tested second, and so the changes do not seem to have made a great difference in terms of efficiency.

For all tasks combined, the data from both cases has statistical significance, which means that is possible to draw some potential conclusion from it. The data shows a higher percentage of increase in performance for the case where prototype A was tested first. As with task 6 both prototypes consistently performed better when being tested second, meaning that the effect of the changes made are difficult to pinpoint. Prototype B did perform on average 10% better than prototype A when tested second, and so it is possible that the combined changes made to the Existing Design have had a positive effect on the efficiency of the system.

Lastly, as assumed for all tasks individually, the results were similar to the ones for all tasks combined. Both prototypes performed slightly better when being tested second, and prototype B performed on average 9% better than prototype A when tested second. The conclusion is then that there has likely been a sequence effect involved, causing participants to perform better in whichever prototype was tested second. Such an effect is expected when running a within subject experiment, and so the interesting observation is that the new design performed on average 9-10% better regardless of order.

5.2.4 Analysis of the User Satisfaction Questionnaire

Each user was asked to answer the same questionnaire used in the initial testing phase (see Section A.1). This questionnaire was given to the participants right after all tasks had been completed, and before the beginning to complete the task in the next design. As before half the participants completed the tasks in the existing design first, and the other half in the new design first. The results from all questionnaires concerning the existing design were then combined, as were the results for the new design. It was then possible to see which of the prototypes had the highest score of user satisfaction.

Statement 1 - Navigating through the system was simple and enjoyable

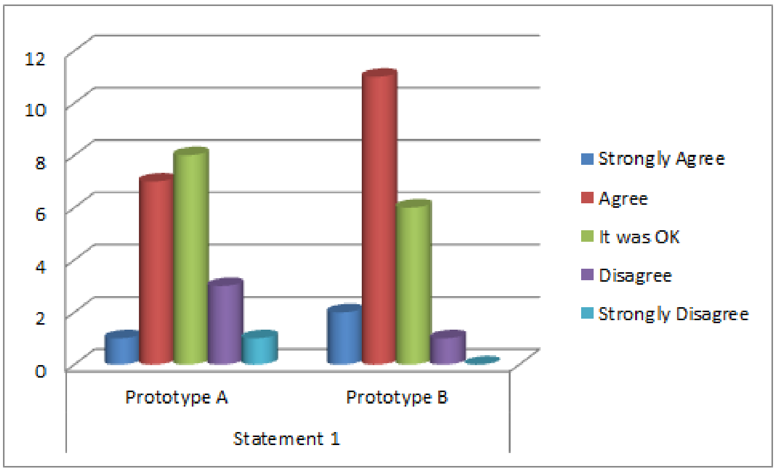


FIGURE 5.21: User Satisfaction: Navigation

Independent of whether a participant started with the existing design or the new design, there seems to be a slight increase in user satisfaction when it comes to navigating through the system. An increase in general satisfaction is assumed when more participants tend to agree or strongly agree with a positive statement. In both cases more participants agree with the statement, and less participants disagree (see Figure 5.21). As there were no major change in the way of navigating through the system, aside from the addition of the proceed button while selecting single tickets it is hard to say exactly why this increase in satisfaction occurred.

Statement 2 - I can accomplish what I want with few clicks

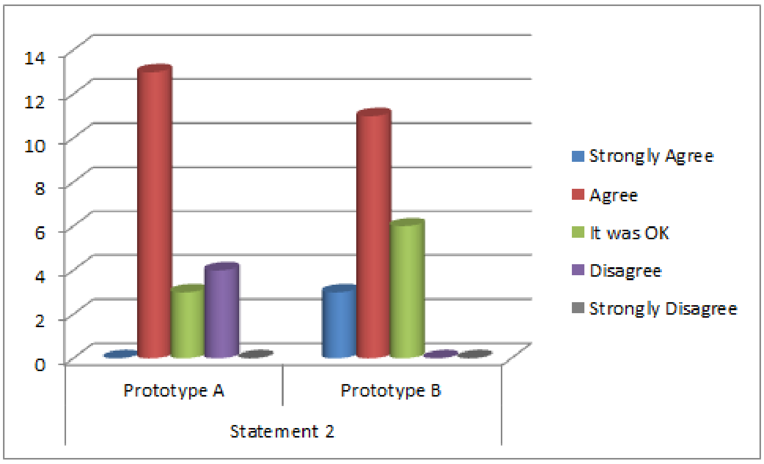


FIGURE 5.22: User Satisfaction: Number of Clicks

In both cases the participants also seemed to have a higher level of satisfaction when it comes to the efficiency of the system (see Figure 5.22). In fact while using prototype B first, there is a significant decrease in satisfaction, even though the participants still agree that prototype A is fairly efficient. The changes made based on the integration principle were designed to raise efficiency, and so this might be one of the reasons why the participants were more satisfied with the new design.

Statement 3 -Recovering from a mistake was quick and easy to do

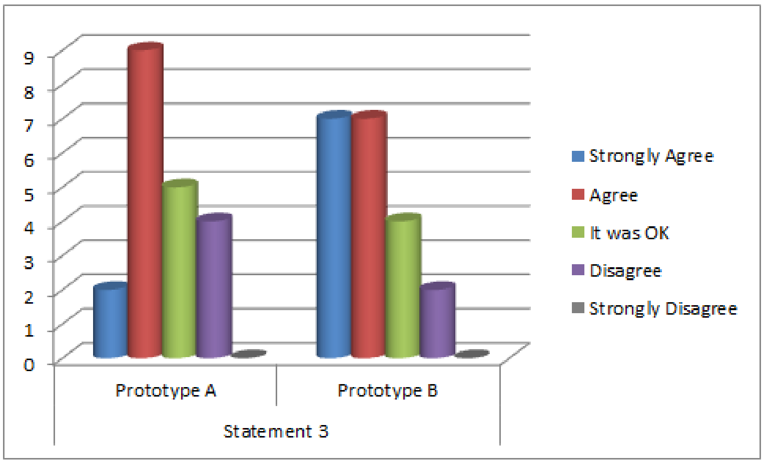


FIGURE 5.23: User Satisfaction: Recovering from Mistakes

When it comes to being able to recover from mistakes, no major changes were made. In both prototypes the participant had the choice of a cancel button and a back button. Even so the new design of the prototype still had the best results in the questionnaire, regardless of the order in which the prototypes were tested (see Figure 5.23). Possible reasons could be that participants were overall more satisfied with the new design and so felt that they performed less mistakes.

Statement 4 - Using the system is difficult

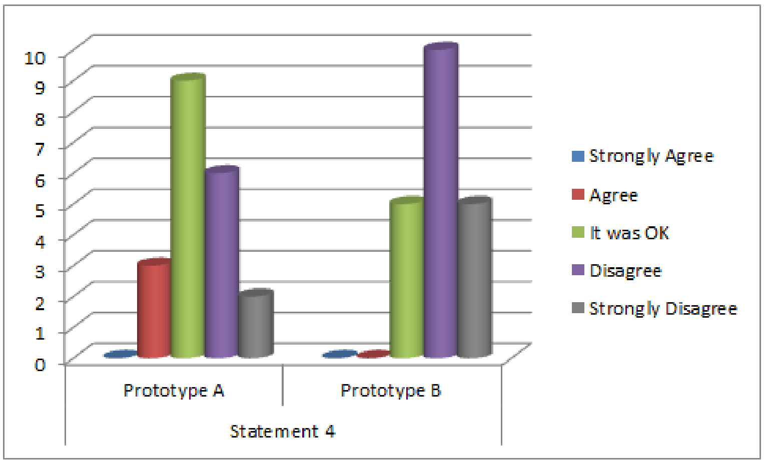


FIGURE 5.24: User Satisfaction: Is the System Difficult

When participants were asked if they thought the system was difficult to use, there was a significant change in the case where prototype B was tested first (see Figure 5.24). In general the participants did not feel any major difficulty while using the system, aside from a few users that tested prototype A first. As the question is generalized to the whole system, it is not possible to say exactly which change in the prototype that has lead to the perceived decrease in difficulty.

Statement 5 - It was easy to get help from the system if I needed it

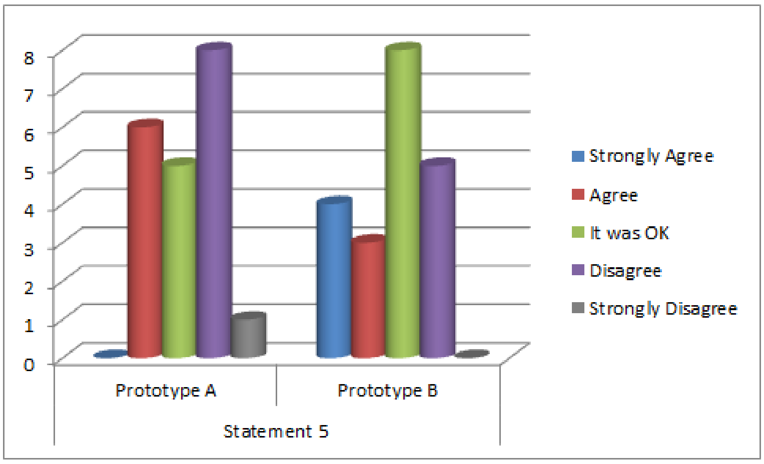


FIGURE 5.25: User Satisfaction: Getting Help

One of the major changes to the initial prototype was to the information screen, and the reduction to the amount of information buttons. As there has been an increase in perceived user satisfaction in both cases (see Figure 5.25), this change is likely one of the factors for the increase in positive answers.

Statement 6 - All functions were clearly visible and easy to find

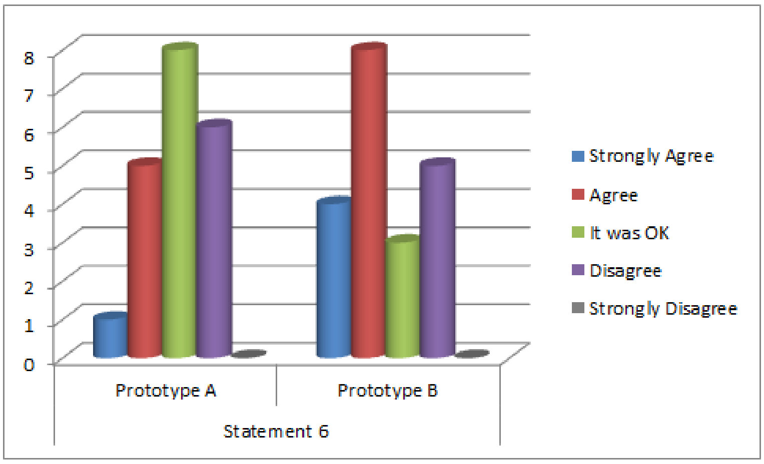


FIGURE 5.26: User Satisfaction: Visibility of Functions

In terms of the visibility of functions, there was also a slight increase in user satisfaction in both cases (see Figure 5.26). Based on comments that were made during the tests, one of the reasons for this was the enlarging and specification of the map buttons. Another possible factor can include the increase in font sizes.

Statement 7 - I don't notice any inconsistencies as I use the system

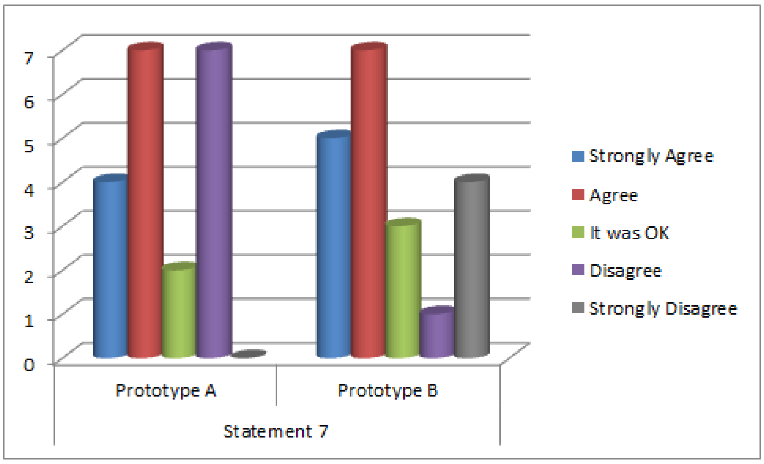


FIGURE 5.27: User Satisfaction: Consistency of the Layout

When asked if the participants noticed any inconsistencies in the system, the results were fairly spread between an increased and decreased sense of inconsistency (see Figure 5.27). In retrospect, one of the main issues with the new design of the prototype is that there is different information about the system under the same information button. Therefore some of the participants did not think to check the information button after scanning their card, as they assumed they would be given the same information as in the main information menu. This is likely the main reason why some users were unsatisfied with the consistency of the system.

Statement 8 - Overall, I am satisfied with this system

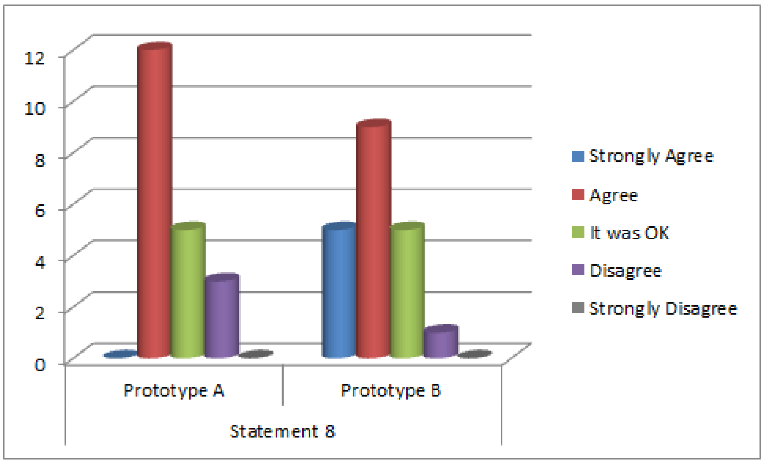


FIGURE 5.28: User Satisfaction: Overall Satisfaction

In terms of overall satisfaction with the system, most users were fairly satisfied with both versions of the prototype. For both cases the level of satisfaction is slightly higher for prototype B, indicating that for many users, this was their preferred version.

5.2.4.1 Summary

The participants were able to use both interfaces without needing much help. With a few exceptions of participants who could not find the map in prototype A, and some that could not find the information about "FleksiSkyss" prototype B. In general the results show a tendency of higher user satisfaction for prototype B. In fact only statement 7, concerning inconsistencies, showed a slight decrease in user satisfaction. As mentioned the likely explanation for this, based on comments made by participants, is that there is two different types of information under the same information button. One for the "with card" page and one for the "without card" page. Although the thought behind this was that users without cards should not need the information pertaining card based tickets, it seemed to confuse participants when being faced with tasks that required them to find information on both screens.

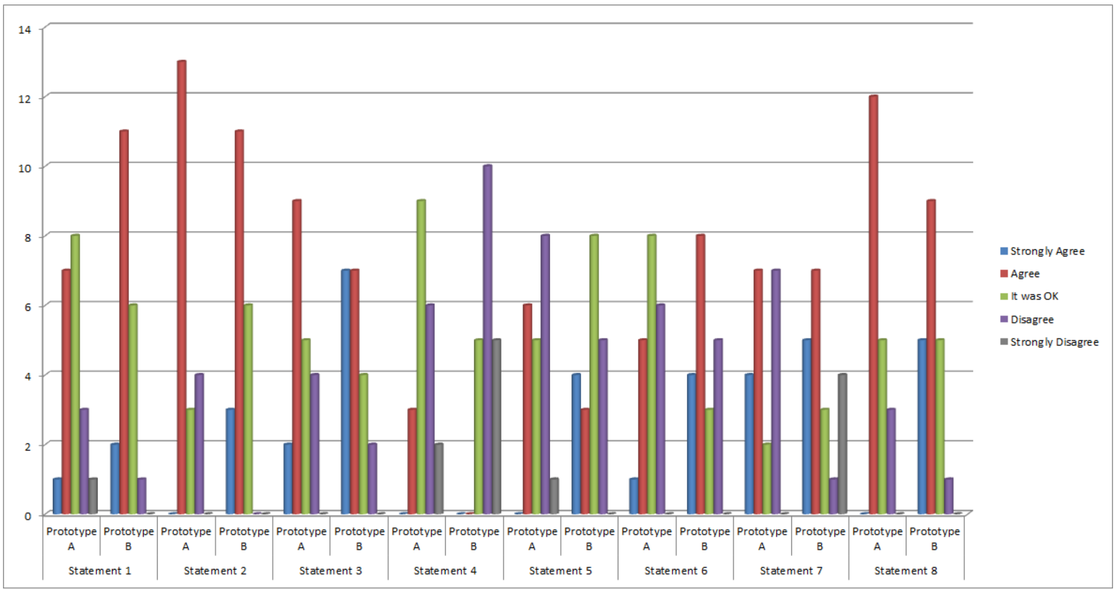


Figure 5.29: User Satisfaction: Overview

As seen in the within subject experiment in section 5.2.3.1, prototype A actually performed better than prototype B in many cases. This regardless of order. Even though much of the data yielded no statistical significance, it is worth considering why certain aspects of the existing design gave better results. A possible factor is that some participants could have had previous experience with the ticketing system, which could affect their efficiency as it would be familiar to them. Aside from actual preference, there is also a possible factor of why participants might have had higher satisfaction with the new design of the prototype. As the existing design is based on the current system, some participants might have been aware that prototype B was the new version, and as such could introduce some bias to their answers. Keeping in mind that some participants had used the existing design on a regular basis, and still found the new design to be more satisfying is still a positive result in terms of improvements based on the new design principle.

Chapter 6

Results and Discussion

This chapter will discuss the findings and thoughts about the self-service principles and their potential improvement of current affairs. Throughout this thesis the goal has been to explore methods of conducting research on self-service technology, and to discover design principles that has the potential to improve the usability of this type of technology.

During the design process it proved increasingly difficult to understand how to grasp the users, and their infinite number of opinions. One interesting observation made was that an issue pointed out by many participants during the initial study, the excessive use of information buttons, turned into a different issue when removed. Based on feedback from the focus groups and on consistency and simplicity the number of information buttons was reduced. The surprising discovery revealed itself during the final testing phase, when some participants after having tested both prototypes seemed to prefer the version where each ticket option had its own information button. This underlines the importance of a varied set of participants, and of working iteratively throughout the study.

In the following section the research questions will be presented once more, with focus on how the research has progressed in trying to answer them.

6.1 Main Research Question

How well do the existing design principles support the ongoing development of self-service systems and are they sufficient?

To answer this question a series of methods were implemented. A literature review with focus on design principles and similar theories such as the seven stages of action and heuristics. An observation process that focused on errors that real users made while interacting with the Skyss TVM. And finally a user testing phase with a prototype of the existing design, and a following set of focus groups.

The literature review helped uncover many of the existing principles. Many ties between the theories on design principles, heuristics and on the seven stages of action were also found. Ideas such as good feedback, consistent designs, constraining users and the affordance of interfaces are presented in some shape or form throughout the literature. What these principles often have in common is that they can be very general, and cover good practice in terms of several types of systems. There do exist heuristics for specific domains such as websites, but as of yet there are none for self-service technologies.

The observation uncovered many issues with the system, and possible solutions to these issues based on existing design principles. By making a connection between errors made and design principles it was possible to see whether or not the existing principles could have been used to remedy the errors. In essence many of the errors made by the users could have been either prevented, or made easier to avoid by adding either more visibility, better feedback or by adding constraints. In terms of efficiency it seemed as though one of the most time consuming actions that users encountered was to buy several tickets for a group of people. Several travelers went through the whole process of buying a single ticket, twice or more times even though a function exists for purchasing several tickets at the same time. In this situation visibility could be used to make the function more explicit, but might not be enough to improve the amount of travelers who use it.

Through the user testing of a prototype that simulated the existing design, and through focus groups with many of the same participants from the user testing, a clear view of existing issues with the TVM in particular emerged. Several of the issues uncovered during the observation were also mentioned during the focus groups, but not to the extent that was expected. One participant also had knowledge of design principles, and mentioned a lack of consistency in terms of the information icons displayed in the system. The suggestions made during the focus groups were tied to existing design principles where applicable, and also influenced the design of the new prototype. A common concern for the participants in terms of what was expected of a self service terminal, was that of a quick process. The existing design principles often addresses ways to make functions easier to perceive and to use. While this in turn can affect the efficiency of a process, there is no principle specifically detailing a way to increase how long a process takes.

Together these methods created a solid background for the thesis, and in addition gave way to the design principle of integration. They also showed that existing design principles do support the ongoing development of SSTs, but that there is still a need for more research on both principles and methods for further development of such systems.

6.2 Sub-Question 1

How can novel design principles be used to improve the usability of self-service technologies?

Following the formulation of the design principle of integration, this research question had the potential of being answered with the use of an actual novel design principle. The principle of integration was along with existing principles used to create a new version of the existing design for the Skyss ticketing system. Through a within subject experiment and following user satisfaction questionnaire the potential improvement to both efficiency and user experience has been determined.

While there was a lack of statistically significant data for individual tasks performed in the prototypes, the general results were positive in that there was a slight overall increase in efficiency. This also held true for the perceived user satisfaction, as the percentage of users who were satisfied with several aspects of the design had steadily increased for almost all questions posed. One important thing to be aware of is introducing confusion while implementing new design principles, or breaking with existing principles in the process. An example being that the participant was given different information after pressing the same information button in separate parts of the system. While the solution integrated the information with the functions around, it also introduced inconsistency across the design. Thus any new principle should be thoroughly tested with users before any final decisions are made.

In light of these results a good way to incorporate novel design principles in SSTs is to use them in combination with existing frameworks, and to perform specific changes in designs with them. This while keeping in mind possible conflicting design principles.

6.3 Sub-Question 2

Which methods are optimal for researching self-service technologies?

In the course of this thesis a variety of methods have been used to collect and analyze data. This data has been used to research a particular self-service technology. That of a self-service ticketing machine. Although the thesis has had a specific technology in focus, the methods used are thoroughly researched and should also be applicable for other self-service technologies.

In existing research a similar study on self-service ticketing machines has been performed by Siebenhandl et al. (2013), where the focus was on developing an improved TVM using a user centered approach. One of the main methods used in their study was an observation of the actual users of the machines. As self-service technologies rely so much on users being capable of operating them without assistance, such observation is a vital part of any study involving these types of technologies. This will reveal patterns of behavior that is otherwise very difficult to elicit under controlled circumstances. That being said, this thesis would not have been complete with observation alone. In order to analyze suspicions about why users make mistakes in SSTs it would also be advised to talk directly to the users in either focus groups or interviews. Performing more than one method will both raise the validity of findings, and produce more types of data to make the process of analysis much simpler.

It is therefore believed that a user centered approach is the most solid way of continuing research on self-service technologies. This does not mean that methods involving expert users are not valuable. They might become more prominent once SSTs become more understood, and start taking up an even bigger part of our daily interactions.

Chapter 7

Conclusion

This chapter will conclude the thesis with a summary of findings, discuss some limitations of the study and finally present ideas for future research on the topic.

7.1 Summary of the Thesis

This study was performed as an attempt to add something valuable to the field of HCI, specifically to the ongoing development of self-service technologies. The main focus being to expand on the existing design principles. As the media had recently been focused on passengers being wrongfully fined for not having a valid ticket, the decision was made to focus on the ticketing system used for Bybanen and the buses in Bergen. One main research question and two sub-questions were used to form the goal of the thesis.

- How well do the existing design principles support the ongoing development of self-service systems and are they sufficient?
  - How can novel design principles be used to improve the usability of self-service technologies?
  - Which methods are optimal for researching self-service technologies?

The findings showed that the existing principles are still very useful, and that existing research has its place in the domain of self-service technology. One novel design principle was produced with methods of observation, user testing and focus groups as the grounds for data gathering. This principle of integration was incorporated in a prototype that aimed to make tasks more efficient, and to raise the general satisfactions users have while using the system.

After rigorous user testing and analysis of data, the findings also indicate that the participants have experienced a sense of increase in satisfaction with the resulting prototype. On the other hand much of the data concerning the efficiency of separate tasks in the second prototype was statistically insignificant, and so the validity of some results must be questioned. When it comes to the findings based on the whole experiment the results are promising, showing slightly better results for the cases where the refined prototype was tested second.

7.2 Limitations of the Study

In all academic work there are some final considerations that must be taken into account. Would the method used produce the same results under the same circumstances, did the evaluation methods measure what they should, are the results distorted in any way and how much of the findings can be generalized?

As observational studies involve random users it is not possible to guarantee similar results, even if another researcher would perform an identical observation as described in this thesis. Therefore it would be advised to perform the full extent of methods in order to compare results. The evaluation methods followed strict rules, and the data gathered in both the within subject experiment and user satisfaction questionnaires are thus assumed to have measured what they were meant to. The largest limitation in this study is the limited sample of users, and the possible distortion in terms of participants potentially being aware of which prototype was the new version. In order to properly generalize and test the implementations proposed, a larger study would therefore have to be conducted. The study would have to include elderly participants, and a representative sample of users. It would also have to include several more iterations of a prototype in order to fully isolate the effects of any additional design principles.

7.3 Future Research

The field of self-service technology is still in its infancy, and will continue to grow at a rapid pace. Understanding which methods to use while evaluating and developing such technology requires more research. While this thesis focused mainly on the development of a new design principle, further research should also focus on how this process of gaining new insights can move forward. Thus leaving two very relevant areas of further research, methods of self-service research and a more exhaustive set of design principles or heuristics for self-service.

Bibliography

Axure Software Solutions (2013). *Interactive Wireframe Software & Mockup Tool.* URL: http://www.axure.com/.

Benson, Jim (2013). *Personal Kanban.* URL: http://www.personalkanban.com/pk/.

Blandford, Ann, Anna Cox, and Paul Cairns (2008). *Controlled experiments.* Cambridge University Press.

Boring, Ronald Laurids (2001). "User-Interface Design Principles for Experimental Control Software." In: *CHI.*

Bryman, Alan (2012). *Social Research Methods.* 4th. Oxford University Press.

Cai, Xinyuan (2009). "Principles of Human-Computer Interaction in Game Design." In: *2009 Second International Symposium on Computational Intelligence and Design*, pp. 92–95. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5368956.

Difi (2014). *Universell Utforming for IKT.* URL: http://uu.difi.no/.

Fagervoll, Maria (2014). *Gunnar ble lagt i jern ved Bybanen.* URL: http://www.bt.no/nyheter/lokalt/Gunnar-ble-lagt-i-jern-ved-Bybanen-2687433.html\#.Uw8Z3_mHh8E.

Fog Creek Software (2013). *Trello.* URL: https://trello.com/.

Geest, T van der (2013). "Introduction to the Special Section: Designing a Better User Experience for Self-Service Systems." In: *IEEE Transactions on Professional Communication* 56.2. URL: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6525145.

Hagen, Simen and Frode Eika Sandnes (2010). “Toward accessible self-service kiosks through intelligent user interfaces.” In: *Personal and Ubiquitous Computing* 14.8, pp. 715–721. URL: http://www.springerlink.com/index/10.1007/s00779-010-0286-8.

Hollan, James, Edwin Hutchins, and David Kirsh (2000). “Distributed cognition: toward a new foundation for human-computer interaction research.” In: *ACM Transactions on Computer-Human Interaction* 7.2, pp. 174–196. URL: http://portal.acm.org/citation.cfm?doid=353485.353487.

Holmlid, S (2007). *Interaction design and service design: expanding a comparison of design.* Linköping, Sweden.

James, Barnes. G, Dunne. A Peter, and William. J Glynn (1999). *Handbook of Services Marketing and Management.* SAGE Publications, Inc, pp. 89–101.

Kaptelinin, Victor (2013). “Affordances.” In: *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Ed. by Mads Soegaard Dam and Rikke Friis. Aarhus, Denmark: The Interaction Design Foundation. URL: http://www.interaction-design.org/encyclopedia/affordances_and_design.html.

Kvamme, Lars (2014). *Frøya (16) fikk bot selv om hun hadde billett.* URL: http://www.bt.no/nyheter/lokalt/Froya-16-fikk-bot-selv-om-hun-hadde-billett-3068591.html\#.Uw8BcPmHh8E.

Laerd Statistics (2013). *Wilcoxon Signed-Rank Test using SPSS.* URL: https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php.

Molich, Rolf and Jakob Nielsen (1990). “Improving a human-computer dialogue.” In: *Communications of the ACM* 33.3. URL: http://dl.acm.org/citation.cfm?id=77486.

Nesse, Irene (2013). *Personal Communication.*

Nielsen, Jakob (1990). “Heuristic Evaluation of User Interfaces.” In: *CHI.* April. URL: http://web.vtc.edu/users/cad03090/hci-r/heuristic.pdf.

– (1995). *How to Conduct a Heuristic Evaluation.* URL: http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/.

Nielsen, Jakob (2006). *F-Shaped Pattern For Reading Web Content.* URL: http://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/.

Nielsen, Jakob and Don Norman (2013). *Human Computer Interaction.* URL: http://www.nngroup.com/topic/human-computer-interaction/.

Norman, Don (2013). *The Design of Everyday Things: Revised and Expanded Edition.* Revised Ed. Basic Books.

Nunamaker, J F, M Chen, and T D M Purdin (2001). "Systems development in information systems research." In: *Journal of Management Information Systems* 7.

Oates, Briony J (2006). *Researching Information Systems and Computing.* Vol. 37. 18. SAGE. Chap. 1.

Pearson, Jennifer, George Buchanan, and Harold Thimbleby (2010). "HCI design principles for ereaders." In: *Proceedings of the third workshop on Research advances in large digital book repositories and complementary media - BooksOnline '10*, p. 15. URL: http://portal.acm.org/citation.cfm?doid=1871854.1871860.

QSR International (2012). *NVivo 10 Research Software.* URL: http://www.qsrinternational.com/products_nvivo.aspx.

Reed, B (2013). *Netflix and YouTube are the Internet's bandwidth consumption kings.* URL: http://bgr.com/2013/11/11/netflix-youtube-bandwidth-consumption/.

Robertson, Paul, A Szymkowiak, and Graham Johnson (2010). "Investigating the future of self-service technology." In: *Proceedings of the 24th BCS ...* URL: http://dl.acm.org/citation.cfm?id=2146354.

Sharp, H, Y Rogers, and J Preece (2011). *Interaction design – beyond human-computer interaction.* John Wiley & Sons, Ltd, p. 585.

Siebenhandl, K, G Schreder, M Smuc, E Mayr, and M Nagl (2013). "A User-Centered Design Approach to Self-Service Ticket Vending Machines." In: 56.2, pp. 138–159. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6515185.

Skyss (2013). *Kontanttillegg påbussen i Sone Bergen.* URL: http://www.hordaland.no/Aktuelt/Arkiv-nyhende/2013/Mai/Kontanttillegg-pa-bussen-i-Sone-Bergen/.

Ssemugabi, Samuel and R de Villiers (2007). "A comparative study of two usability evaluation methods using a web-based e-learning application." In: *Proceedings of the 2007 annual research ...* pp. 132–142. URL: http://dl.acm.org/citation.cfm?id=1292507.

StatisticsLectures (2012). *Wilcoxon Signed-Ranks Test.* URL: http://www.statisticslectures.com/topics/wilcoxonsignedranks/.

Stewart, D W (2007). *Focus groups: Theory and practice*, pp. 109–134. URL: http://books.google.com/books?hl=en\&lr=\&id=Rb9KlLtpGe8C\&oi=fnd\&pg=PR9\&dq=Focus+Groups+Theory+and+Practice\&ots=m05PmroYoO\&sig=fcrr40z377HtVCWpEYdFbo7pbzg.

Teräs, Sampo and Mari Mäkelä (2012). "Service design determinants for user value design: online store case study." In: *Proceedings of the 24th Australian Computer- ...* URL: http://dl.acm.org/citation.cfm?id=2414625.

Tognazzini, Bruce (2013). *Ask Tog.* URL: http://www.asktog.com/basics/firstPrinciples.html.

Vaishnavi, V and B Kuechler (2004). *Design Research in Information Systems.*

Wickström, Gustav and Tom Bendix (2013). "The "Hawthorne effect" —what did the original Hawthorne studies actually show?" In: *Scandinavian Journal of Work, Environment & Health.*

Wild, P J (2008). "HCI & the analysis, design, & evaluation of services." In: *HCI 2008*, pp. 207–208.

Appendix A

Questions and Tasks

A.1 User Experience Questionnaire

Question 1: What is your gender?

Question 2: In which age group do you belong?

The following are statements, where the participants chose an element from a likert scale ranging from strongly agree to strongly disagree.

Question 3: Navigating through the system was simple and enjoyable.

Question 4: I can accomplish what I want with few clicks.

Question 5: Recovering from a mistake was quick and easy to do.

Question 6: Using the system is difficult.

Question 7: It was easy to get help from the system if I needed it.

Question 8: All functions were clearly visible and easy to find.

Question 9: I don't notice any inconsistencies as I use the system.

Question 10: Overall, I am satisfied with this system.

A.2 Initial User Tasks

Task 1: Buy a single ticket for "Voksen" in "Sone Bergen" without a Skyss-card.

Task 2: Buy a single ticket for "Honnør" to "Haakonsvern 1" without a Skyss-card.

Task 3: Buy a single ticket for "Barn" using the quick choice.

Task 4: Buy two "Voksen" tickets at the same time for "Sone Bergen".

Task 5: Look up general information on how the ticketing machine works.

Task 6: Find the map of "Sone Bergen".

Task 7: Buy a "PeriodeSkyss" for 30 days with a Skyss-card for "Buss and Bybanen".

Task 8: Figure out how much money is on your "KontantSkyss".

Task 9: Play around with the prototype for a few minutes.

Task 10: Answer the Questionnaire.

A.3 Within Subject Design User Tasks

Task 1: Buy a single ticket for "Voksen" in "Sone Bergen" without a Skyss-card.

Task 2: Buy a single ticket for "Honnør" to "Haakonsvern 1" without a Skyss-card.

Task 3: Buy a single ticket for "Barn" using the quick choice.

Task 4: Buy two "Voksen" tickets at the same time for "Sone Bergen".

Task 5: Find information about the general use of the ticketing machine.

Task 6: Find the map of "Sone Bergen".

Task 7: Find information about how to use "FleksiSkyss".

Task 8: Buy a "PeriodeSkyss" for 30 days with a Skyss-card for "Buss and Bybanen".

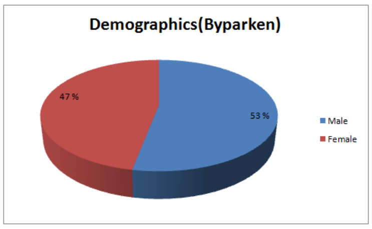
Task 9: Figure out how much money is on your "KontantSkyss".

Task 10: Find out how many days that are left on your "PeriodeSkyss".

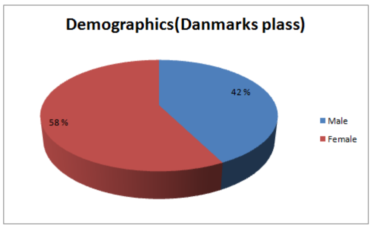
Appendix B

Figures and Tables

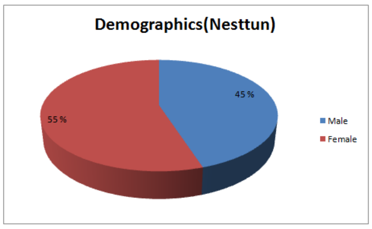
B.1 Charts from the Observations



(A) Genders at Byparken

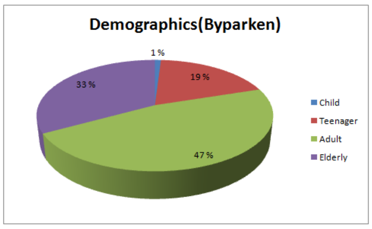


(B) Genders at DP

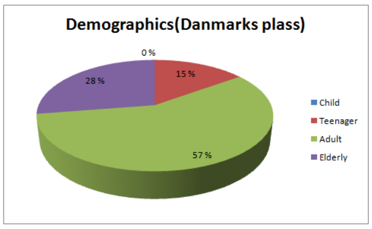


(C) Genders at Nesttun

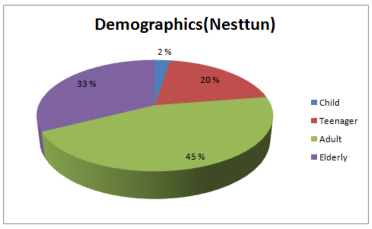
FIGURE B.1: Genders at the Stations



(A) Ages at Byparken



(B) Ages at Danmarksplass



(C) Ages at Byparken

FIGURE B.2: Ages at the Stations

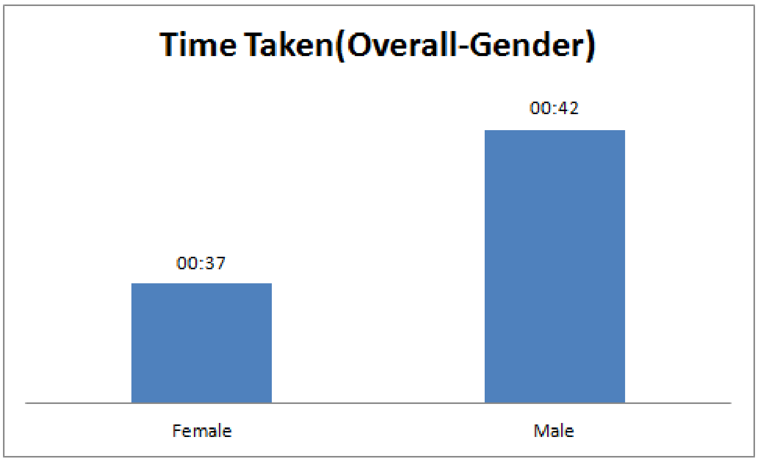


Figure B.3: Times Taken: Genders

Figure B.4: Times Taken: Tourists vs Locals

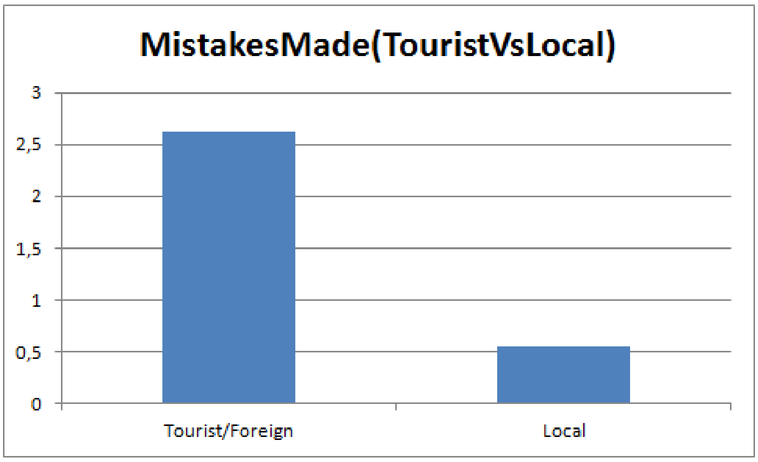


Figure B.5: Average Mistakes: Tourists vs Locals

Appendix C

Consent Form

Consent form for participation in a research project

I am asking you to be a subject in a research project called *Design and evaluation of a self-service application using novel design principles*. The purpose of this project is to evaluate self-service devices and to develop design principles that will improve their usability.

This requires you to evaluate a prototype and answer a small number of questions related to the prototype. It will take you about 10-15 minutes to complete the survey. I might also ask you to later participate in a focus group in which we will discuss some issues concerning your experiences with the system. This process will be repeated twice more with improved versions of the prototype.

The information and data I gather will be completely anonymous and you are free to withdraw your participation at any time. So no one else from work or from your family will ever know what your answers were. If you sign this sheet, it means that you read this form and that all of your questions were answered.

Dato: ______________________________ Navn: ______________________________