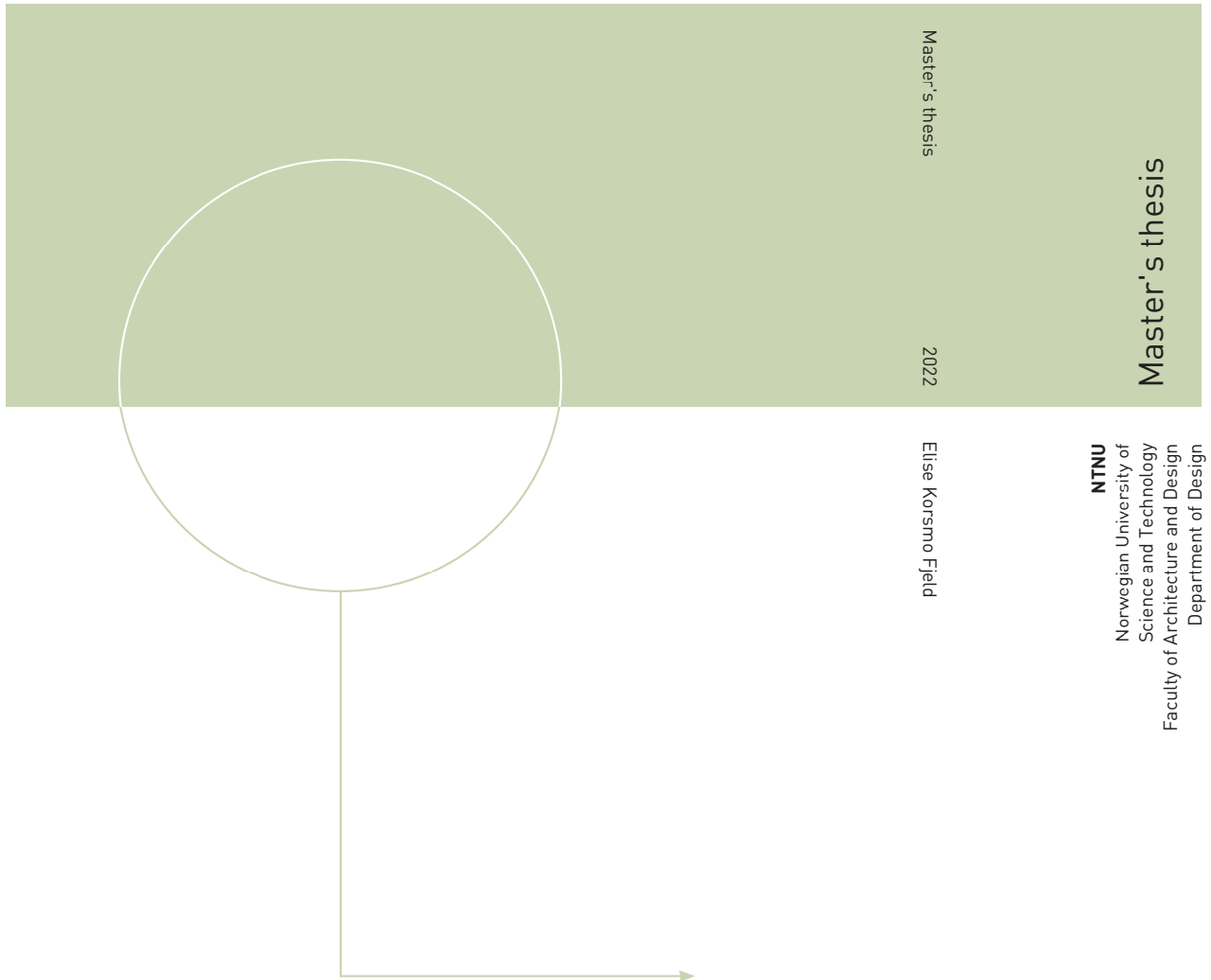


Master's thesis

2022

Elise Korsmo Fjeld

NTNU
Norwegian University of
Science and Technology
Faculty of Architecture and Design
Department of Design

Elise Korsmo Fjeld

Evaluating usability evaluation methods in the context of walk-up-and-use systems

June 2022

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Evaluating usability evaluation methods in the context of walk-up-and-use systems

Elise Korsmo Fjeld

Industrial Design
Submission date: June 2022
Supervisor: Erik Styhr Petersen
Co-supervisor: None

Norwegian University of Science and Technology
Department of Design

Evaluating usability evaluation methods in the context of walk-up-and-use systems

Master's thesis
Elise Korsmo Fjeld
Industrial Design, NTNU
Spring 2022

Preface

The report

This report is a documentation of my master's project at the Institute of Design at NTNU Trondheim during the spring semester of 2022.

The report is a collection of detailed descriptions of how I applied human-centred design and usability evaluation methods to two different walk-up-and-use systems, as well as reflections regarding the subsequent results, lessons learned, and my own efforts.

Acknowledgements

Although the work I conducted throughout this semester was all me, it would not have been the same without much needed help and support.

First, I want to thank my supervisor, Erik Styhr Petersen, who has steered me in a sensible direction every time I veered into a new rabbit hole, who calls me out on my antics, and who always helps me see things from a different perspective.

Thank you to everyone who participated anonymously in my many usability evaluation sessions. Without you, this project would have never happened.

Thank you to my colloquium who offered a place for me to express my frustrations and allowed me to sustain a realistic idea of what's expected of me throughout the project.

Lastly, I want to thank all my friends who have endured, not only the last semester beside me, but the last five years, and who have created a friendly and supportive environment in which there was always help, support and fun for those who needed it.

To the reader

Who is this thesis aimed at?

It may be enjoyed by academics who are interested in how a novice designer interprets certain established usability evaluation methods or ISO 9241-210:2019. It may be of inspiration for future master students wondering about formatting. It may be for my family and friends who are curious about what I've been learning these last five years.

But mostly it's for myself as a record of personal and professional growth, and of failures and accomplishments. I've been confused, angry, bored, and exhausted. I have also been curious, excited, eager and ambitious.

This master's thesis portrays *my* perspective on and application of human-centred design and usability evaluation, and what I have learned from the experience I have gained.

Abstract

Background

ISO 9241-210: 2019 is possibly the ultimate cornerstone of modern user-centred design. The essence of the standard is well known by most people in the field, and is seemingly easy to understand with its four steps. Usability evaluation, the third step in described in the standard, is an important basic principle to ensure that design solutions cover specified user requirements.

Walk-up-and-use systems are very simplistic in that they are not specialised towards specific user groups. In this thesis, the effectiveness of usability evaluation methods is examined by using such methods on two different walk-up-and-use systems.

Goal

The aim of this project is to contribute to literature with examples of how certain evaluation methods, with ISO 9241-210: 2019 as a basis, work in the context of walk-up-and-use systems.

Process

The project addresses two different walk-up-and-use systems, both of which are evaluated using usability evaluation methods. The results from these evaluations then form the basis for either making a design recommendation or a redesign. In the latter case, the redesign is re-evaluated. Based on all these exercises, a thorough reflection is made on the execution and the effectiveness of the methods in the given contexts.

Result

The result of the project is documentation of how the selected usability evaluation methods worked when used on selected walk-up-and-use systems. The results from the exercises show that the methods work excellently for what they are intended for, but that it can be valuable to look at a system within a larger context. In addition, it is made clear how important it is to do a thorough job in the earlier phases of ISO 9241-210: 2019.

Sammendrag

Bakgrunn

ISO 9241-210:2019 er muligens den ultimate grunnsteinen i moderne brukersentrert design. Essensen til standarden er godt kjent av de aller fleste innen fagfeltet, og er tilsynelatende lettfattelig med sine fire steg. Innenfor standarden finner man blant annet brukbarhetsevaluering som et viktig prinsipp for å sikre at designløsninger dekker spesifikke brukerbehov.

Gå-bort-til-og-bruk-systemer er svært enkle i form av at de er lite spesialiserte mot spesifikke brukergrupper. I denne avhandlingen blir effektiviteten til brukbarhetsevalueringsmetoder undersøkt ved å bruke slike metoder på to ulike gå-bort-og-bruk-systemer.

Mål

Målet med dette prosjektet er å bidra til litteraturen med eksempler på hvordan visse evalueringsmetoder, med ISO 9241-210:2019 i grunn, fungerer i kontekst av gå-bort-og-bruk-systemer.

Prosess

Projektet tar for seg to ulike gå-bort-og-bruk-systemer som begge blir evaluert ved hjelp av brukbarhetsevalueringsmetoder. Resultatene fra disse evaluteringene ligger så til grunn for å enten komme med en designanbefaling eller et forslag til nytt design. I det siste tilfellet blir redesignet evaluert på nytt. Ut i fra alle disse øvelsene blir det gjort en grundig refleksjon rundt utførelse og metodenes effektivitet i de gitte kontekstene.

Resultat

Resultatet av prosjektet er dokumentasjon på hvordan de valgte brukbarhetsevalueringsmetodene fungerte i bruk på utvalgte gå-bort-og-bruk-systemer. Resultatene fra øvelsene viser at metodene fungerer utmerket til det de er ment for, men at det kan være verdifullt å se på et system i en større kontekst. I tillegg kommer det fram hvor viktig det er å gjøre et grundig arbeid i de tidligere fasene av ISO 9241-210:2019.

Table of contents

Introduction

Motivation

Last semester I explored ISO 9241-210:2019 in the context of maritime industry. Through that project I was confronted with the fact that the standard is not as straight forward as it seemed to be – especially to me as a novice designer.

With that in mind, I wanted to delve deeper into the world of human-centred design, only this time with the focus on usability evaluation methods.

Project description

In the project description as seen on the next page I stated a clear goal of improving the user experience of a selection of designs that people struggle with. The project description was purposely kept open to make room for new discoveries down the road, but it turned out to be insufficient.

In the first half of my project after evaluating the design of AtB's ticket machines, I concluded that I was wrong by hypothesising that the machines have low usability, thus contradicting my project description.

Instead of starting anew with a different design, I found my findings very interesting and wanted to tweak the goals of the project accordingly. I changed my focus from the increase in usability gained by involving users to the evaluation methods themselves and how they served me. This also allowed me to widen my toolbox and explore evaluation methods that are not user-based.

Thus, my new goals are:

- To evaluate the usability of a small selection of designs.
- To provide design fixes or design recommendations based on the aforementioned usability evaluation.
- To provide a commentary on the effectiveness of the chosen evaluation methods and my application of them.

NTNU
Fakultet for arkitektur og design
Institutt for design

Masteroppgave for student Elise Korsmo Fjeld

Tittel Utforsking av viktigheten av brukerinvolvering i produktutvikling
Title Exploring the importance of user involvement in product development

Bakgrunn for oppgaven:
Dag til dag blir man ofte konfrontert med løsninger som er mye vanskeligere å skjønne enn det burde være. Burde man måtte lære seg å bestille bussbilletter, åpne en dør eller stille vekkerklokka? Når slike ting ikke er åpenbart fra første gang aktiviteten utføres kan det føre til mye frustrasjon og misnøye, noe som er helt unødig når løsningene sannsynligvis bare trenger å designes litt bedre.

Oppgavens innhold:
I dette prosjektet skal jeg utforske fasene i HCD som involverer brukere og stakeholders, altså innsiktsarbeid og brukertesting. Jeg skal sette meg inn i relevant teori, og bruke brukerinvolvering i mitt forsøk til å forbedre brukeropplevelsen av et lite utvalg produkter som mange sliter unødig med.

Oppgavens gjøremål:
- Samle innsikt om og brukerteste utvalgte produkter og tjenester
- Foreslå forbedringer av produktene basert på innsikt og brukertester
- Utforskning av temaet gjennom litteratur og annen eksisterende informasjon

Oppgaven utføres etter "Retningslinjer for masteroppgaver i Industriell design".

Ansvarlig faglærer (hovedveileder ID): Erik Styhr Petersen
Biveileder:
Bedriftskontakt:

Utleveringsdato: 7. januar 2022
Innleveringsfrist: 7. juni 2022

NTNU, Trondheim, dato — Sara Brinch

Navn — Sara Brinch
Ansvarlig faglærer — Instituttleder

Scope

I chose to limit my scope strictly to the usability itself and the designs as closed systems. I want to argue that when a user starts to interact with a design that is meant to be walk-up-and-use, it is irrelevant who the user is, why the system was brought into existence, and how it creates value for the owner of the system; If a user wants to use a system, they should simply be offered a good user experience.

This is an argument I have struggled to accept myself throughout the project, as no systems exist in isolation in the real world, and I have time and time again felt the urge to look outwards. This is, however, irrelevant when my goal is to comment on the effectiveness of my chosen usability evaluation methods. It does not matter where I set the limits for what I evaluate, as the methods will provide results either way.

The project

To allow myself to focus on the methods rather than the domain, I chose to apply the methods to so called walk-up-and-use systems. From there the basis of my project is very simple: Evaluate the usability of a couple of walk-up-and-use products and provide redesigns or design suggestions based upon said evaluation.

In the end I opted for two different systems: AtB's ticket machines and Flybussen.no. I wanted to roughly stay within the same domain and coincidentally landed on ticket purchasing for public transport. Otherwise the two systems are fairly different as one is a physical construction while the other is a website, one is simple while the other is more complex.

Walk-up-and-use-systems

What are they?

There is no single definition of what a walk-up-and-use system is, but as the "walk-up-and-use" phrase implies, it generally means that anyone should be able to walk up to the system and use it effectively without any prior training. This is how I choose to define it. Examples include parking-metres, lifts and self-checkout machines.

Usually, a design is aimed at a certain user group, but when designing a walk-up-and-use system you are designing for everyone. This term is relative as it can be argued that a service's user group is limited to the people who have access. Thus, it is strictly speaking not for *everyone*.

Who is everyone?

It wouldn't be a good idea to invest enough resources into an interface design that even an alien could use it. A design will never be universally inclusive and at some point, a line must be drawn.

In this thesis I will be evaluating the usability of two products that I consider to be walk-up-and-use: AtB's ticket machines and Flybussen's website. I have no access to any customer data, and I won't be pursuing any, but I wish to argue that it's irrelevant. As soon as I have defined a service as walk-up-and use, I am saying that everyone is a potential user.

In the case of both AtB and Flybussen, I assume everyone to be anyone who travels in Norway, be it commuting or vacationing. There are heaps of user groups within this definition who are unlikely to use these services, but it's an important principle that they should be able to do so.

Am I the user?

Although the answer should be an obvious no, if I'm being honest with myself, I do catch myself thinking the opposite, since I most definitely fall within my own definition of who the user is.

The False-Consensus Effect

Biases may be pervasive even when the designer is not part of the user group, as they may have an opinion on how they would use the design and assume that the users are similar to themselves. This effect is called *The False-Consensus Effect* and was first described in Ross et al. (1977). It can be summed up as a bias caused by observers (e.g. designers) overestimating the degree to which others share their beliefs, opinions, habits etc. They will perceive alternative response as peculiar or rare.

Falling for this effect as a designer can be catastrophic as it may affect an unlimited amount of people's lives, depending on the product or system. This highlights the importance of testing your designs.

Being aware of my own bias

I believe it's important for me to point this out explicitly as it undoubtably will become a pervasive theme throughout both my subprojects. It is not possible for me to measure my own biases and I don't believe I will be able to eliminate them, but my hope is that by being continuously aware of them I will be able to minimise their effects.

I will be attempting a particularly risky activity in the case of AtB's ticket machines where I conduct a usability test with myself. This may seem counterintuitive as I just explained how dangerous it is to confuse yourself with the user. My motivation is mainly to use the test results to design subsequent usability tests, and not to substitute myself with a user.

As for Flybussen's website, I did a similar exercise, although not as thorough, with the same motive of learning about the system before evaluating its usability.

I am not collaborating with either AtB or Flybussen and won't have access to information they might be sitting on. Therefore, with me being the one to assume and define the main functionalities of the two systems, I acknowledge that both the analysis of evaluation data and subsequent designs may be skewed by my biases.

Figure 1: An example of a walk-up-and-use system in the shape of a railway ticket machine. Photo by Sam Balye on Unsplash.

Applied theory

In this chapter I will be going through the basic theory that my project is based upon and briefly how they will be applied. ISO 9241-210:2019 defines the entire human-centred deisgn process from planning to end. As already presented in the introduction, I will be focusing on usability evaluation methods, which is is only one step in the four step process. Therefore, I will be omitting many of the otherwise important steps in the process, and compressing others.

Human-centred design

Human-centred design (HCD) is defined in ISO 9241-210:2019 by the International Organization for Standardization and aims to provide "*requirements and recommendations for human-centred design principles and activities throughout the life cycle of computer-based interactive systems*" (ISO, 2019).

The HCD process can be used to create new design solutions that cover certain user needs, but also to improve existing design solutions, which is what I will be doing in this thesis. The next four subsections will briefly explain each respective step of the HCD process, as well as how I plan to utilise them during my project.

Understanding the context of use

In the initial phase of the HCD process (after the planning), information is collected in order to, not surprisingly, understand the context of use. Only when I understand the context in which the system will be used will I be able to identify the user needs that I subsequently will design a solution to cover (ISO, 2019).

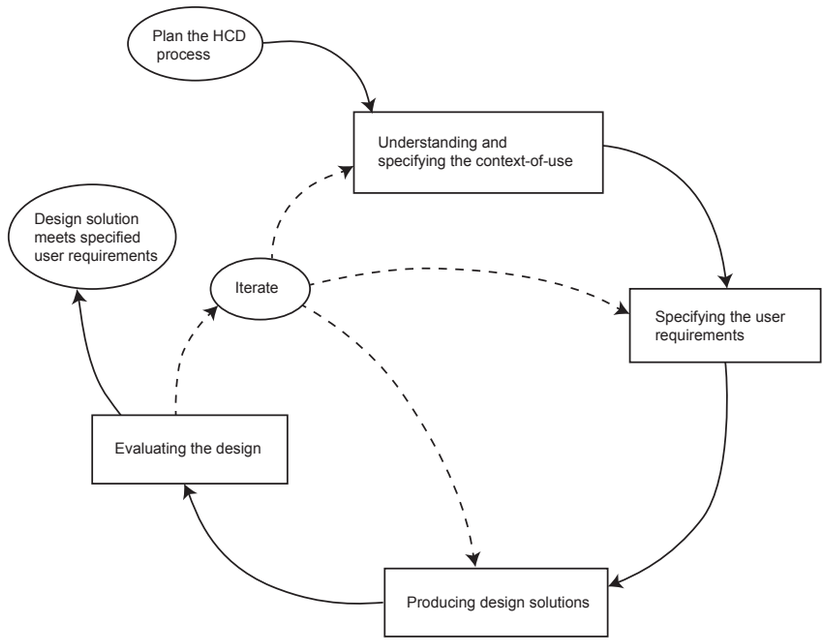


Figure 2: A graphical representation of the human-centred design process.

A thorough understanding of the context of use should include the following (ISO, 2019):

1. An overview of all the different users and stakeholders and their needs, including their relationship with the development of the system to be designed.
2. User characteristics such as skill, demographics, experience etc.
3. User goals and tasks in terms of how users carry out tasks, and whether they can be carried out incorrectly.
4. The environment of the system in terms of technical aspects, the physical area, cultural context, lighting etc.

Since I will be working with walk-up-and-use systems, and to limit my scope, I will be making several assumptions of the context of use in terms of all four list items. I already defined the user group on the basis of what a walk-up-and-use system is. I am in a way conducting the project in a vacuum, without a project owner or a development team, and I acknowledge that results may

differ if I were to involve the companies as stakeholders. However, as I have scoped this thesis to mainly revolve around evaluation methods, I chose to not seek out this information.

Regarding the environment of the systems (list item 4) I base my assumptions on what appears to be the main functionalities of the system, which in both subprojects is ticket purchasing.

Specify user requirements

After collecting the information required to understand the context of use, an explicit statement is formulated. The specification of the user requirements should include the following (ISO, 2019):

1. The intended context of use.
2. The context of use, in terms of physical limitations and environment or safety concerns.
3. Any ergonomic requirements or requirements related to e.g., standards.
4. Objectives and measurable usability goals.
5. Organisational requirements directly affecting the user.

As already stated, I will not be taking business goals into consideration since I have no businesses to answer to. I am solely concerned with the usability of the systems and the usability evaluation methods.

User requirements will be formulated based on what I consider to be the main functionalities of the systems, my assumptions about the context of use, and any other information I may gather through getting to know the system.

I am not concerned with quantitative scores or measures as any improvement will be enough to prove the effectiveness of the method.

Produce design solutions

With a solid statement of user requirements ready, a design solution can be produced. Creating design solutions should include the following (ISO, 2019):

1. User tasks, and an interface that takes the whole user experience into consideration.
2. More concrete design solutions, e.g., through creating prototypes, scenarios etc.
3. Changes to the design based on results from user-centred evaluation if it has previously been conducted.
4. Handing off the designs to whoever is responsible for implementation.

After conducting a usability evaluation on either system, I will methodically "correct" the mistakes based on the results. Handoffs will not be relevant as my design solutions are not to be implemented.

Evaluating the design

Arguably the most important, and last step of the HCD cycle, a usability evaluation is conducted in order to test whether the design solution fulfils the specified user requirements, and whether any unexpected usability problems are to be found. The activity includes the following (ISO, 2019):

1. Obtaining resources for any usability evaluations that may be conducted during the development or lifetime of the design.
2. Planning the evaluation to make sure it can be completed within the project's timeframe.
3. Conducting the usability evaluations.
4. Analysing the results, and subsequently prioritising tasks and suggesting design solutions.
5. Communicating said design solutions to the design team.

This is the step in the HCD process that I have chosen to focus on in my thesis.

It's relevant to note that user-centred evaluation does not have to be user-based. This means that users aren't required to be involved in all evaluation processes. The alternatives are either an inspection-based evaluation, which I will be doing with Flybussen's website, and long-term monitoring where user input is collected over time (ISO, 2019).

ISO (2019) also argue against always involving users in the evaluation because of cost and time constraints, especially in early phases of the project.

Usability testing

Usability testing, or user-based testing (ISO, 2019), is a category of usability evaluation methods where users are involved. The purpose of such a test can be to either, uncover usability problems in a design, to learn more about the user, or to uncover potential for improvement (Moran, 2019).

Usability testing can be conducted any stage in the HCD process, including the conceptual stage, prototype stage, and high-fidelity stage (ISO, 2019). In my case, with both AtB's ticket machines and Flybussen's website, the usability testing will be conducted at the stage where the system has already been implemented.

Usability testing is predominantly used in the "evaluating the design" step of the HCD cycle, but I will also be using it partly to specify the context of use in the case of AtB's ticket machines.

Usability testing according to Nielsen

The usability testing I will be conducting in my project will be based on Nielsen (1993). The method is described in somewhat general terms but can be summed up in four stages. The following sections will present a short description of each respective step.

Preparation

At the preparation stage the responsible experimenter makes sure that practical matters, such as meeting rooms and technical equipment, are in the correct state for the test to begin. Any scripts need to be laid out and the correct screens need to be turned on (Nielsen, 1993).

Introduction

The introduction is the first stage the test participant will be present for. The facilitator communicates practical information such as the purpose of the test, what the agenda is, what is expected of the participant etc. It's also important to state clearly that the purpose of the test is to evaluate the design, and not the participant, and that any data will be used to improve the design (Nielsen, 1993).

The test itself

During the actual test the participant is given predefined tasks that involve the design in question. At this point, it's crucial that the facilitator does not interfere with or interrupt the participant too soon. How the facilitator should act can be summed up in four simple rules (Nielsen, 1993):

- Don't interact with the user
- Don't express anything that may be interpreted as hints or feedback
- Don't help the participant unless they are very clearly stuck
- Have only one person act as the talking facilitator

To gain the best insights possible, it's important to see how the participant solves tasks without any help or guidance. Any communication with the participant should therefore be limited to signals that the facilitator has heard the participant, or any practical information that the participant may need (Nielsen, 1993).

Usually, the participant is asked to "think out loud". The participants vocalise their thoughts while performing the given task, which allows the facilitator to both understand which aspect of the interface the participant is processing and why. This version of the method is called *Concurrent Think-aloud* where the thinking aloud happens in parallel to the test. This is the most common version. The alternative is *Retrospective Think-aloud* where the participants complete the task in silence and give comments after the fact while watching a replay of the tasks being performed (Martin & Hanington, 2012).

Debriefing

After the test is completed, the facilitator follows up the test with any questionnaires they may have, and a discussion about the test. This stage is good for clearing up any questions that may arise during the test, such as why the participant made certain choices or what they found difficult (Nielsen, 1993).

After the participant has left, the facilitator should document the results in as much detail as possible. It's important to not wait too long as memory fades quickly (Nielsen, 1993).

Usability inspection

To evaluate the usability of Flybussen.no I wanted to use usability inspection. I opted for this evaluation method instead of usability testing because I had never practiced it before and saw it as a great learning opportunity.

In layman's terms, a usability inspection aims to improve a system, not by evaluation with users, but rather with usability experts or other professionals. Because of its cost effectiveness it can in some cases act as a substitute for usability testing, it can be used to eliminate bigger usability problems before involving users, and it can be a good complementary method to usability testing (ISO, 2019).

In this chapter I will give a brief introduction to usability inspection as a method, with a closer look on the usability inspection method I chose for my project and why.

Usability inspection methods

Similarly to usability testing, I wanted to keep to one single "recipe" of the method. I chose Nielsen and Mack (1994) as my main source. The book describes eight different usability inspection methods (Nielsen & Mack, 1994):

- **Heuristic evaluation:** A selection of usability experts evaluates the usability of an interface according to a particular set of heuristics.
- **Guideline reviews:** A user interface is evaluated up against a comprehensive list of usability guidelines. This method is similar to heuristic evaluation but is much more comprehensive as the list of guidelines can be very long and require a high degree of expertise.
- **Pluralistic usability walkthrough:** A group composed of usability experts, developers, user etc. walk through a scenario together.
- **Consistency inspection:** Designers representing other design projects check if the design acts similarly to the product they themselves are working on.
- **Standards inspection**: An interface is evaluated according to a standard by an expert on the standard in question.
- **Cognitive walkthrough:** A detailed scenario is used to simulate a user's choices and problem solving at each step in a dialogue to see if it's likely to leads to the intended next step in the dialogue.
- **Formal usability inspection**: A six-step process where participants with specific roles evaluats an interface both individually and in groups.
- **Feature inspection:** A method mainly meant for engineers to systematically go through different features and sequences of clicks.

A usability inspection can generally be said to have two different goals: 1) To improve the system, or 2) effects beyond the system itself, such as to educate the development team or to develop a new usability inspection method (Karat, 1994). My case aligned mostly with goal 1, but I also wanted to see how this method compares to "regular" usability testing, which falls under the umbrella of goal 2.

More specifically, I had the same goal with the usability inspection as I would have had with a usability test, which is to identify usability problems and subsequently use the findings to improve the system, and to document my experiences with how the method works in practice

Therefore, when choosing a usability inspection method to use, I needed to carefully consider my constraints and be clear about my requirements. The method needed to be cheap, flexible, require few different roles, and be executable with only one facilitator.

Heuristic evaluation

I opted for heuristic evaluation as Nielsen and Mack (1994) claims it to be cheap, fast, easy to use, and even describes it as "*one of the main discount usability engineering methods*."

The different stages of heuristic evaluation can be summed up in four steps, which will be explained briefly in the next four subsections. The following procedure is based on a case study described in Nielsen and Mack (1994).

Training session

The training session is meant to mainly get the evaluators up to speed on three areas (Nielsen & Mack, 1994):

- The domain in which the design to be evaluated resides.
- Heuristic evaluation as a method and the heuristics that will be utilised.
- The scenarios the evaluators will be basing the evaluating on.

Such a training session is particularly important if the design and/or domain is complex. This is to ensure that the evaluation data reflects real use cases to the maximum extent (Nielsen & Mack, 1994). Walk-up-and-use systems, on the other hand, are meant to be usable without training and their purpose should be obvious. Therefore, it could be argued that teaching the evaluators about the system beforehand could influence the evaluation results negatively.

Individual inspection

Each evaluator conducts the evaluation individually. Their main task is to identify as many usability problems as possible, and state them explicitly. The evaluators should also state which of the given usability heuristics the identified usability problem violates. However, focusing on the heuristics during the test may be distracting and result in fewer usability problems found. Therefore, it may be favourable to allow the evaluator to do the categorisation after the evaluation is completed (Nielsen & Mack, 1994).

Debriefing

The debriefing session should be conducted together with everyone who participated in or observed the evaluation. The main purpose of the debrief is to discuss the interface and collect ideas for high-level redesigns that will fix several identified design problems in one go (Nielsen & Mack, 1994).

Severity rating

As time and resources can be scarce and usability problems many, it is useful to prioritise them sensibly. If there isn't enough time to fix all the problems, then at least the most important ones should be fixed. When deciding the severity of a usability problem four factors are taken into account (Nielsen & Mack, 1994):

- How frequently the usability problem occurs.
- How much impact the usability problem has on the user experience.
- To which degree the user can learn to work around the usability problem.
- Market impact.

A severity survey, including a complete collection of usability problems identified by all evaluators, is sent out to said evaluators who are asked to rate the severity of each problem. Doing it this way is beneficial as the usability problem each evaluator finds will not completely overlap. The survey should be answered individually (Nielsen & Mack, 1994).

Heuristics

An experimenter may choose whichever usability heuristics they wish (Nielsen & Mack, 1994). I chose to use Jacob Nielsen's 10 general principles for interaction design (Nielsen, 1994) as the usability heuristics for the evaluation, which I assumed everyone to be familiar with. It's not a heuristic per se, but I added "bugs" to the list because I knew it would be likely that many usability problems would fit into this category.

Evaluating AtB's ticket machines

In this chapter I will be going through my process of evaluating the usability of my first walk-up-and-use system of choice. In this first instance it will be AtB's ticket machines, a fare option for public transport in Trondheim.

About

Trondheim's local bus company, AtB, offers many fare options, as shown in the picture below (Figure 3). AtB heavily promotes their AtB app where the traveller can both purchase tickets and plan their journey, which is probably the most popular option as it has all ticket options available, including period tickets.

In 2018 AtB started upgrading a large portion of their bus stops, many of which included brand new ticket machines, and in 2019 introduced their new metro buses. Another significant change was the end of ticket sales on board the buses. Large signs on the bus stops are now clearly stating that tickets must be bought before boarding a bus, while passengers could earlier buy their tickets directly from the bus driver.

Ticket machine design

It's not uncommon to stumble upon a badly designed ticket machine.

I could tell you about my own experience taking the train from Gatwick to London, finding out on my arrival that I bought the wrong ticket, which of course ended up costing me a penalty fee. It turns out I'm not alone. A 2017 article from the Guardian presents various concerning statistics regarding ticket purchases made from machines, reporting that as much as 20 % of rail passengers buy the wrong ticket (PressAssociation, 2017).

If you google "why are ticket machines so bad" you will inevitably get countless hits with blog posts ranting about the horror of bad ticket machine design. That's probably to be expected given the search word but switching out "bad" with "good" will actually land you with a very similar search result.

With all this in mind, I hypothesised that I would find several usability problems with AtB's ticket machines.

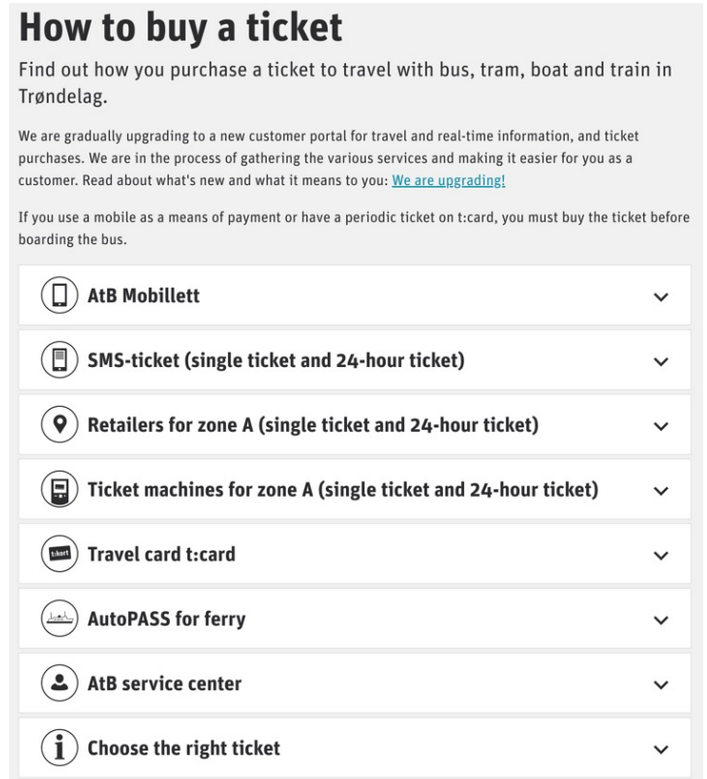


Figure 3: Screenshot from atb.no/billetter showing all the different ticket purchasing options that AtB offers.

Defining the context of use

Not knowing anything about the ticket machines and how they work, I wanted keep the definition somewhat open and have the usability tests provide me with more information that could help me define it further.

However, I have already defined the users to be everyone, AtB to be an additional stakeholder but not one that I will bring into the mix, and the physical environment to be anywhere a ticket machine is placed at any time of day.

I realised that the method of usability testing could do more for me than simply evaluating the usability of the machines. It could also provide a chance to more clearly define the context of use, and thereby the user requirements.

At this point, I felt like I knew too little about the machines to make satisfactory assumptions about the context of use.

First, I needed to learn how the ticket machine is operated.

Testing the ticket machine myself

I have previously never used AtB's ticket machines. Having lived in Trondheim for several years, I noticed when they popped up around town on seemingly random bus stops, but I never had an interest in taking a closer look. Therefore, I had no preconceived notion of how they may work. This all laid an almost ideal foundation for testing a ticket machine myself. The idea was that by testing it with myself as a user before testing it with recruited participants, I would be able to get to know the system and its flaws, and subsequently design better user tests.

To prevent myself from giving myself hints for the usability test I avoided any information that would help me complete the tasks I set for myself, such as where the machines are placed*.

* Out of curiosity, I later tried to see if I could find this information online. I could not find anything on my first try, but a few weeks later I stumbled upon it by accident while looking for some other information.

Test design

To make it as difficult as possible, I decided I would start at a bus stop without a ticket machine, go after dark and to not use my phone for help. By upping the difficulty of the task, I theorised that I would find more usability problems than I would otherwise.

The scenario I set up for myself was that I was going to take the bus, but I had forgotten my phone and card (maybe I had one of those phone covers with slots for cards that many people have). I would "coincidentally" find some cash in my pocket.

With some tweaks in the story, I assume this to be a realistic scenario, for senior travellers – also native to Trondheim as the ticket machines are relatively new. Seniors are statistically less likely to own a smartphone with only 79 % of the ages between 75 and 79 owning a mobile phone or a smartphone, compared to 97 % of the ages between 16 and 74 (SSB).

It might be their first bus trip after AtB stopped selling tickets onboard each bus and they have therefore not yet tried the ticket machines.

Test execution

In this section I will present how the service works through a detailed recount of my own user experience from beginning to end.

I arrived at Dragvoll bus stop, a bus stop without a ticket machine. I already knew beforehand that the poster on the bus stop displays a map with directions (Figure 4) for the closest ticket machine. In this case it was Lohove bus stop. I had never been to Lohove bus stop before and was not sure how to get there.

Normally, I would have followed the GPS on my phone. Not giving myself that opportunity I did not feel confident that I would 1) be able to remember the map, and 2) be able to locate my destination with the map as my only help. Looking at the map, however, I was a little reassured when I saw how (literally) straight forward the walk would be. It appeared that all I had to do was to follow the main road until I reached the next bus stop.

I did not walk far before being faced with an unexpected challenge. The pavement stopped and I had to choose between walking on the road or on a large parking lot parallel to the road (Figure 5). Because of the dark I could not see whether the parking lot transitioned back to a pavement at the end, and I was for that reason unsure whether it would be a dead end. Not being comfortable walking on a trafficked road in the dark, especially without anything reflective on me, I chose to walk across the parking lot. When I reached the end, I could see that the lot only led back to the road and not a pavement. At this point I was very confused. The map on the bus stop had clearly told me to follow the main road, but that did not feel safe to me. I had decided prior to the test that I was not going to use my phone as help, but as I could see no other option than putting myself in danger by walking on the road, I decided to use my phone to check if there was any other way to get to my destination.

Figure 5: A screenshot from Google Maps showing my approximate view. I could either walk on the parking lot on the left or the main road on the right.

Figure 4: The poster on the bus shed displays a map with directions for the nearest ticket machine.

Figure 6: The map on the bus shed displays a map with directions for the nearest ticket machine.

With the help of the satellite photo, I could see that the path went from the parking lot through a small but densely forested area before continuing along the main road (figure 6). Because the parking lot, especially the edges, were poorly lit I would never have spotted the path on my own (Figure 7).

Finding the bus stop from there was easy enough. Although the ticket machine was hidden beside the shed, it was easy to spot because of its very bright screen (Figure 8). At this point I began recording while thinking out loud. I filmed the test with my phone strapped to my chest to keep my hands free and to allow myself to keep my focus on the task.

The screen only displayed the AtB logo with no instructions for how to use the machine. There were no

Figure 7: The end of the parking lot which leads back to the main road. The arrow points to the hidden path.

Figure 8: Lohove bus stop. The arrow points to the ticket machine which is hidden behind the shed. (The photo was taken during day light hours a few months later.)

buttons. By instinct and process of elimination I clicked the screen to initiate the process, which turned out to be the correct course of action.

When I reached the payment page, I realised that it did not offer me to add more tickets to my "basket". When I tried to go back one step to see if that's where I could add more tickets, the back button (an x) brough me all the way back to start forcing me to start the process all over again.

The payment page also informed me that the machine neither provides change

nor accept notes. As I added coins to the machine the screen displayed the remaining amount. When the count reached 0, the ticket slot blinked and I could hear the machine doing something, but I could not see the ticket. I was confused about where my ticket was until I put my hand inside the slot and found it tucked underneath outside my field of vision.

Findings

My self-test allowed me to understand how the user experience of the ticket machine may be for someone having to use the ticket machine for the first time. I was also able to see how the ticket machine service works.

For later reference, the following pictures on the next page and descriptions display the different screens of the ticket machine display in the same user story I created for myself. I chose to draw the screens as glare made it difficult to photograph. I don't consider any important information to be lost looking at the drawings rather than a photograph.

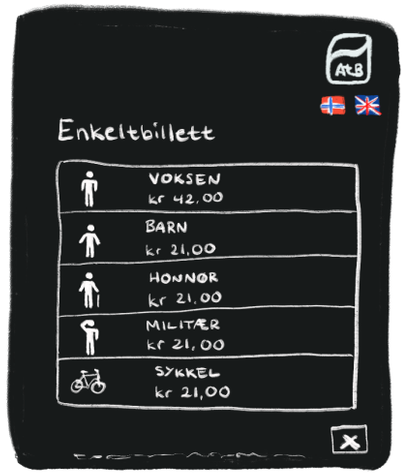
Figure 9: The ticket machine at Lohove bus stop. The photo was taken during day light hours a few months later.)

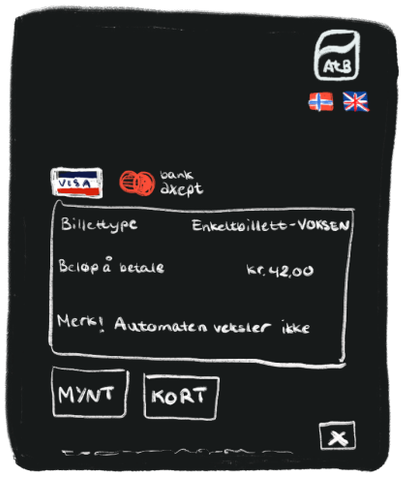
1 The default screen shows the AtB logo and nothing more.

2 After clicking the inital logo screen the interface tells you to click the screen again to buy a ticket. Language options also appear in the top right.

3 The user can now choose between either a single ticket or a 24 hour ticket. The X in the bottom right exits the process and takes the user back to screen 2. The small scribbles at the bottom represents contact details for customer service. The language options are available throughout the rest of the user journey.

4 When opting for a single ticket, the different ticket options are adult, child, senior, military, and bicycle. All tickets cost 21 NOK except adult tickets which cost 42 NOK.





5 The screen heading prompts the user to choose between the two available payment options: coin and card. Info box informs the user of the ticket type, the amount to be paid, and a warning that the machine does not provide change.

6 After opting for coin the machine prompts the user to insert coins into the machine, and repeats the change warning and the price to be paid.

7 The content of the info box changes when clicking it. It now informs the user that they can pay by card underneath, and what the amount to be paid is. This number decreases parallel to the value of coins inserted into the machine.

8 When the full amount has been paid, the ticket is printed. To indicate where the ticket will be printed from, a green light blinks right underneath the hatch marked with “BILLETTER / tickets”. There is a similar green light above the coin slot which blinks during the coin insertion process.

I found the ticket machine to be overall straight forward to use, and it worked the way I expected it to. My only annoyance with the GUI was the lack of a back button. Instead, there was the X in the bottom right which would reset the machine when clicked. Although I wouldn’t categorise it as a malfunction, the hidden ticket was a result of flaws in the physical design of the machine or the unpredictability of paper.

The biggest problems I experienced with the entire test were either in the category of finding the machine or missing functionalities. My **personal** criticisms after finishing the test are:

- The machine does not provide change. There is a decreasing chance that I am carrying cash at all, but it’s even less likely that I am carrying a combination of coins that add up to 42 NOK in total.
- The machine does not accept notes. As pointed out above, I am unlikely to carry cash, but if I am, it’s most likely in notes as coins are heavy and not likely to be useful for anything.
- I could not see the ticket. After testing other machines with users later, I now believe this to be a problem either specific to the machine I used or a somewhat rare occurrence that can happen with all machines.
- The machine does not offer the opportunity to purchase more than one ticket at a time.
- The map did not give provide satisfactory directions.

Regarding the ticket machine’s physical interface, I didn’t feel like the placement of the different elements made much sense. At the bottom you can find the card terminal in three parts (the PIN pad, contactless pay surface, and card slot) and a compartment for coins if the user were to cancel the payment process or if the payment failed. These are placed very low, forcing the user to crouch to see them properly. I theorise that this is done to accommodate wheelchair users, but in that case the coin slot is probably too high up.

As with many other mechanical machines, the answer probably lies with engineering and physical constraints. For example, it makes sense that the coin slot is higher up than the coin outtake as to take advantage of gravity.

To compensate for odd placement, the components are clearly marked with text and flashing lights are placed by each component to attract the user’s attention. I must admit that I did not notice the lights during the test even though they are very bright, but it did not hinder me in understanding which component I should be interacting with.

Conclusively, looking at the ticket machine as a closed system, it worked fine. Saying that, I am left with many complaints related to my user experience.

There are many open-ended questions as to why certain design decisions are the way they are, but that should be irrelevant if the user is able to achieve their goal without having to wonder what to do next or why something is not acting as expected.

This is what I will subsequently be testing with users.

Testing with users

Test design

I designed the usability test almost identically to the test I did with myself. I did not see it necessary to expand the test as most features were already covered. I also thought it interesting to compare the differences in results between myself and users.

Recruitment

I recruited test participants from my own pool of friends and acquaintances, as well as colleagues of one of my friends. As I did not have much to offer in terms of compensation, I approached most of the participants face to face, which made it easier to play on their charitable side and more uncomfortable for them to say no. I also recruited two participants by sending out a mass text, which prompted the participants to contact me voluntarily. To sweeten the deal, I made it clear that I was offering a free bus journey of their own choice and some chocolates which had a total value of 62 NOK.

The participants I was able to recruit were all in their 20s, had lived in Trondheim for several years and were highly educated. They were therefore probably not the most ideal group to test on. This is something I will be taking into consideration in my analysis and conclusion.

Scenario and test setup

I asked the participant to meet me at the bus stop of their choice and to take the bus to their desired destination. I gave them the following scenario:

"When you get to the bus stop you realise that you have forgotten your phone and wallet. Because of an appointment you don't have time to go home and retrieve them, but you find some cash in your pocket."

The participants were required to travel with a valid bus ticket, and they were not allowed to use their phone. I gave them some cash to emulate the money they were to find in their pocket in the scenario. The money consisted of one 100 NOK note, one 50 NOK note, two 20 NOK coins, one 5 NOK coin and one 1 NOK coin. The money added up to 196 NOK.

Consistency vs realism

To maximise the value and reliability of the test results I needed to balance realism and consistency between each test session.

If I were to maximise consistency every participant should travel between the same two bus stops at the same time of day. I chose to tip the scale a little bit more toward realism where the participants would travel a journey of their own choosing in terms of both route and time of day.

Adding more variables into the mix has its disadvantages. I would not be able to test for the same challenges in each test, and the tasks would be easier for some participants and harder for others. I would not be able to check whether the participants' would share my user experience at Dragvoll bus stop.

On the other hand, compromising the consistency of the tests gave me three advantages:

- It would provide an opportunity to confirm that certain usability problems are shared between several machines. This would prevent me from suggesting fixes that are irrelevant to all machines but one.
- I would probably be able to find more usability problems.
- Recruitment was made much easier as sacrifice on the participants' part was minimised.

Debrief

I joined the participants on the bus for the debrief. I wanted to do it as soon as possible, as that's when both of our memories are the freshest, and I would not need to book more of the participant's time.

I asked everyone the same questions:

- Have you used AtB's ticket machines before?
- What are your initial thoughts after finishing the test?
- Would you use one of AtB' ticket machines again?
- Did you experience any challenges?
- Was there anything you liked?

I also asked any follow-up questions when anything interesting came up and made sure to ask about specific occurrences and choices made during the test.

In the cases where the participant started at a bus stop with a ticket machine, I made sure to inquire about how they would solve the task otherwise. This was done at arrival, as these bus stops (fortunately) did not have ticket machines.

Results

7 tests were conducted, all within one week of each other. The table below shows the diversity of journeys made during each test. I was able to test 4 unique ticket machines at all different times of day. Two participants started the test at a bus stop with a ticket machine.

Malfunctions

In participant 1's case, the ticket machine's (at Høgskoleringen bus stop) screen was unresponsive, which forced us to cancel the test. I first theorised it was because of the dirty screen, but after similar problems with the machine at Østre Berg (where the screen was clean) the culprit was probably the cold (the tests were conducted in January). This affected three participants (1, 5 and 7).

All three quickly expressed a wish to give up and take the bus without a ticket. As it was important for the usability test that they complete the purchase, I needed to insist that they keep trying.

When they continued to have issues with partly or completely unrespon-

sive screens, I had to offer to help as their hands were colder than mine after touching the screen repeatedly without gloves. This worked with participant 5 and 7, though not perfectly, which strengthens my belief that it was a temperature issue. I did my best to not help them with the task, making sure that I only click what the participants wanted to click. It was certainly not ideal, but as the alternative was to terminate the test it was the preferred option.

Participant no	Bus stop	Nearest machine	Approx. time of day	Successful purchase
1	Gløshaugen	Høgskoleringen	Evening rush	No
2	Studentersamfundet 2	Studentersamfundet 2	Morning	Yes
3	Gløshaugen	Høgskoleringen	Evening	Yes
4	Solsiden	Solsiden	Morning rush	Yes
5	Moholt Studentby	Østre Berg	Morning rush	Yes
6	Nidarosdomen	Studentersamfundet 2	Noon	Yes
7	Moholt Studentby	Østre Berg	Morning	No

Table 1: Overview of routes, time of day and success for each respective test participant.

Participant 7 was able to complete the purchase but ended up buying a children's ticket. During our struggles with the unresponsive screen, they managed to do a "double click" which led to a children's ticket being selected without them noticing. The error was not discovered until the ticket was printed, and by then there were not enough coins left to purchase the correct ticket.

The machine at Høgskoleringen was functional again for test 3, which may or may not be because I reported it to customer service the day prior.

Identified usability problems

The machines malfunctioning is without a doubt a huge usability problem as it hindered nearly half of the participants in trying to purchase a bus ticket. Nonetheless, this is not something I will pursue as it's not a design issue, but rather a hardware issue.

Aside from that, all participants managed to purchase a ticket with ease. During the debrief two participants commented on screen 7 (see page 30), wanting the counter to display the amount of money inserted, not the remainder. All participants expressed annoyance when discovering the machine doesn't offer change, which was especially noticeable to them since I did not provide them with an exact amount in coins.

Participant 7 bought the wrong ticket even though screen 5 displays the ticket type. During the debrief they stated that they didn't notice it. They did however notice the surprisingly low price when they reached screen 6 but did not think to check why.

The lack of usability problems

After taking a step back from focusing on usability problems uncovered during the usability tests, I realised that the absence of usability problems was just as compelling.

Ignoring the troubles caused by malfunctioning machines, no participants struggled with the design of the machine – neither the GUI nor the physical interface.

When asked about it, no one reported experiencing the task as difficult. One participant even expressed delight at the user experience, saying it was much quicker than using the AtB app. It's important to note that they followed up with an explanation of their technical problems with the app.

The feedback I did receive mainly suggested that the machines aren't needed at all. Most participants said they would get on the bus without a valid ticket as the time and effort saved outweighs the risk of getting caught. None reported that using a ticket machine would be a suitable solution even in an emergency. It is not clear whether this is a universal attitude or specific among this certain demographic.

Finding the machines

Although the task I initially set for myself was to evaluate the usability of

the ticket machines as a walk-up-and-use system, my test scenario, which included the task of finding the machine, allowed me to observe the ticket machine within the larger system of bus stops, instead of only looking at the ticket machine as a closed system.

The most conspicuous usability problem to me looking at the entire test beginning to end, was the difficulty some users had finding the ticket machine. One user spent a few minutes roaming around the bus stop confused before finding the correct information on the poster. When they understood that they needed to go to a different bus stop they were very unsure what the exact spot the map was pointing to and thought for a while that they needed to go to the closest grocery shop.

Participant 4 had no trouble finding the ticket machine as it was placed on the same bus stop from which they wanted to travel. When we arrived at our destination, I instead asked them what they would do in the same scenario there. They gave up before finding the information on the poster.

Most participants were able to read the map correctly, but there was a general consensus that being familiar in the area helped a great deal. Both participant 2 and 6 were asked about the map at Nidarosdomen bus stop, and both used the river to orient themselves as it's a very distinct feature of the surrounding area. This illuminates the importance of landmarks.

Conclusion (and checking my biases)

All in all, I would conclude that the ticket machine is mostly unproblematic given the pool of participants I was able to recruit.

I was initially inclined to conclude that the posters and maps were the root of the primary difficulties during the usability tests. Thinking about it again at a later time, I realised that my interpretation of the results was coloured by my own bad user experience. All four participants who started at a bus stop without a machine were able to find their way fairly quickly. Two of the others figured it out easily as well. The last participant gave up, but I did not push this participant to continue the same way as with the other participants. Therefore, it's inconclusive whether they would have succeeded given a few more minutes as with participant 5.

The only "real" usability problem of great impact was the malfunctioning screens which I sadly cannot fix with design. However, the method of usability testing was able to uncover these problems.

There are many small changes that could be done in order to make the user experience slightly better, but the participants I tested on were not part of a group who are interested in the services the ticket machines provide. Therefore, I do not see any value in making changes to the design of the machine.

Reflections

The results of the usability test prove the effectiveness of the method as both the presence and absence of usability problems were uncovered. However, both the method itself and my application of it have faults and weaknesses.

Scenario

I did not specify in the scenario whether it is an emergency or not. If the scenario was real, either one or the other would be true. In an emergency, in this case meaning that the participant doesn't have time to go back home to retrieve their phone, there would probably be panic involved. I theorise that haste to that degree does not afford looking for alterative ticket payment options. The results suggest that users are not unlikely to get on the bus without a ticket even a when there is no emergency. In the opposite scenario, where they do have the time to go back home, they would likely do just that. This is something that I should have asked the participants during the defrief.

I do still believe there is a possibility that this scenario would play out in real life, but I struggle coming up with any that are likely to happen. This could mean that the scenario I created was inadequate, but it's an open question to which degree that matters. The scenario did cover a large percentage of the functionalities of the ticket machines, and for that reason I will consider it adequate. More importantly, I did set out to test the ticket machines, and not AtB's ticketing system as per the scope I set for myself. Therefore, it can be argued that it doesn't matter why the user is using the machine. The usability

of the machine itself does not change.

Recruitment

As mentioned, the group of participants I was able to recruit was not ideal. They had many characteristics which may have made the tasks easier even though none of them had used the machines before. I theorise that the task would have been harder for both elders, tourists, and people with impairments.

With the selection of participants in mind, I believe I was not able to uncover all, or even the most serious, usability problems. Either way, my belief does not matter as it's the **uncertainty** that shall dictate the need for additional usability testing with a more diverse and representative group of participants.

Why do they even exist?

Based on my own anecdotal evidence from having lived in Trondheim for the entire time the ticket machines have existed and from the usability tests with people very similar to myself, I believe that the ticket machines are used very little. This belief was compounded by finding a machine at a very busy bus stop that was malfunctioning over several days (at least).

Purchase, implementation, and maintenance of a single ticket machine is probably extremely pricey, which is apparent as only a small fraction of the bus stops have one.

Because all other payment methods are being promoted first, because the machines only offer single tickets, and because society is increasingly cash free, I believe the ticket machines are an expense that will not pay itself back over time.

I don't believe it's unlikely that the ticket machines in Trondheim will suffer a similar fate as the ticket machines in Oslo. Ruter, the Oslo equvalent to AtB in Trondeim, phased out their ticket machines in the Oslo area in 2021. With only 2.8 % of all ticket sales and 21 tickets sold a day on each ticket machine on average, it's clear that they have become obsolete (Garza, 2020).

Next steps

If the project were to continue, the next step would be to iterate back on one of following three steps in the HCD cycle: understanding the context of use, defining user requirements, and creaing design solutions.

Because I still know so little about the ticket machines, I believe it would be best to look back to understanding the context of use.

Redefining the context of use

After completing the usability tests of AtB's ticket machine I still feel uncertain of what an appropriate definition of the context of use could be. My group of test participants did not prove to be a group who are likely to use the machines, which means their user needs are different from the real user group.

To determine the context of use, it would need to get in contact with users who are likely to use the machine. They may teach me about why they use it, when they use it and how they use it. Knowing this would make it much easier to define user requirements.

Next round of testing

By subsequently testing the machines again with a user group who are likely to use the machines, I believe I would be able to uncover a range of usability problems that were not brought to light in the tests I conducted. Alternatively, I would find that the machine is not problematic with anybody. Testing the machine with a representative user group would also allow me to define clear user requirements which in turn would allow me to make sensible redesigns or design recommendations.

The question is how I would know what a representative user group is. This is not a question that usability testing easily can answer. It did not do so for me.

Next next steps

As mentioned on the previous page, I do not believe the ticket machines are here to stay. I would be wrong, but in the case that I am correct in thinking that they are close to becoming obsolete, that leaves the question of what to do with them when that day comes.

They would most likely be taken down, but I imagine it would be fun to explore what other services the machines can offer. Maybe they could be used to teach seniors about the AtB app, or maybe it could even be used as a game for kids?

Evaluating Flybussen.no

The second walk-up-and-use system I will be evaluating is Flybussen.no. Unlike AtB's ticket machines, which are physical structures that people literally walk up and use, Flybussen.no only exists digitally. In this chapter I will be evaluating the website, and hopefully see if its usability reflects its modern appearance. I will also be making a redesign based on my usability evaluation.

About

Flybussen.no is a website where the main functionality is purchasing bus tickets to and/or from a selection of airports in Norway. The website, which is what I will mainly refer to it as, is the result of a collaboration between four different bus companies in Norway, the so-called Airport Bus Alliance, operating from a total of nine airports.

It's brand new!

In an article on their news page they proudly announce the launch of their new website. The previous version was launched in 2017, making it likely that this newest version is less than a year old. But does that mean it's extra well designed?

Desktop vs mobile

The newest version of the website has been built for mobile first, which they claim to "*make it easier for you to buy a ticket or search up information on the go with just a few keystrokes*". Yet, I will mainly be focusing on the desktop version. Even though the website is mobile first, desktop second should not be neglected.

Defining the context of use

How I define the context of use will solely be based on my own assumptions and what I consider to be realistic and possible. I choose to define the context of use by answering the four points mentioned on page 9:

1. I have two stakeholders: Flybussen and everyone (see definition on page 5). When I evaluate and redesign the website, I will mainly be prioritising the users. Regarding Flybussen as a stakeholder, I will be honouring their graphic profile and their core service.
2. The user characteristics I will be considering are anything that my definition of everyone covers. I define the website to be a walk-up-and-use system, and I will thus keep in mind that users can be completely inexperienced with the service or similar services.

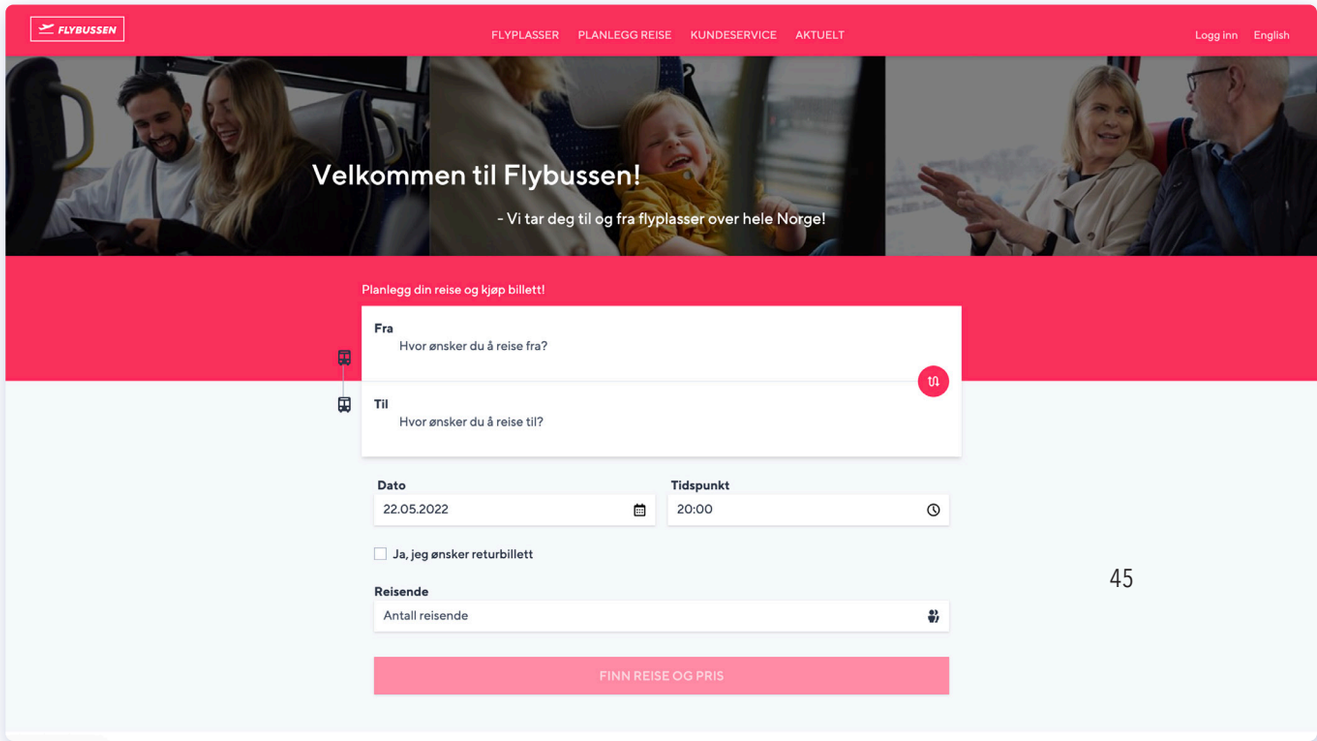


Figure 10: The home page of Flybussen.no.

3. I will initially assume the main user goals to include the following tasks: purchasing a ticket, finding live updates about their bus, and finding out whether the service applies to them (in terms of price, airport, bus stops etc.).
4. The user tasks can be performed anywhere and anytime. I believe it's most likely that the user either plans their journey and buys a ticket in advance at home, buys a ticket on their way to the bus, or checks the app for updates while waiting for the bus.

Usability inspection

Preparation

To prepare for the heuristic evaluation I found it important that I am well versed with the website and the heuristics myself. I mapped out the website as seen in the Figure 11 on the next spread which reveals that the website has more than 40 unique pages. Every airport listed on the airport page has its own news section, which I did not go through*. The most interesting feature found was a live map that enables the user to see where each bus stop and bus is in real time. This feature was a shocking 5 levels deep in the website.

* I later went back to investigate further, and discovered that the website is easily has a hundred unique URL. Most of them were under the many news sequions and were essensially duplicates.

As I did not have any backup participants in addition to never having done a heuristic evaluation before, I conducted a test run with a classmate. Aside from much needed practice, they also pointed out many usability problems that were missed during the "real" usability evaluations later on.

Recruitment

I recruited four evaluators per Nielsen and Mack (1994)'s recommendation of 4-5. My only criterium was that they could be considered "usability experts" which I did not specify further. All but one evaluator had a design related PhD and many years of experience in either academia, industry, or both, with the fourth evaluator soon to be finishing their PhD.

Setup

Three out of four evaluations were conducted digitally. My setup consisted of one screen where I kept notes and one screen where I could observe the evaluator's screen. As the evaluator identified usability problems, I wrote them down on post-its in Miro.

During the one physical session, I only brought my laptop for note taking in Miro and let the evaluator to conduct the evaluation on their own laptop. This way I could both see what the evaluator was doing and simultaneously have my own laptop freed up, making it a very similar setup as with the digital sessions, the only difference being not having a screen recording after the fact.

Evaluations

As a replacement for the training session, I made sure to carefully explain what the method is and what their task is. None of the evaluators had practiced the heuristic evaluation formally in the past, but they were all familiar with it. Everyone was also familiar with Jacob Nielsen's 10 general principles for interaction design (Nielsen, 1994) but did not have them memorized.

Beforehand, I had set up a strict schedule, I was not able to book the evaluators for as long as Nielsen and Mack (1994) suggest so I had to create a compressed version:

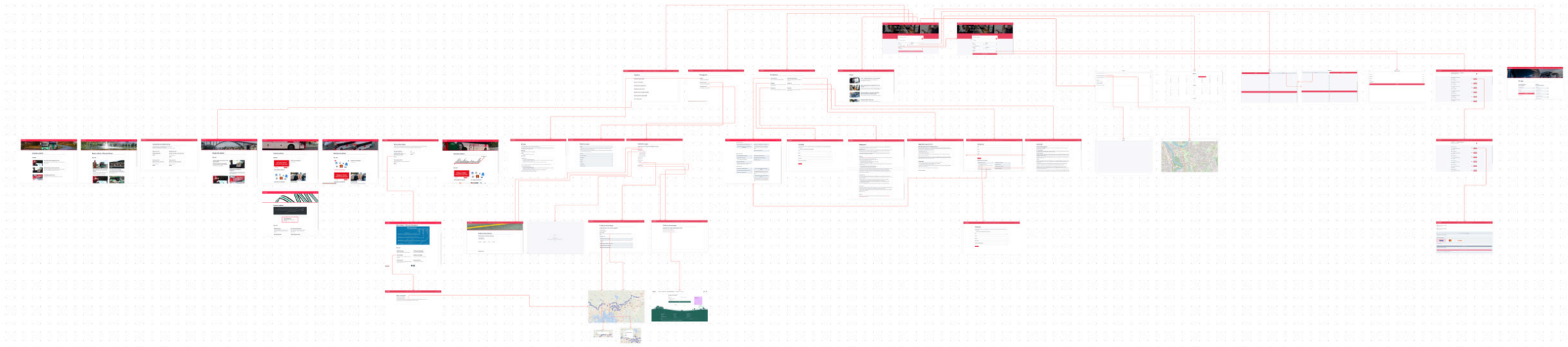


Figure 11: A mapping of Flybussen.no excluding the many news pages.

1. 5 minutes intro
2. 30 minutes of evaluation
3. 25 minutes of categorisation and debrief

Introduction

In the introduction I started by explained the purpose of the session. I continued by asking whether they had any knowledge or experience with the method, and subsequently stating their task. I sent them a link to the heurisitcs and made it clear that I would be writing down all the usability problems they find so that they could categorise them later, but they were allowed to categorise as they go if that's what they prefer. A couple of evaluators started the test by categorising usability problems as they found them, but quickly forgot to do so and focused solely on the finding instead.

Evaluation

My initial plan for the evaluation was to allow the evaluators to roam around the site as they wished but I had a few backup scenarios they could use if necessary. Ideally, I would encourage the evaluator to go over the website twice; once to get a general overview, and once more to investigate the details, which was the recommended procedure in Nielsen and Mack (1994). There was unfortunately not enough time for this, and I instead had to urge them to focus on the main functionalities.

My backup scenarios came in handy immediately during the first evaluation. The evaluator expressed signs of confusion and uncertainty of what to enter the text input fields for picking bus stops, so I suggested a travel route. To keep each session consistent, I provided the same route to everyone.

I made sure to pick a route where the start and end points were places, not bus stops, the route being Oslo Airport to Ås. Since three of the evaluators were not Norwegian and thus did not have a Norwegian keyboard, they were immediately stuck as the website did not offer any way to get around it. I then switched Ås out with Oslo Central Station, where the nearest bus stop is Oslo Bussterminal. This is not something a user is likely to know unless they are familiar in the area.

After the evaluators had finished the ticket purchasing dialogue, I found that it was necessary to provide a new scenario for them to go through. The scenario was as follows:

"The user is waiting at the bus stop, the bus is late and they want to figure out what's going on."

I believed this to be a realistic use case as public transportation is frequently delayed causing users wondering about what's happening*. Through this scenario I wanted to guide the evaluators towards the traffic announcements and live map. The map is very difficult to find, which would lead the evaluators to cover a much larger portion of the website.

Debrief

The debrief mainly consisted of the categorisation of the identified usability problems. I screenshared the Miro board and had the evaluator tell me which heuristic they felt each usability problem fell under. With evaluators who had identified a considerable amount of usability problems, I cut the sorting short as I saw little value in the exercise and was more interested in their general thoughts.

I ended each debrief by asking the evaluators about possible design solutions. I hoped to receive big creative ideas that would affect the overarching structure of the website's interface or service, but they mostly gave me simple and obvious solutions to individual usability problems they had found. I would get suggestions such as "*Increase the colour contrast*" to problems like "*The colour contrast is too low*".

This was probably a consequence of the evaluators not having enough time to get to know the website and service in its entirety. They only had the time to test and evaluate it superficially, and naturally their design solutions would reflect that.

* In subsequent user test I found that users are not likely to be looking for such functionality on the website.

Identified usability problems

Complete list

My documentation of the identified usability problems was, as previously mentioned, done live in Miro while the evaluators stated each problem as it was found. The evaluators identified 31, 33, 25 and 20 usability problems respectively. This number is an approximation as it may be skewed by my own note taking. I did not find the heuristics to be useful for my continued work and decided to disregard them. See Appendix 1 for screenshots of the Miro boards with all documented usability problems.

Refined list

I went over the complete list of identified usability problems from each evaluator and deleted duplicates. Some were rewritten to make them easier to understand and more appropriate for the subsequent severity survey. The list was created in parallel with the survey. I added screenshots by most of the list items to minimise the risk of misunderstandings. See appendix X for the list with pictures as listed in the survey.

In no particular order, the refined list of usability problems is as follows:

1. There is no way to search for places or bus stops containing the letters Æ, Ø or Å.
2. Expected a dropdown list or active input field when clicking the first input field instead of being sent to a separate page.
3. The suggested return time is sometimes earlier than the selected departure time.
4. It's unclear what the grey text means, making it look like there are many Oslo Airports or many stops near Oslo Airport (in the example below)
5. Searching for "Oslo Central Station" gives "Oslo Airport" (incorrect) as the first suggestion and "Oslo bussterminal" (correct) as the second suggestion.
6. There are too many separate screens, causing the user to lose context.
7. The user has to select travel times twice: First on the front page, then once more after clicking the "Find trip and price" button.
8. On the "Airports -> Oslo Airport" page the user has to know and remember which bus company they have booked with.
9. Norwegian text on English page ("Gå til side").
10. The website is not giving a clue as to where to find information about why the bus is late.
11. Bus stops are not searchable on the real-time map.
12. Buses are not searchable on the real-time map.
13. The real-time map is too small for a phone.
14. The image below looks like a real-time map but is only a static image.
15. There is no predicted time of arrival for the buses on the real-time map
16. Info on the real-time map has not been translated to English.
17. It's unclear how the blue and red pins are different unless the user has read the instructions before entering the map.
18. The naming of the "Plan your trip" page communicates poorly what the page actually is.
19. There is no info on which currency the price is in.
20. It's difficult to identify the language setting on small screens.
21. The edit (pencil) button changes the language back to Norwegian.
22. Payment page is not available in English
23. Some links are not available in both languages. (Only one of the links in the image below exists on the English page)
24. Input information is deleted when switching between languages.
25. The hamburger menu (on small screens) shows no info on how to buy a ticket.
26. Some pages are just a wall of text.
27. The website is missing a quick way to find live updates.
28. The info boxes have too much whitespace, which may be communicating that information has failed to load.
29. The list of traffic announcements is blank, which may be communicating that information has failed to load.

30. There is no direct way for the user to say that they haven't found what they are looking for in the Q&A.
31. Phone number is too hidden.
32. The "Find nearby stops" feature only shows a blank page.
33. The "Find stops in map" feature shows a blank page if location sharing is turned off.
34. When the screen is blank, it is completely up to the user to troubleshoot the problem.
35. Unclear if travel time on the home page is selected for arrival or departure.
36. The "Find trip and price" button is hidden on smaller screens until the user scrolls down.
37. The edit feature is tedious if all the user wants is to see earlier departures.
38. It's unclear from the search suggestions that "Oslo Airport" is the same airport as "Gardermoen"
39. Departure/arrival time can be confused with when the plane lands
40. It's unclear to the user what FB5 means.
41. Absolutely nothing happens when the user searches for places that don't exist in the database.
42. It's not possible to delete text inputs.
43. The contrast is too low on the "Find trip and prices" button.
44. Lack of feedback when a departure is selected.
45. If the user wants a return ticket the layout on the two departure selection pages are almost identical.
46. Many of the input fields, buttons etc are too wide resulting in a lot of scrolling that would have been avoided with better use of space.
47. The text is too far apart making the prices to the right hard to spot and hard to see which text is on the same line.
48. The "Place order" button is hidden at the bottom of the page outside the initial viewport.
49. Bad error message when the user enters an invalid email at the checkout page.
50. The different sections on the "Customer service" page are not ordered according to importance.
51. There are news on the "Airports" page under most airports even though there is a separate news page.
52. There is no available info on when the user is charged after clicking "Place order."
53. Info about tickets being valid all day is too difficult to spot.
54. The prices at checkout are only available in Norwegian.
55. Info icon is not functional.
56. It's not clear how the search suggestions are sorted, making it unclear which bus stop is closest in distance.
57. The website does not remember the user selecting "Arrivals" instead of "Departures," and will then only show bus journeys departing AFTER the user wants to arrive.
58. The browser back button does not take the user to the previous page from any of the input type of pages.
59. The "find stops in map" features are not always available.
60. There is no way to check prices without going all the way to the checkout page.
61. It's not possible to choose travellers before choosing places to travel to and from.

Severity survey

Following the method of heuristic evaluation is described in Nielsen and Mack (1994), I created a survey containing all usability problems identified during the evaluation sessions (see appendix 2), adding up to 61 in total. I also added a few usability problems I had found myself and from the test run with my classmate.

The survey consisted of one question per usability problem asking the participant to rate the severity of the usability problem. I used the exact word by

word rating system as described in Nielsen and Mack (1994):

- "*I don't agree that this is a usability problem at all*
- *Cosmetic problem only – need not be fixed unless extra time is available on project*
- *Minor usability problem – fixing this should be given low priority*
- *Major usability problem – important to fix, so should be given high priority*"

The Nielsen and Mack (1994) also included a fifth, more serious rating: "*Usability catastrophe – imperative to fix this before product can be released*". I opted to not include this rating as the product is already released.

I added an open question at the end of the survey allowing the participants to leave feedback on anything they desired.

The survey was created in Google Forms and the answers were anonymous.

Analysis

The survey results were meant to help me 1) prioritise design decisions, and 2) design a usability test. A direct and simple way to create such a prioritisation list is to plot all problems against the perceived time they would require to fix or redesign.

Severity rating

First, I translated the four severity levels into numbers between 0 and 3, with 3 being the most serious. With a quantitative score, I could easily calculate an average score for each respective usability problem and sort them accordingly.

Time estimation

All that was left to do to create my plot was to estimate the length of time I would need to redesign each respective usability problem. This was a highly unnatural exercise for me to do. Most problems were very small and fairly quick to fix, so it made the most sense to enter my time estimates as minutes. Looking back, it would have made more sense to range them on a simple difficulty scale with a similar number of steps as the severity rating. It would have been an even better idea to follow a formal prioritisation method such as a MoSCoW Analysys or Imact-Effort matrix (Gibbons, 2021). Both separate the usability problems into four categories, which is much easier to tackle than what I did and ended up doing.

The time estimate added up to approximately 25 hours. Because I believed this was already a very conservative time estimate, I wanted to filter out the least important problems. Simultaneously, drawing a hard line of where to stop the redesign could be unwise as I both could not trust my time estimates completely and I did not know how much time I had to spare for the redesign. I did so anyway.

Plotting the data

I was not satisfied with the digital graphing tools as I wanted to visualise the many overlaps. Therefore, I opted to draw the plot manually with pen and paper.

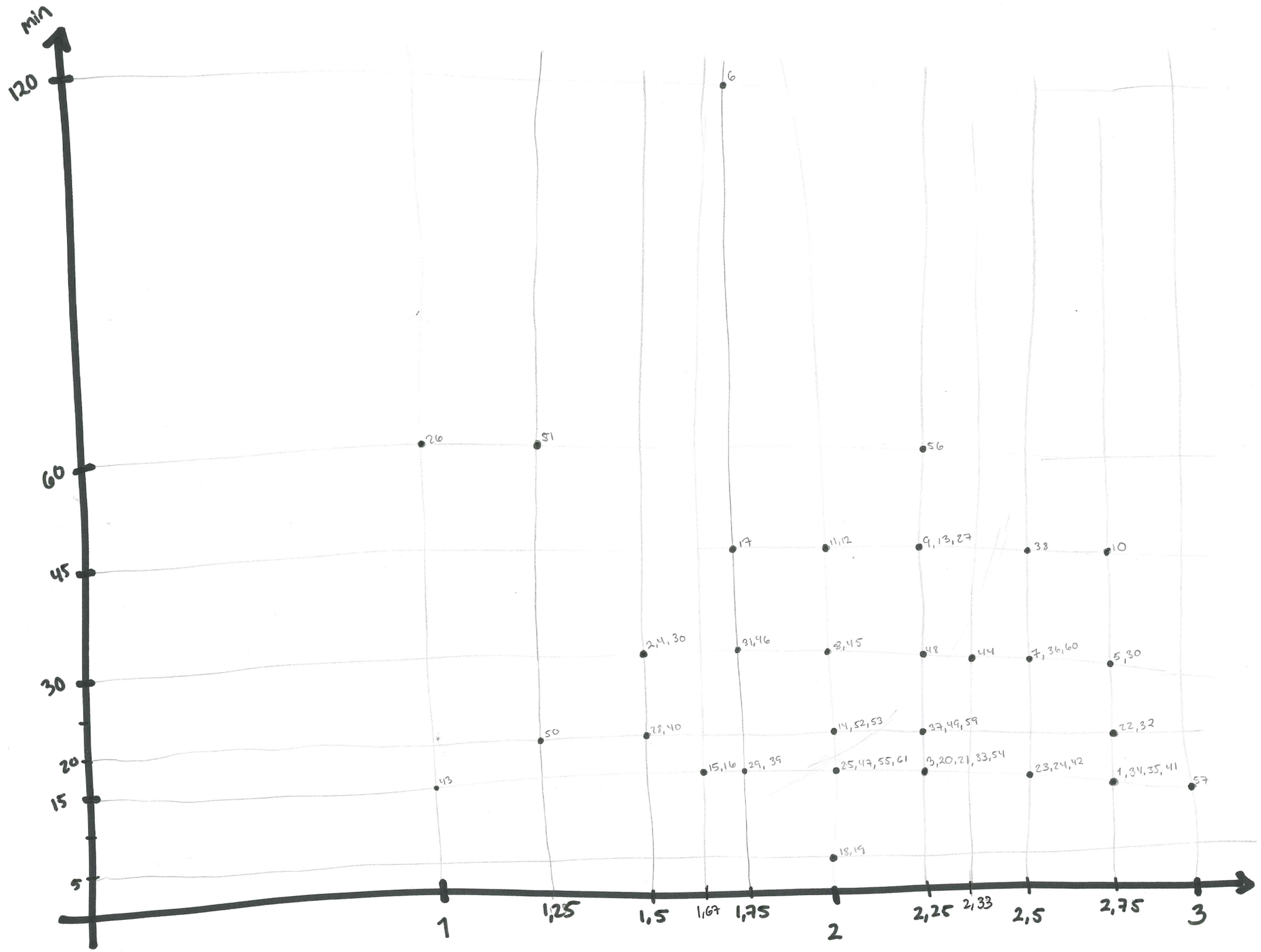
First, I drew the axis and a simple grid, and entered all the data points, making sure to write the ID numbers of each problem next to each point. That way I could see where all individual problems were placed on the plot. But since the *number of problems* were more interesting to know than *what* the problems are, I chose to redraw the plot, only this time I marked each point with a number indicating the number of overlapping points. Tthe plot is not linear, which meant I still had to make a judgement call for when I wanted to prioritise severity or when I wanted to prioritise time.

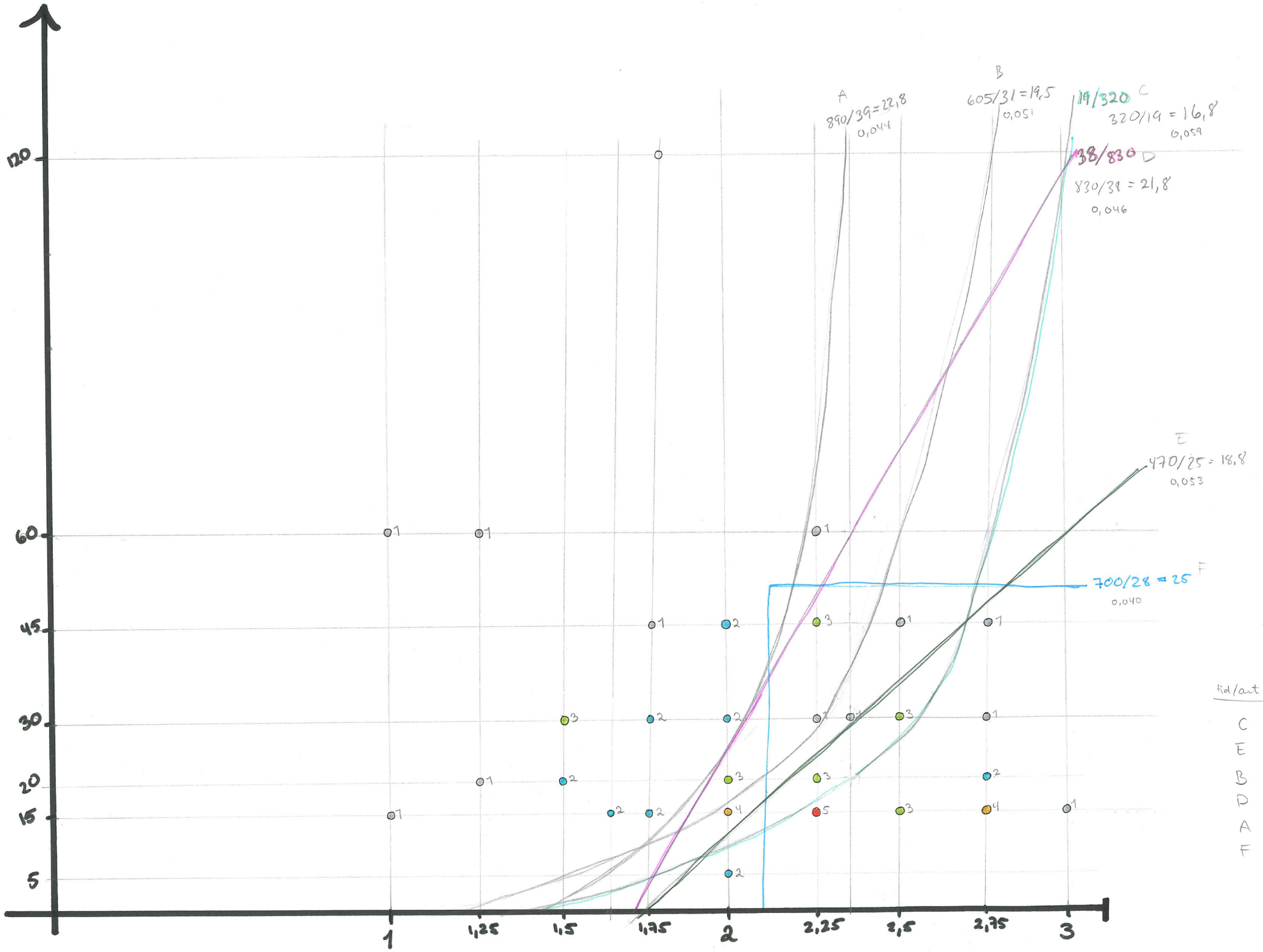
I drew a few lines and curves trying to filter out some of the problems. I summed the time estimates together and tried to see if I divided the number of problems by the total amount of time, I could pick one group over the others based on the number of problems I could do on the shortest amount of time possible. I decided that this was not an ideal method because my time estimates were still not to be trusted.

When I landed on where I wanted to draw the line, I still needed to make some decisions on how to prioritise within this selction. I could not come up with a good solution to this, so I went with my gut feeling. Sometimes time would trump severity, and other times the other way around. See the list in Appendix 3.

Figure 12: Seen on the next page is the first plot I drew by hand. The x-axis represents severity, while the y-axis represents my time estimate. The numbers represent each usability problem, and uses the same numbering as the list on page x.

Figure 13: The second plot can be seen on page x. This time the numbers on the plot represents the amount of overlapping usability problems. The notes on the side are my attempt at making sense of the plot and choosing where to set the limit for which problems to fix.


min
120
60
45
30
20
15
5
1
1,25
1,5
1,67
1,75
2
2,25
2,33
2,5
2,75
3
6
26
51
56
17
11,12
9,13,27
38
10
2,4,30
31,46
8,45
48
44
7,36,60
5,30
50
28,40
14,52,53
37,49,59
22,32
43
15,16
29, 39
25,47,55,61
3,20,21,33,54
23,24,42
1,34,35,41
57
18,19



tid/ant	ant/tid
C	F
E	A
B	D
D	B
A	E
F	C

19 =

Redesign

When fixing the usability problems found in the heuristic evaluation, I used the prioritised list made prior (Appendix 3). I simply started from the top and worked my way down. Assuming that the list is prioritised correctly, I was sure to always be working on the usability problem with the highest payoff in terms of severity and time spent. In this chapter I will be going through each individual change I did in the design of the website.

In this exercise I chose to attack the problems from a design perspective, meaning that I did not try and research technical solutions which would have been a developer's job.

Timing myself

As a small side-experiment, I timed myself redesigning the respective problems as I worked myself down the prioritised list. In retrospect, this was a bad idea as I suspect my preoccupation with the time affected my designs. However, I was able get an approximate measure of how well my time estimates were, with which I could conclude the quality of my prioritised list.

Skipping a few steps

The website is already designed and launched. Therefore, I saw no reason to resketch the whole website by hand. With a bigger overarching redesign this may have been the most sensible thing to do.

Instead, I used screenshots from the website and manipulated them with simple tools available in Figma. I made it a point to reuse as much as possible in terms of colours, styles and other design system components as the styling and brand identity were not identified as usability problems.

The redesigns

Out of 61 usability problems, I set out to fix the 39 most top rates ones. After filtering out what I considered to be non-design related problems, the actual number of fixes I did was 34. Additionally, some of the fixes I did also solved some lower priority usability problems which brought it down further to 26.

The numbering of each respective usability fix indicates [priority ranking]/[ID][estimated time]. For example, "Problem 4/13 (20)" would mean that the problem was number four on the priority list, number 13 on the refined list of usability problems, and that I estimated the redesign to take me 20 minutes to complete. Please refer back to the refined list on page 50 for an overview of which usability problems the following design solutions aim to fix.

Problem 1/57 (15)

I perceive this to be a problem entirely in the hands of the developers, unless I want to redesign the whole website to not need this functionality.

The most highly rated usability problem is also one of the easier ones to fix. I perceive this problem to be a small slip up in the code. Therefore, I believe a redesign would not be worth it as a developer could most likely reprogram it much quicker.

Problem 2/1 (15)

I do believe this problem could be fixed through building a cleverer search algorithm. An example of a solution could be to count e. g. AA as Å. Entur is an example of a related service (journey planning) with this feature.

I believe a developer would be a good help in with this problem as well, since I don't have the competence to know how viable different solutions are. Assuming that changes to the search is unviable, I added a keyboard button inside the input field. A small keyboard with letters not found on an English

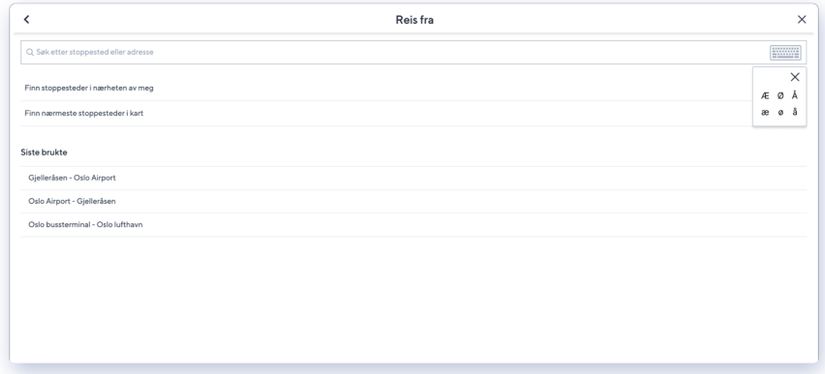


Figure 14: The fix for problem 2/1.

keyboard appears when clicking the icon. By clicking a letter, the same letter will appear in the text input field as if the user typed it with their keyboard.

Problem 3/35 (15)

In the case of an error, the user should be presented with a descriptive error message. The error message should make it clear to the user what the problem is and how they should approach it. In this particular case, I don't know what the problem is, and I cannot design a fitting error message.

There may occur errors that are not considered by the development team. Therefore, there should be a general error message for these cases. In this message I made sure to both inform the user that an error has occurred, and what steps they can take to fix the problem.

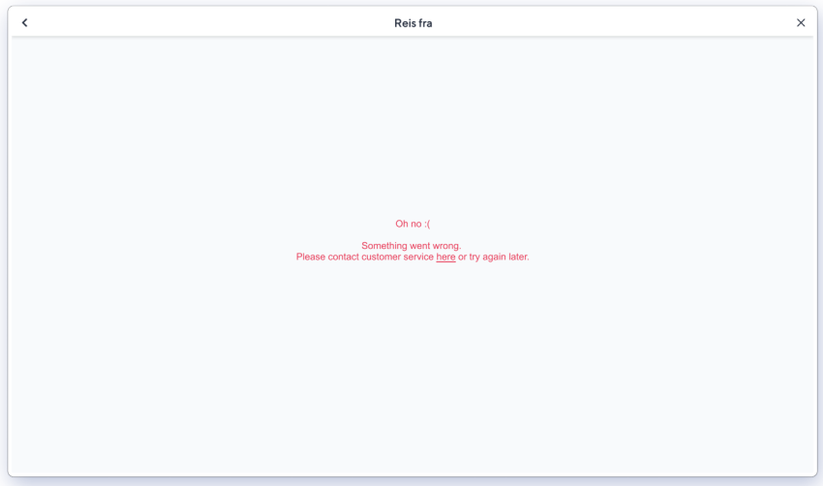


Figure 15: The fix for problem 3/35.

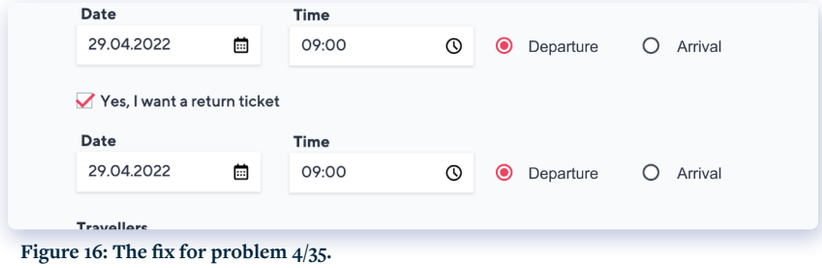


Figure 16: The fix for problem 4/35.

Problem 4/35 (15)

An important point in user friendliness is to keep information in the world rather than in the head (Norman & Norman, 2013). In this instance we are talking about whether the user has selected departure time or arrival time. To me, the obvious answer was to bring this information to the main page instead of hiding it in a menu. Additionally, by using a common interaction pattern it's more likely that the user understands the concept.

Problem 5/41 (15)

By adding a short message telling the user that their search word yielded no results, there is no confusion about why no results are appearing.

Problem 6/22 (20)

The payment page is external and standardised. Therefore, I have no power as a designer to fix it. However, it's not unthinkable that the payment service provider has an English version available. That leaves a couple of options: 1) There is a setting that the flybussen developers have control over, 2) the translated page has not been requested from the provider. Either way, it cannot be fixed with a simple UI tweak.

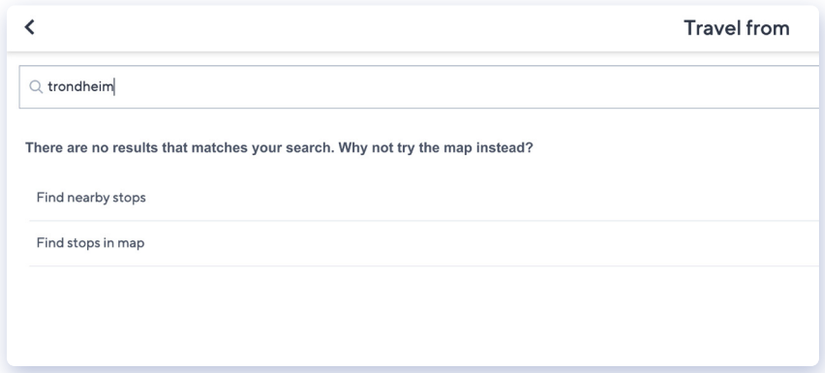


Figure 17: The fix for problem 5/41.

Problem 7/32 (20)

This problem overlaps with problem 3/34 where the problem was that the user had no way of knowing how to handle the empty page. And as I already mentioned when fixing that problem, I believe it's most likely a technical bug and therefore not something I can design around except for good error

handling which has already been solved in 3/34.

Problem 8/30 (30)

With a shortcut to the customer service contact page from the Q and A page, there is no question where the user should direct their questions if they can't find satisfactory answers on the Q&A page.

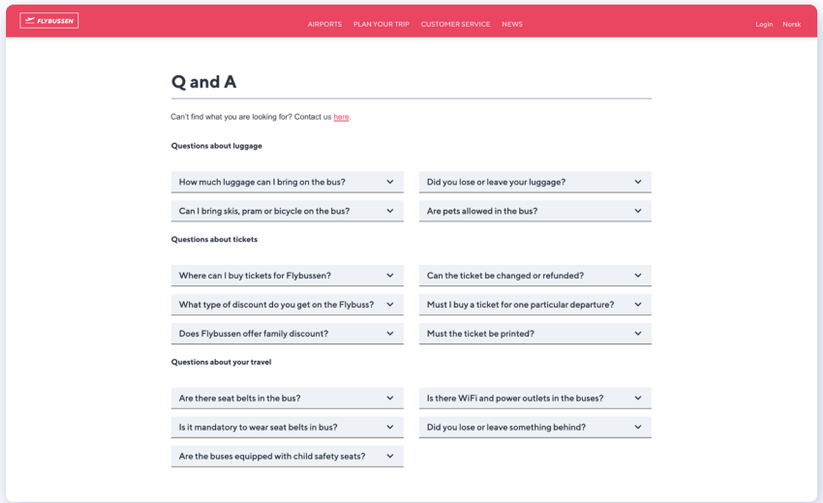


Figure 18: The fix for problem 8/30.

Problem 9/5 (30)

This problem is more complicated than it looks at first glance. Similarly to problem 2/1, it's an issue with the search, only in this case it results in hard to interpret search suggestions. To look at it from a design perspective, I think it's possible to use the same search and data and display them in a manner that helps the user pick the correct bus stop. A map could be beneficial for this purpose.

In this case I used the same live map which shows all bus stops. The bus stops can be clicked to display its name. I also imagine this info box to pop up when the user hovers the search suggestion. This would however, not work on mobile. I am very unsure how this feature will work and will for that reason make sure to pay extra attention to how participants respond to it during subsequent usability tests.

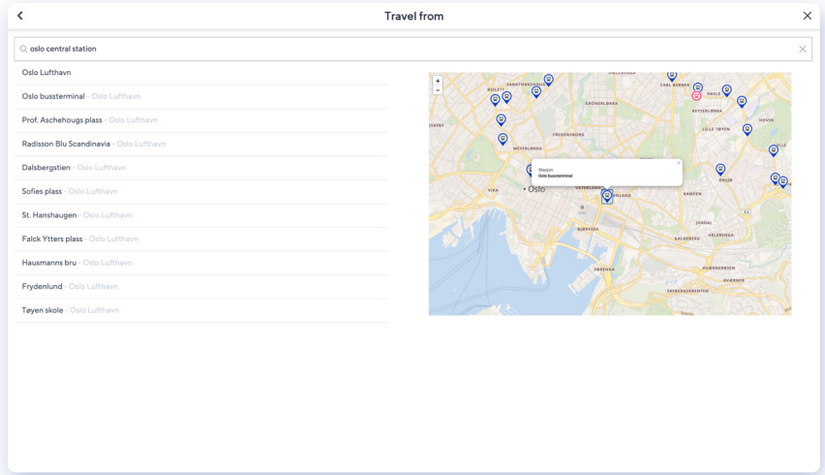


Figure 19: The fix for problem 9/5.

Problem 10/10 (45)

The problem states that the website does not give any info on why the bus is late. I argue that it's more important that the user can find out if they should give up on and find alternate transportation. The bus could be right around the corner, but it can also be blocked by a car accident far away.

The website already has the live map, which allows the user to see how far away the bus is, and traffic deviations, that allows the user to see reported delays. This information is very well hidden. Therefore, I collected all of it onto one single page and added a direct link in the header. This reduces the number of clicks to access the map from 4 or 5 to 1.

Problem 11/23 (15)

This is yet another problem that I believe should be fixed in code by adding the necessary hyperlinks. Instead, the fix of problem 9/10 could cover this problem as well. By including all updates from all bus companies/airports on the same page there would be no need to link to each one. Additionally, traffic announcements are very rarely posted so it makes sense to group them all together.

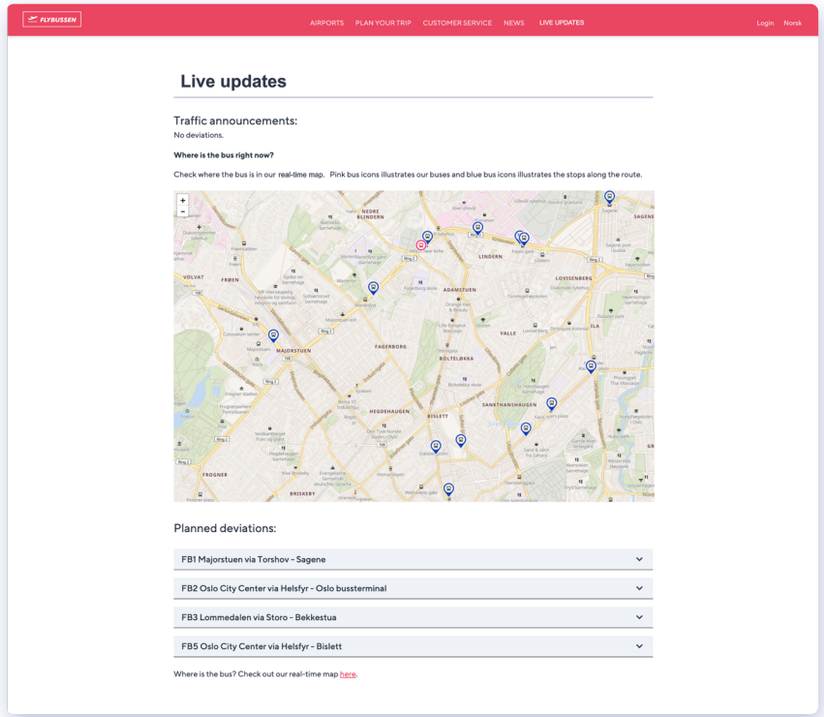


Figure 20: The fix for problem 10/10.

Problem 12/24 (15)

This problem can only be fixed in code.

Problem 13/7 (30)

From an outside perspective there is no obvious reason why "select departure" pages exist at all if a ticket is valid the entire day anyway. A believable hypothesis could be that the bus company wants an estimate of how many passengers they will have at any given time. With this data they could schedule their buses more economically. From the user's point of view, however, it's completely unnecessary. It can also add a lot of stress if the user doesn't see the fine print that says the ticket is valid all day regardless of the departure they select. Therefore, I will simply remove this stage in the user journey.

Problem 14/36 (30)

After testing it thoroughly myself on both mobile and a narrow desktop window, I was not able to reproduce this problem. It is not a case of having a

different phone than the evaluators as all evaluations were conducted on PCs. I therefore judge it to not be a problem.

Problem 15/60 (30)

Having already removed up to two steps in the user journey (of purchasing a bus ticket) this usability problem had already been softened. By looking at other ticket purchasing services I was able to see a similar theme regarding hidden prices. Therefore, it's not unthinkable that it's a business or marketing decision and not a design decision as the customer may be more likely to complete their purchase if they're already on the last of several steps to get there.

With the user in mind, I designed the front page to include the total cost. I debated displaying the price of each individual ticket type on the main page but decided against it to not overload the page with information.

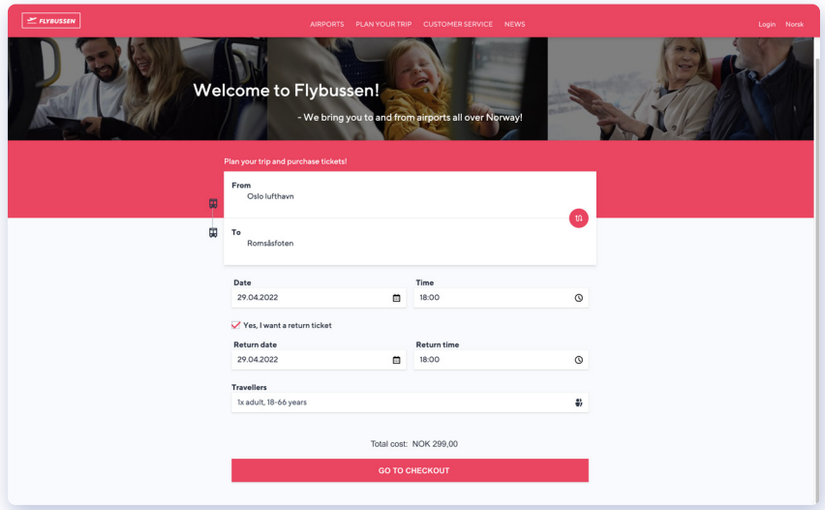


Figure 21: The fix for problem 15/60.

Problem 16/38 (45)

This is a problem that is essensially the same as problem 8/5. The difference is Oslo Airport stop is special as it's always either the one you start at or end up at, so it's always a 50 % change that the user wants to use this bus stop.

When the user clicks the pin or hovers over Gardermoen Parkering the

map will display a path between this bus stop and the Oslo Airport bus stop warning the user from picking a bus stop far away from where they want to be.

After finishing this redesign I came up with an even better idea. As already mentioned, there is always a 50 % chance that the user wants to select Oslo Airport, so it would make sense to always display Oslo Airport as a search suggestion – even before the user start typing anything in the search field*. The map may also show the airport as a default if location sharing is turned off.

* I later realised that I had forgotten the fact that Oslo Airport is far from the only airport Flybussen operates from. A modified version of this design solution could therefore be an accordion menu with a list of all airports.

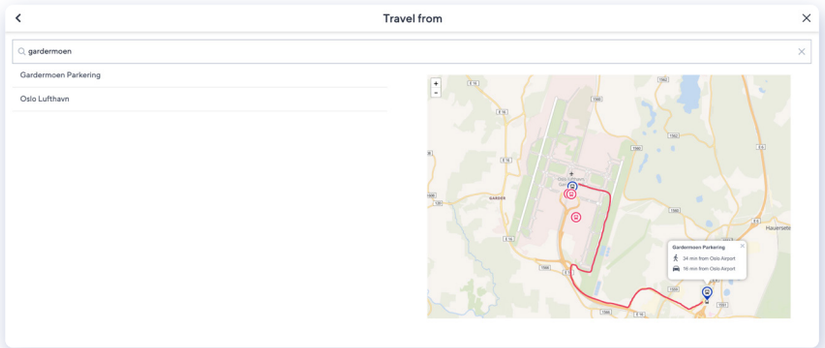


Figure 22: The fix for problem 16/38 (a).

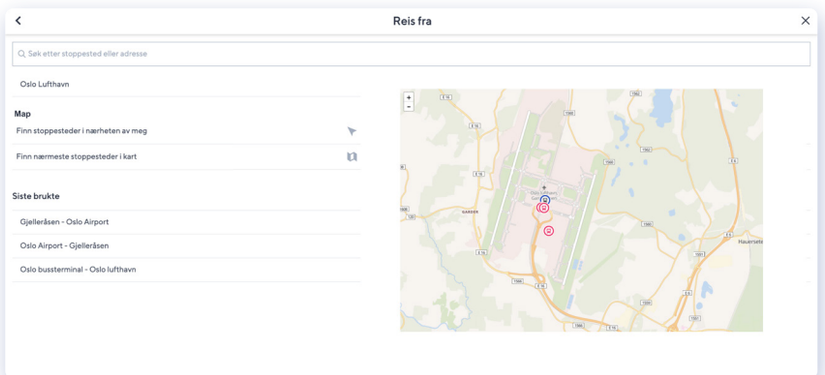


Figure 23: The fix for problem 16/38 (b).

Figure 24: The fix for problem 19/20.

Problem 17/44 (30)

This usability problem was solved via problem 13/7 when the "select departure" step was removed.

Problem 18/3 (15)

When checking the website again at a later point it seems as if it has already been fixed.

Problem 19/20 (15)

By adding a map next to the language setting, it's both much easier to spot and much easier to understand what the setting is.

Problem 20/21 (25)

As the whole page has already been removed, this problem is no longer an issue. It would either way need to be fixed through code.

Problem 21/33 (15)

I believe there needs to be a default map position. Alternatively, it could be enough to show an error message. As mentioned in the redesign of problem 16/38 Oslo Airport can be used as a default map position.

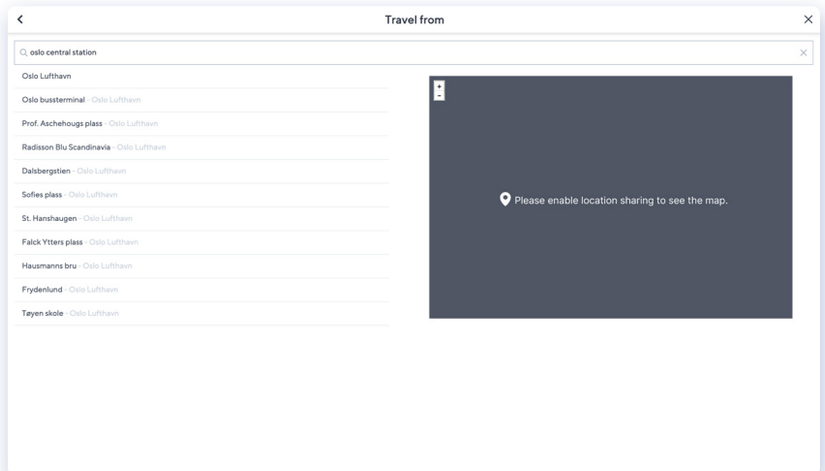


Figure 25: The fix for problem 21/33.

Problem 22/54 (15)

Translation issue only. No redesign needed.

Problem 23/18 (5)

To make the menu title more descriptive, I changed it from “Plan your trip” to “travel information”.

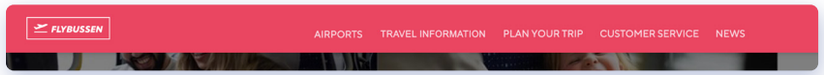


Figure 26: The fix for problem 23/18.

Problem 24/19 (5)

The problem is no longer relevant because the “select departure” page has been removed.

Problem 25/37 (20)

The problem is no longer relevant because the “select departure” page has been removed.

Problem 26/49 (20)

After testing it myself I realised the error message described the usability problem only occurs when nothing has been entered to the input field at all. When the user enters an invalid email address however, that’s when the strange error message appears. The aforementioned use case is handled by automatically scrolling up to the input field, highlighting it with a red stroke and adding an error message. The error message says “payment.emailRequired,” which isn’t so great either but it’s very simple to write it better. That’s what I did.

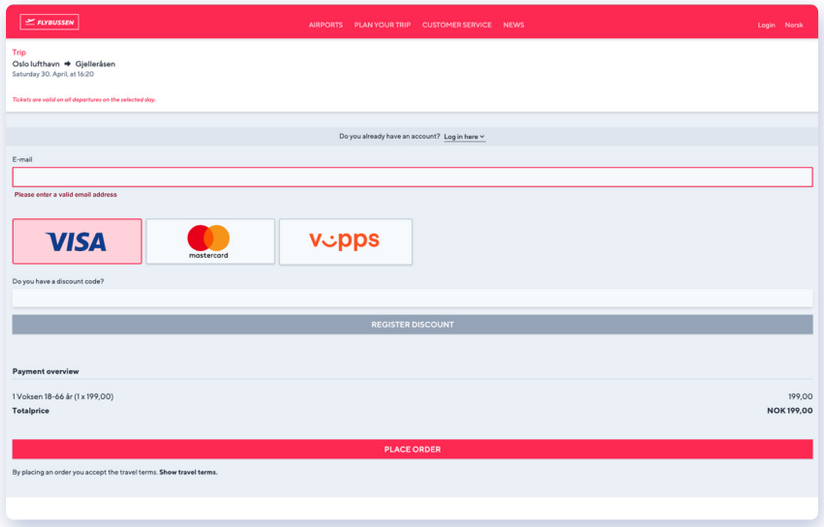


Figure 27: The fix for problem 26/49.

Problem 27/59 (20)

When an airport is entered into one of the input fields on the home page the next input box is selected automatically. This is when the map options aren’t available. I have no idea why it would be like that, but to fix the usability problem it is sufficient to simply always provide the map options.

Problem 28/48 (30)

I was able to move elements around to not take up so much space and without making it look crammed or cluttered.

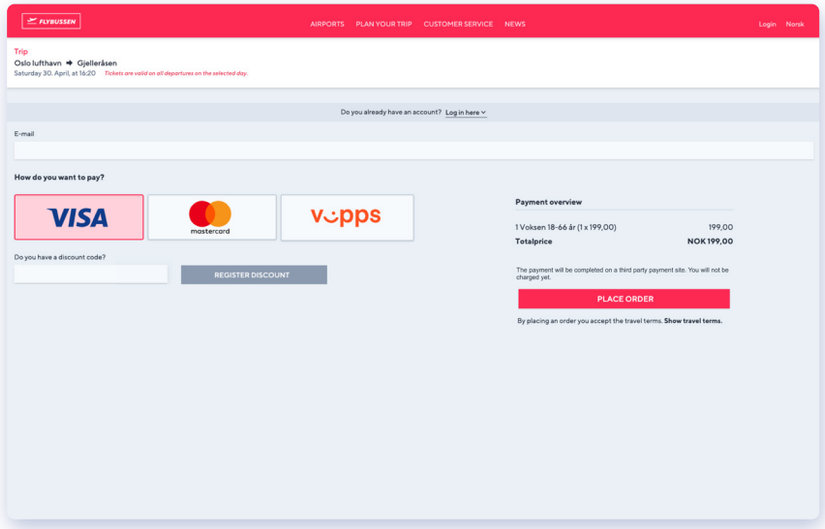


Figure 28: The fix for problem 28/48.

Problem 29/9 (45)

Translation issue only. No redesign needed.

Problem 30/13 (45)

Comparing the map on mobile to other map services, such as Google Maps, I cannot observe that the Flybussen map is any smaller. The user is able zoom in on the map, and the icons are no smaller than on other map services. Therefore, I see no usability problems here to be fixed. I theorise that this usability problem was a consequence of conducting the evaluation on a much bigger desktop screen.

Problem 31/27 (45)

This problem was solved in the through the design fix of problem 10/23.

Problem 32/25 (15)

The problem is easily fixed by adding a “Buy tickets” link in the hamburger menu.

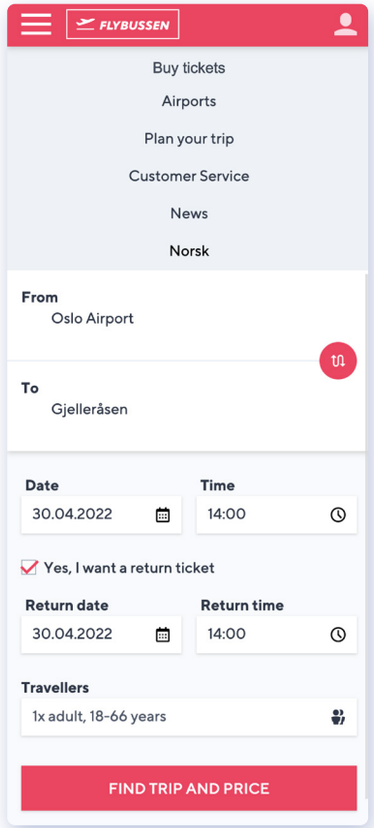


Figure 29: The fix for problem 32/25.

Problem 33/47 (15)

This problem was solved in the the design fix of problem 28/48.

Problem 34/55 (15)

It appeared as if the icons were not working (at least in Google Chrome), but they were in reality very slow. This is something that can be reprogrammed.

Problem 35/61 (15)

Can easily be reprogrammed, and does not require a redesign.

Problem 37/56 (60)

The stated problem implies that the search suggestions are certain to have something to do with physical distance to the entered keyword. This is not guaranteed. In one way it doesn’t matter as a ticket is valid from all bus stops. On the other hand, the search function is the only way for the user to find out which bus stops exist. The exception is the live map but it is not placed anywhere on the website that is natural for the user to look if they are searching for this information. This can result in a user missing their bus if they are looking for a bus stop that doesn’t exist where they are.

Ideally, the search results would suggest the bus stop that is closest to the place the user searches for, which is what I would recommend. Entur is an example of a similar service that provides this feature. However, I have already decided to design around what I know the search can do as it is.

One thing I do know is that the search results are not sorted based on which stop is closer to the place you want to leave from. This is the only information I can provide in a redesign.

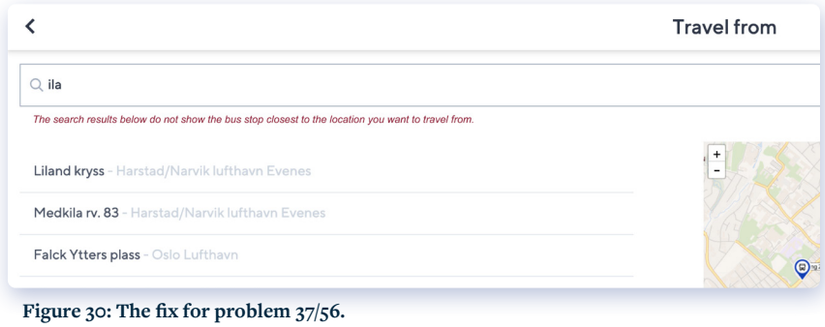


Figure 30: The fix for problem 37/56.

Problem 38/14 (20)

This problem was solved in the through the design fix of problem 9/10.

Problem 39/52 (20)

I added the missing information right above the “Place order” button. By placing it above the button rather than below, it is more likely to be read since the reading direction is downwards. Still, I consider the information to be low on the hierarchy of importance, and should thus not be attention grabbing.

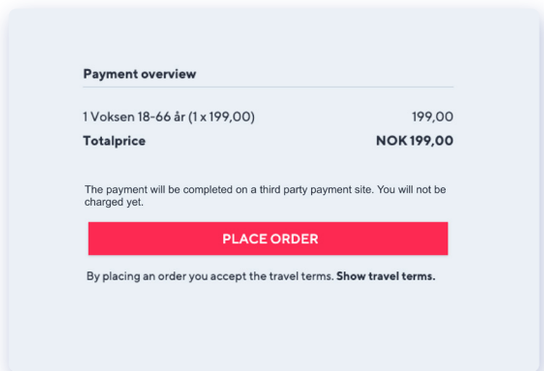


Figure 31: The fix for problem 39/52.

Problem 40/53 (20)

The problem isnt exactly the same anymore since the whole page has been removed, but the information still needs to exist. I chose to put it on the main page.

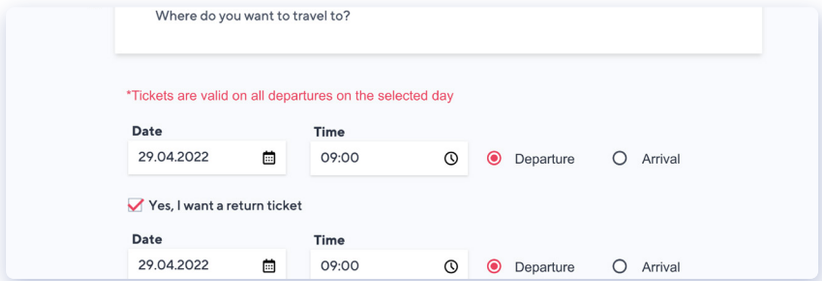


Figure 31: The fix for problem 40/53.

Reflections after doing the fixes

Time

I fixed 40 out of 61 identified usability problems. I had estimated the process to take me 15 hours and 20 minutes, but I ended up spending no more than 4 hours and 34 minutes in total. Only one time estimate was too pessimistic, taking 10 minutes longer than I expected.

The biggest time saver was discovering that many of the problems were not problems I could fix with design. These problems were mainly related to unsatisfactory code and lacking translations.

Design vs CS related problems

Although many of the usability problems could be blamed on unsatisfactory programming, it's too easy to put all the responsibility on the developers when it's part of the designers' job to test the website. My results after the heuristic evaluation prove that many programming errors could have been uncovered with more thorough testing.

It became apparent to me that the line between design and programming can be ambiguous. For example, it could be the developer's "fault" that an error occurred, but it's also the designer's job to design good error handling, such as descriptive error messages or backup functionalities. It is also the designer's job to test the design thoroughly to uncover these errors.

Dialogue elements can also be designed to omit user mistakes entirely. For example, an input field can be designed so that only certain inputs are allowed (e. g. choosing from a list of options), instead of allowing the user to enter invalid inputs.

The missing development team

Designs are ultimately made to be implemented. Consequently, there needs to be a running dialogue between the designers, developers, and the rest of the development team. Part of a designer's job is to communicate to developers what they are to develop, while part of a developer's job is to communicate back what is and is not technically viable.

As I did not have a team, I chose to approach task of redesigning very conservatively. For example, I assumed that the search algorithm cannot be changed.

I found it very difficult and unnatural to make these decisions on my own, and therefore chose to err on the side of caution.

Evaluating the redesign(... and the original website again)

After producing design solutions, the next step in the HCD process is evaluation of the design. This step will suggest whether the usability problems found during the heurist evaluation are still present, and it may give me new insights into user needs and how my design can be improved further (ISO, 2019).

Evaluating the original website again?

In addition to testing my redesign, I decided I wanted to evaluate the original website once more. Usability testing would be my evaluation method of choice, and I could henceforth compare the two methods on equal ground.

Heuristic evaluation may uncover many usability issues that are not found with usability testing, but the opposite is also true (Nielsen & Mack, 1994). Both ISO (2019) and Nielsen and Mack (1994) argue the benefits of using both.

Two tests in one

I stuck to the same basic steps of usability evaluation of both the original website and the redesign, but because of obvious differences in the two versions it was not sensible to conduct the tests identically.

For example, the prototype had only been built to test the fixed usability problems, which only amounted to 7 different screens. It's expected that I would find additional usability problems that I was not yet aware of, but they would be limited to the same 7 screens. The original website, on the other hand, is fully functional (meaning that everything that looks interactive is interactive) and adds up more than 100 unique URLs within the same domain. There's certainly bound to be something on the other 93+ screens not existent in the prototype.

Having the whole website available to explore, also allows for greater freedom during the usability test. The user can to a larger degree "customise" their own scenarios to make for a more realistic use case. It also allows the user to make "mistakes", meaning that they can attempt to complete the task differently than I expect them to but the links are still functional. This fact is extremely valuable in assessing the usability of the website as a whole, the basic service design, and whether my understanding of the user and context of use is correct.

A/B-ish Testing

To save myself some time, I decided to test both the original website and the redesigned website with the same users and chose to borrow some tricks from classic A/B testing.

A/B testing is usually used to compare two different versions of a design to optimise it in one way or another. It's highly quantitative and detail focused in its nature and will not give greater insights into the users and how they use the design (Martin & Hanington, 2012; Nielsen, 2005).

I used a sample size of five users. It is self-explanatory that I would not be able

to collect statistically significant results from such a small data pool. Therefore, I fully kept the qualitative qualities of usability testing and followed the method as I otherwise would except for testing two versions at once.

There are obvious downsides to doing it this way as the test participant will be coloured by the first test when starting the second. The fact that I am testing both a prototype and a fully implemented website may also contribute to confusion between the two. It is possible that the participants will prefer the original website as it behaves more as expected, while only a few functionalities are available in the prototype.

Designing the usability test

Besides from the previously mentioned A/B inspired test setup, I followed much the same procedure as with AtB's ticket machines. I used the guidelines laid out by Nielsen (1993) as a base and modified it to fit my own 'use case'.

I chose to have the participants use their own computers during the test. This freed up my own computer so that I could make quicker notes, the participants are likely more comfortable with their own devices, and since some of the tests were conducted digitally, they were forced to use their own computers anyway. This meant that I could not record the screen for the physical sessions.

There were both pros and cons in doing it this way.

Pros:

- easier to recruit people spontaneously
- no need to book rooms in advance
- much simpler and much more laid-back setup
- cheaper

Cons:

- no recordings to look back on

Since I did not have any observers to take notes for me, I would generally have more to think about at once, increasing the likeliness that I miss something. I thereby modified the test template to make it easier for myself to catch as much information as possible.

I made sure to pause the test at natural points so that I could write notes. I also asked follow up questions throughout the test instead of in one go at the end. I informed the participant beforehand that I would be doing this. I believe this technique would be beneficial with recordings and observers as well since I would not have time to review the data before the debrief anyway. It may also make it easier for the participant to remember why they made certain choices and for the facilitator to remember exactly what the participant did that was worth commenting on.

Since this would be the first and only time I'd be able to talk to 'real' users about Flybussen and their website, I wanted to take the opportunity to ask some related questions about they travel habits and opinions as well. It would either validate my initial assumptions or prove them wrong.

Test setup

The following setup shows a how I would conduct a usability test where I test the prototype first. In the opposite case, they were simply flipped.

The test consisted of an introduction, the actual tests with a list of tasks, and a debrief. This took about an hour in total.

Introduction

I based my introduction on a script by Steve Krug. You can find my full revised script in Appendix X wher you can also find a link to the original. The biggest change I made was to highlight that we are testing both a bare bones prototype and a full-fledged and functional website.

Testing the prototype

I asked the participant to look at the first frame, which was showing the home page. Before giving them the tasks, I started by asking three questions:

- I want you to look at this screen and tell me what you think it is. Have you used it before?
- What other travel or travel purchasing apps and websites do you regularly use?
- What do you expect to find on such a website?

Thereby, I provided the participant with one task at a time. The tasks were as follows:

1. On the 1st of June you are flying to Oslo to visit a friend who will pick you up at Oslo Central Station at 10 o'clock. Buy a bus ticket appropriate for

this trip.

2. Let's go back to start. Now what would you do if you were going to Røa, but you didn't have a Norwegian keyboard?
3. You are standing at the airport but the bus is not coming.
    a. What would you do in real life?
    b. In this scenario you want to find out what's wrong by checking the website on your phone.

Testing the website

My main goal with testing the original website was to compare the method to how I conducted the method of heuristic evaluation. As discussed previously on page 79, I wanted to take advantage of the fact that the website is operational, as opposed to the prototype. This meant that I could investigate the user patterns more thoroughly and broadly. With inspiration from the usability tests of AtB's ticket machines, I wanted look at the website from a wider perspective and investigate how the users would approach tasks where they would need to figure out on their own that Flybussen.no is their correct destination on the world wide web.

I asked the participant to open a new tab in their preferred web browser before giving them the following tasks:

1. You are flying to Stavanger to visit the university there. You need to sort out your own transportation. How would you approach this task realistically?
2. Go to Flybussen.no. On the 1st of June you are flying to Oslo to visit a friend who will pick you up at Oslo Central Station at 10 o'clock. Buy a bus ticket appropriate for this trip.
3. Let's go back to start. Now what would you do if you were going to Røa, but you didn't have a Norwegian keyboard?
4. You are standing at the airport but the bus is not coming.
    a. What would you do in real life?
    b. In this scenario you want to find out what's wrong by checking the website on your phone.

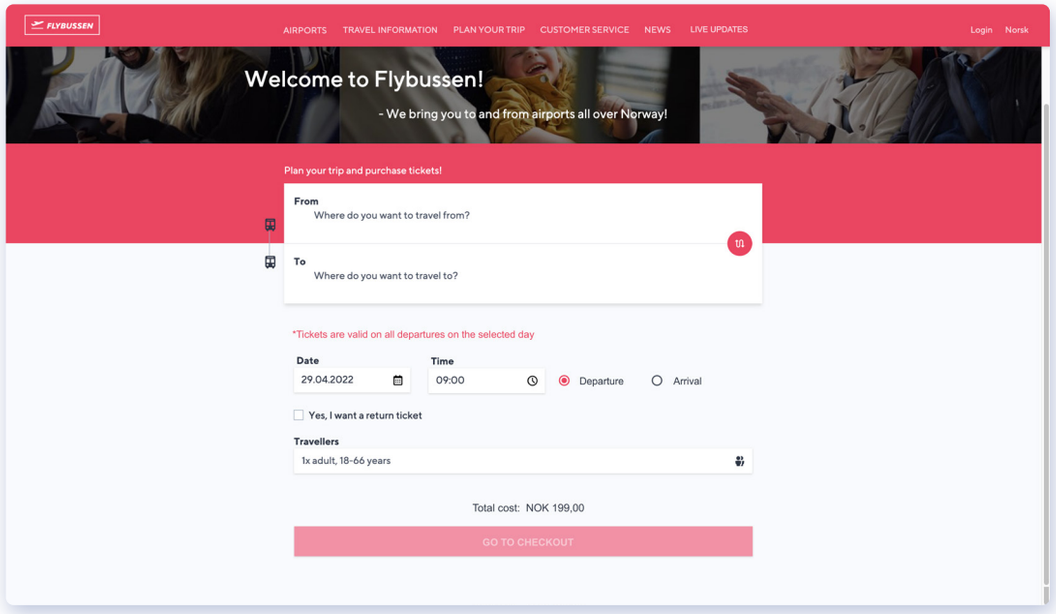
Debrief

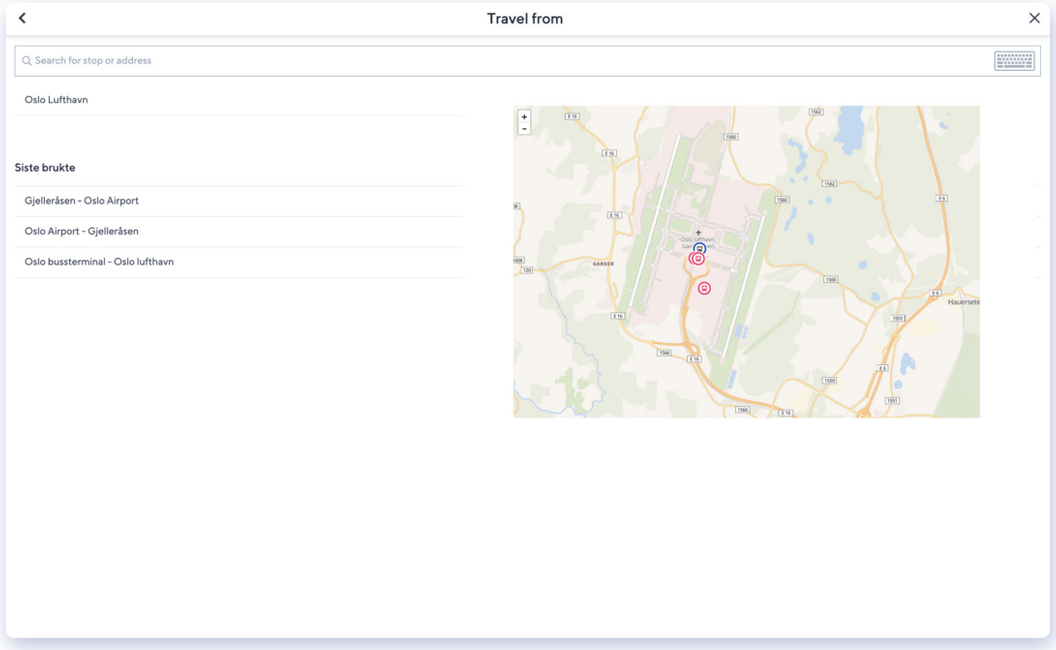
Many of the questions that arose during the tests were asked during the tests as well. Therefore, I used the debrief to talk about the tests in general, what the participant felt about them, and a few questions about their relationship with travel services and about the new features I included in the prototype. I made a list of prewritten questions beforehand:

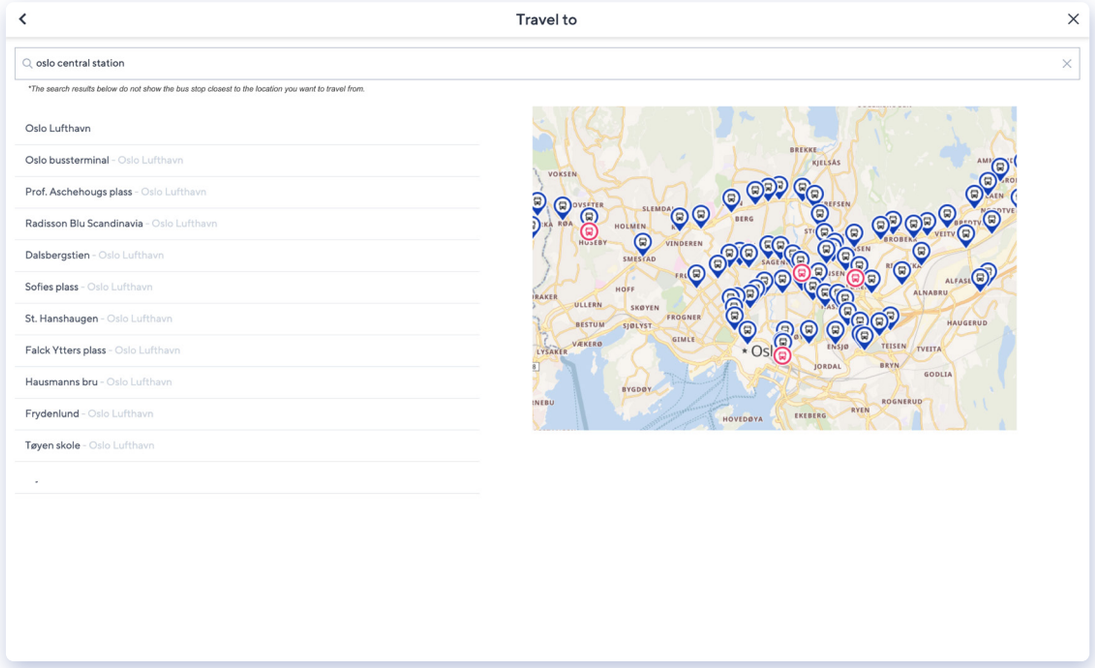
- How do you decide what airport transport service to use?
- Would you have bought a ticket from Flybussen online in real life? If no, what would you prefer to do?
- What do you expect to find on such a website?
- What functionalities do you use on similar websites/apps? (Vy, Flytoget, Værnesekspressen, AtB etc.)
- What would you expect to be able to do with it?
- Do you have any examples of travel apps that you like? What do you like about this app?
- On the search page in the prototype, what did you think of the map?
- How do you feel about not choosing a specific travel time?
- Would you like to know exactly when the bus is supposed to be there? Where do you expect to find this information?

Prototype

After finishing the rededesigns which were presented in the previous chapter, I put all the changes into one interactive prototype in Figma that I could test with users. On the next few pages I will be going through each frame in the order of the test scenario I created.



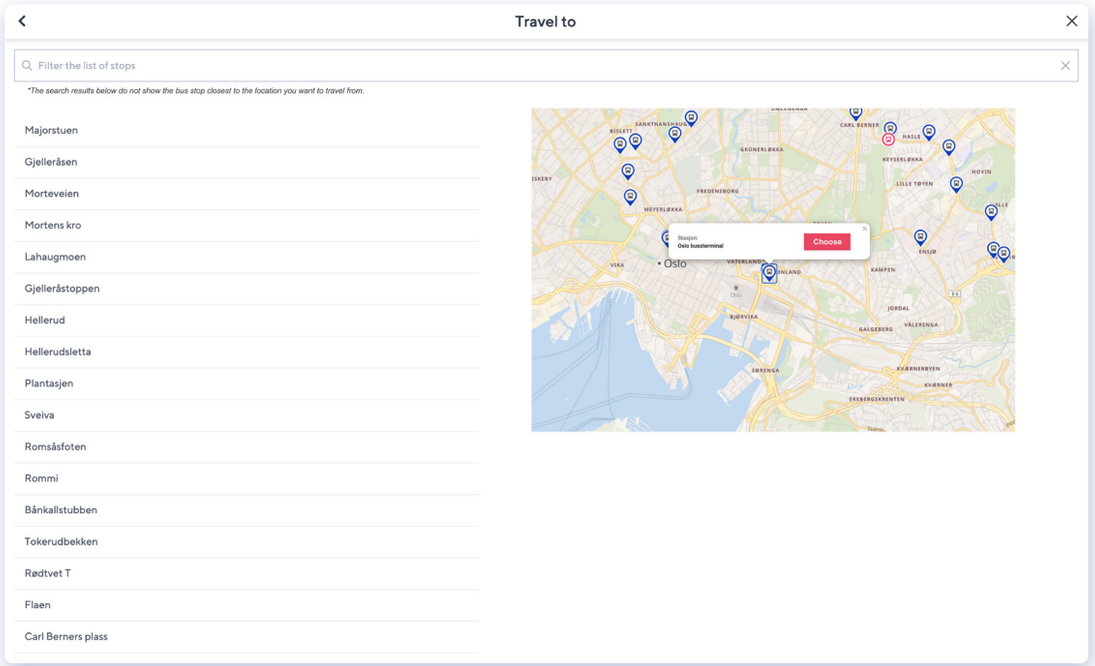


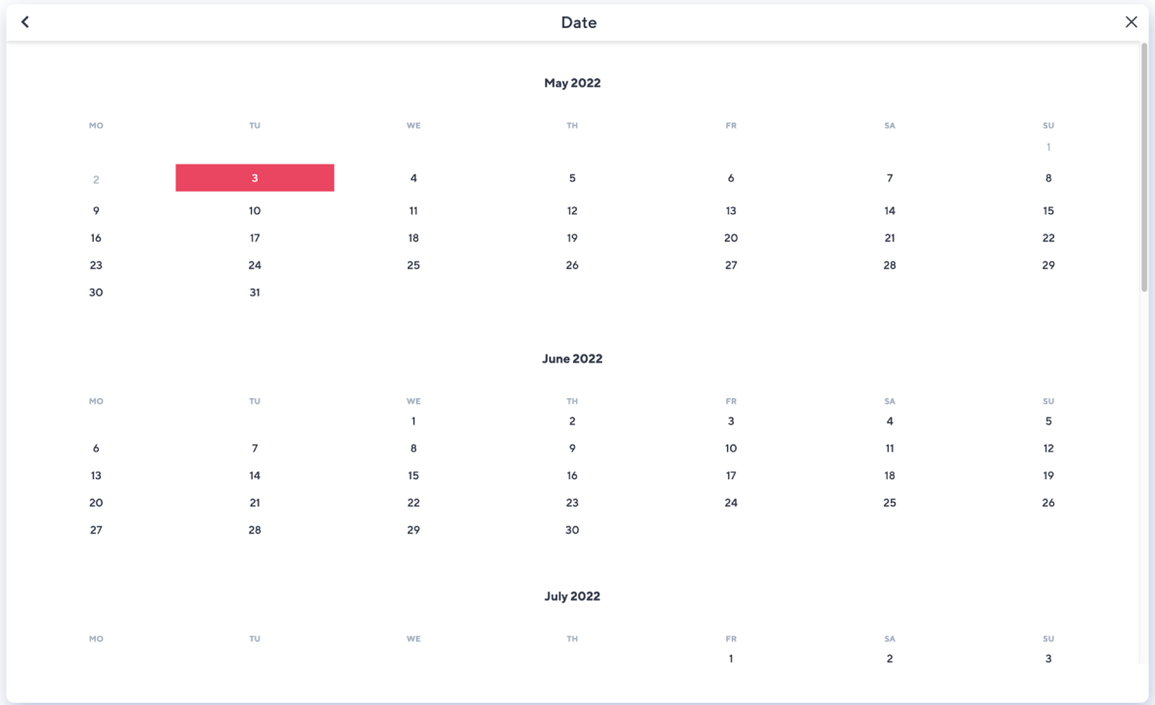


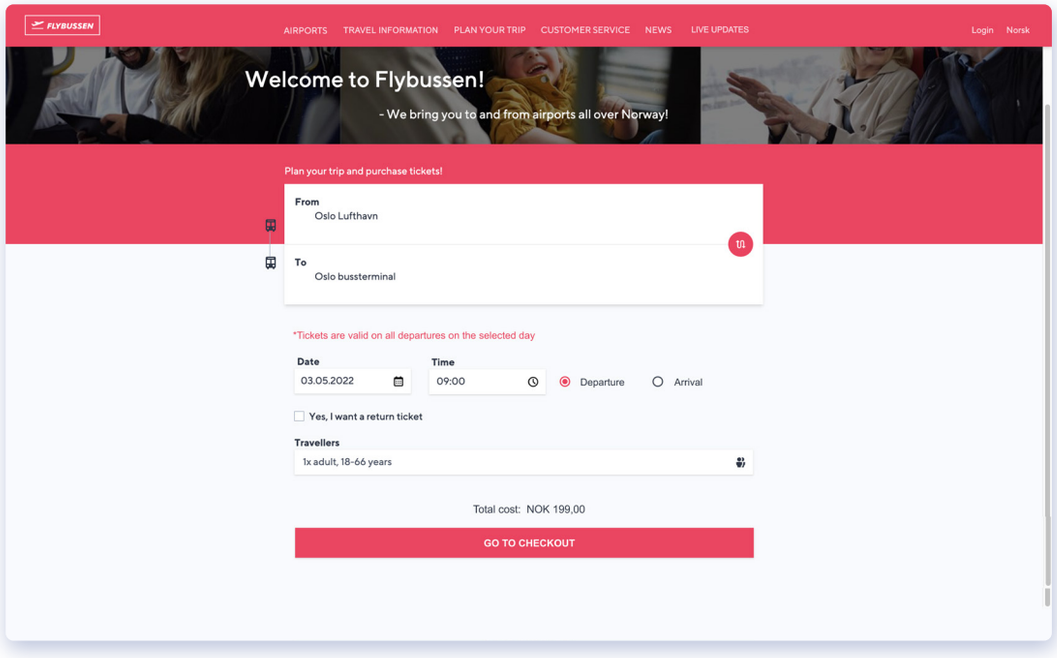
The participants first see the frame on the top left. Their task is to buy a ticket. Assuming they will start filling out the form from the top, they will see the bottom left frame when clicking the “from” input field.

Gardermoen is displayed on the map by default.

When they enter “oslo central station” into the input field, the same results show up as on the original website as seen in the frame above.

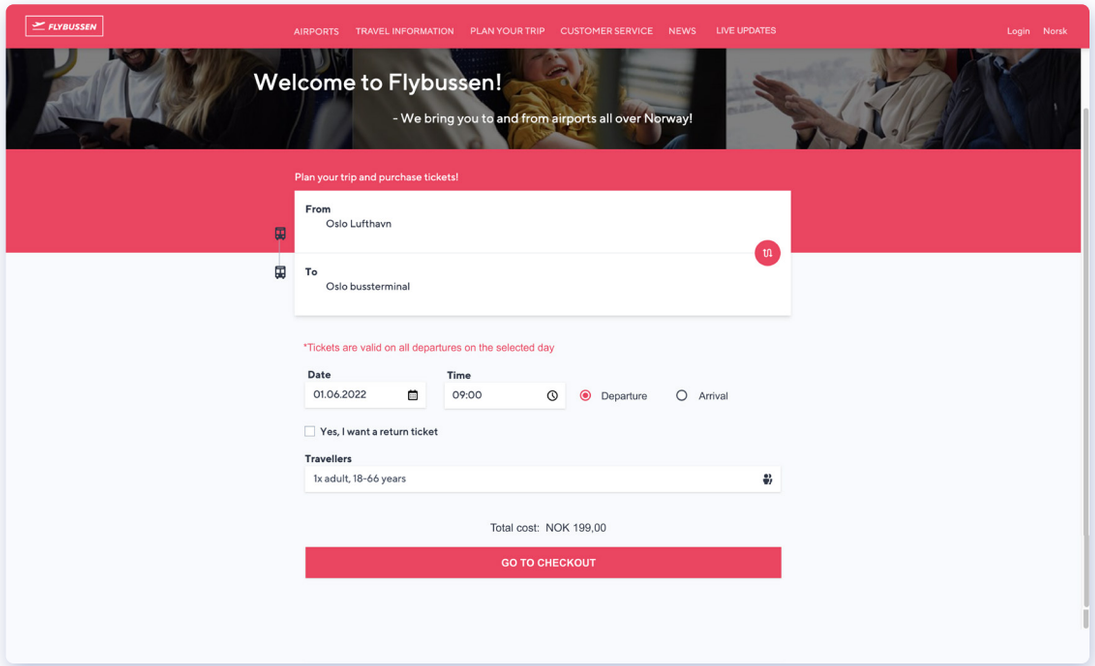


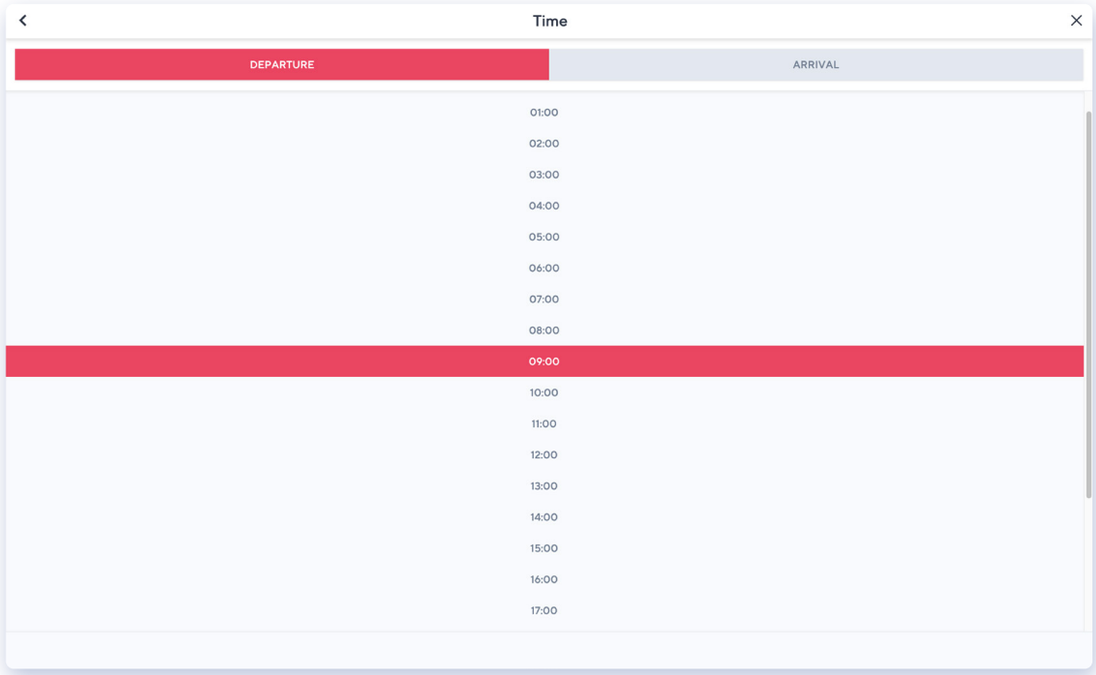


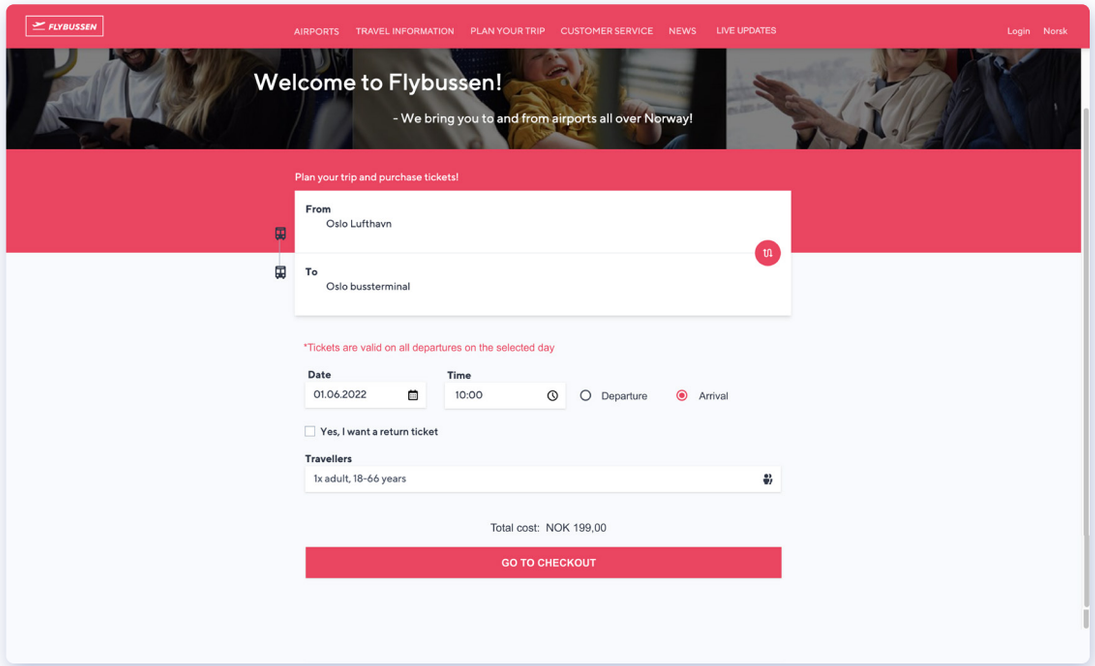


The participants may also click the pins on the map to see details about each bus stop. In the frame in the top left Oslo Bussterminal has been selected, revealing the option to choose a destination from the map instead of the search.

After entering the desired bus stops, the participant should select date and time next.



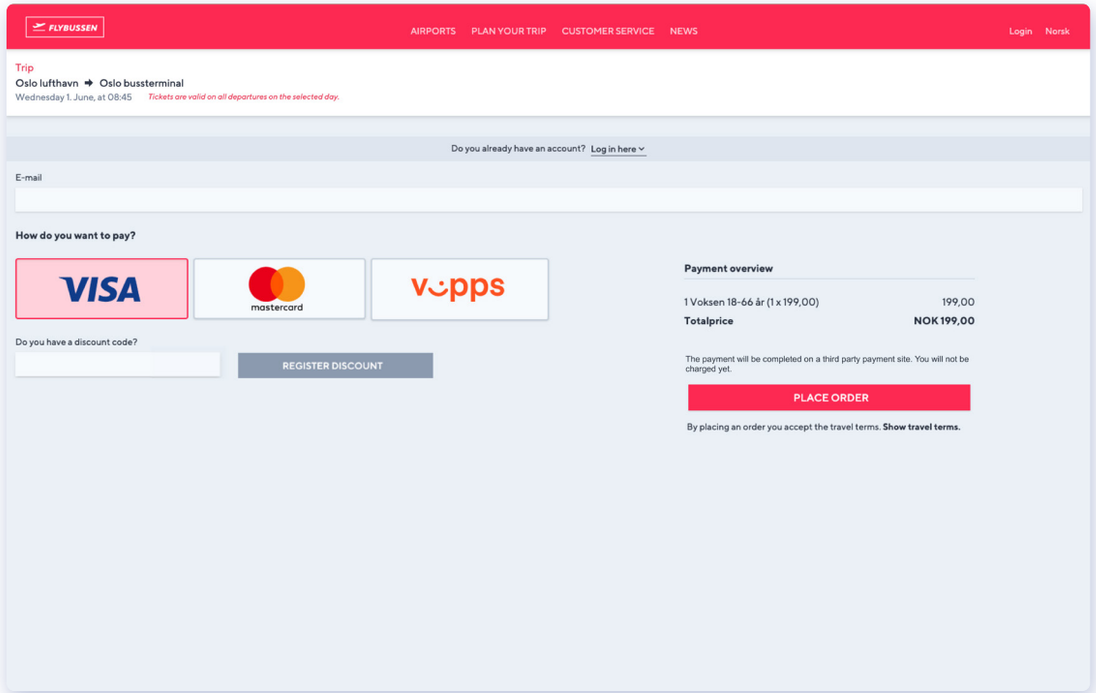


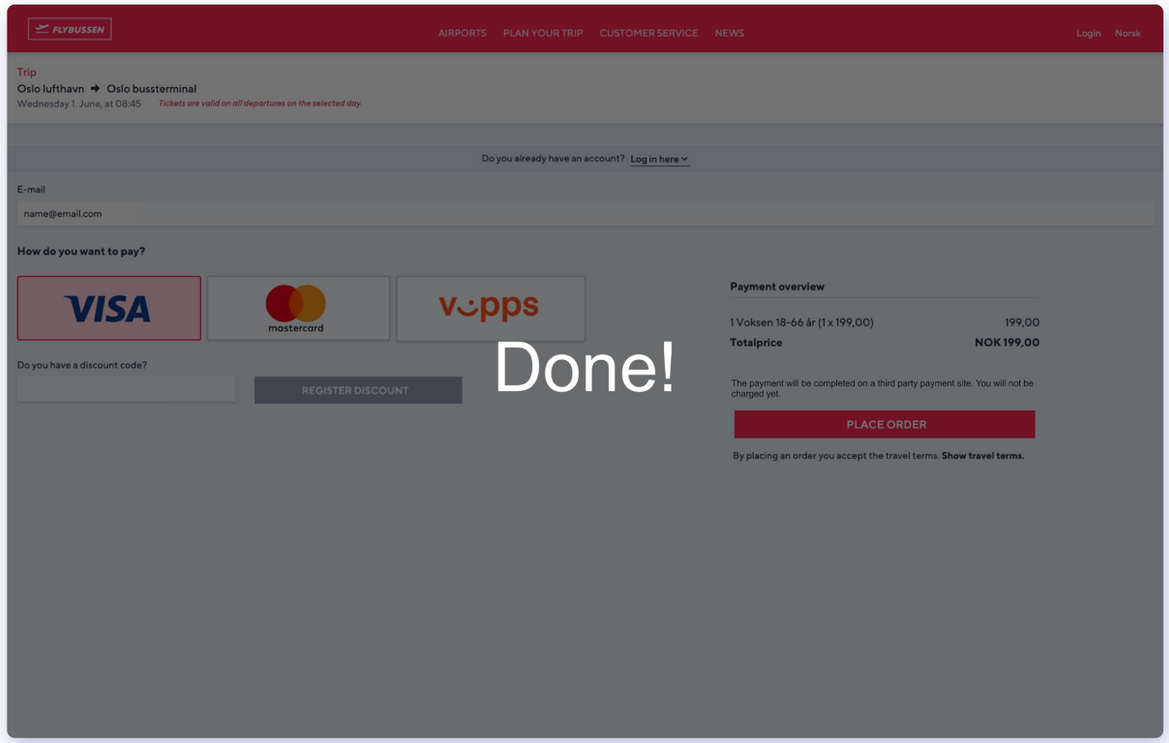


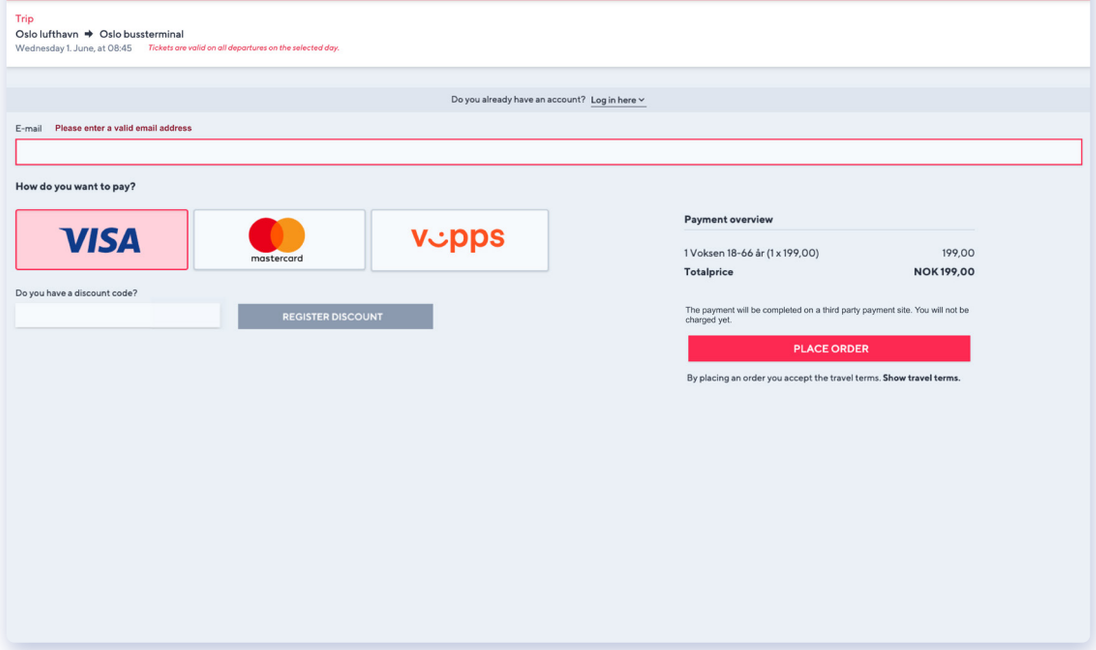
I opted to keep the departure/arrival options in the popup menu, as seen in the frame to the left. This will be carried over to the front page, where it can be edited.

For this scenario the passenger count is correct and the participant should not need to open this menu.

Notice how the total price is displayed the entire time.

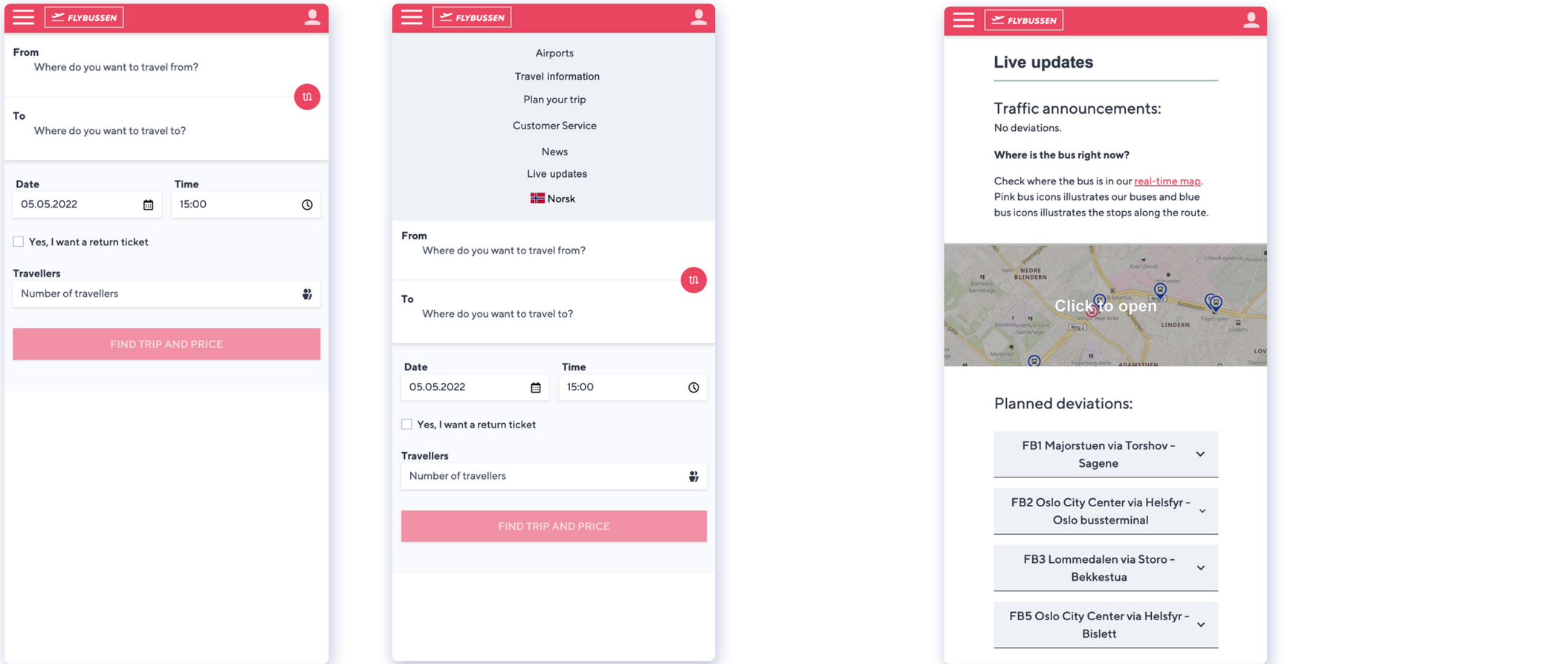






The participant is brought to the top left frame when clicking on “Go to checkout”. If they forget to enter their email before clicking “Place order”, there will be an error message, as shown to the left.

Then the task is completed.



The next task involved a phone. The goal of the participants was to figure out why their bus is late. For this I only needed three frames.

There is the front page to the left, the middle frame shows an open hamburger menu, and the right frame shows the “Live updates” page.

Results & discussion

I conducted the usability tests with 5 participants. Two of them tested the prototype first, while three testes the original website first. Two tests were conducted digitally. To keep it short, I will be discussing the most prominent findings only, which I deem sufficient for conveying the effectiveness of the evaluation method.

How it really went down

Regardless of how structured my agenda and questions were, it didn't go exactly as planned in either test. Many of my prewritten questions were answered spontaneously during a test or I found myself wanting to ask them at certain points where it felt natural. I do not believe this affected my results negatively.

Travel habits

The participants' answers when asked about their travel habits in the debrief varied greatly. One would use Google Maps for all their journey planning. One doesn't travel much and would figure it out as they go. Most stated Entur and Vy as examples of go-to travel apps and would continue to mention these services when talking about Flybussen and its flaws.

A common theme was that no participants seemed to ever actively look for new travel options, meaning that, for example, is not enough to attract new customers. The service either need to be known beforehand, or for the ones who figure it out at the airport, the service needs to be present and easy for the customers to pick.

Deceivingly good looks

When asked about what they think about the website in general, three of the

BUS

To and from Stavanger:

Tickets	Payment options	Stavanger	Grannes	Airport area
Adult (incl. 4 children)* *Family ticket - not applicable for bigger groups/class trips/ sporting teams	On the bus	NOK 180,-	NOK 140,-	NOK 50,-
	Online www.flybussen.no	NOK 149,-	NOK 115,-	NOK 50,-
Return tickets (valid for 75 days when purchased on flybussen.no or in Vipps)	On the bus	NOK 270,-	NOK 210,-	-
	Online www.flybussen.no	NOK 224,-	NOK 173,-	-
Children/Students/Senior (Student discount only valid with proof of education from University or High School)	One the bus	NOK 140,-	NOK 110,-	NOK 50,-
	Online www.flybussen.no	NOK 115,-	NOK 90,-	NOK 50,-

*1 adult + 4 children, one way ticket only

- Travel time: 20-30 minutes
- Timetable: Flybussen.no

Any questions or feedback can be directed to Flybussen.Stavanger@boreal.no or by phone +47 51 59 90 60

Figure 32: The table on avinor.no showing all available travel options to and from Stavanger Airport.

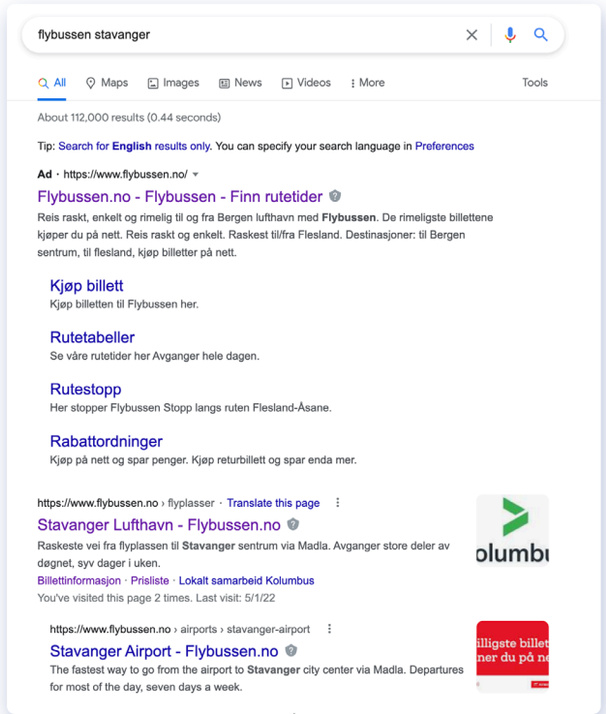


Figure 33: Screenshot of a similar google search to many of the participants. Most clicked on the second result, which mentions Stavanger in the title.

participants immediately commented on its appearance. This was interesting to me as I was mostly interested in, and expecting, answers about the website's usability and functionalities. It was described by participants with adjectives such as "pretty", "modern looking," and "tidy." From my (statistically insignificant) results I could not differentiate the results of these participants with the others in terms of how positively their user experience was. Still, this may suggest the importance of appearance in interfaces to a user as described by the Aesthetic-Usability Effect (Moran, 2017).

Results for the original website

Finding the website

When asked to arrange their own transport from Stavanger Airport to Stavanger University, three participants stated that they would probably just grab a taxi or figure out the bus arrangements at the airport. In these cases, I urged the participants to continue the thought experiment pretending that these options were unavailable.

Everyone utilised either a search engine or Google Maps with search words such as "Stavanger Airport to Stavanger university" or similar. A couple of participants wanted to first find transportation to the city centre where they would take local public transport to their destination and did not consider the possibility that the airport bus could have a stop near the university.

One of the top search results on Google was a link to Avinor's website which gave an overview of transportation options from Stavanger Airport (Figure 32). This is where most participants ended up finding the website in the end.

Since there is no direct link to Flybussen.no in the table everyone opted to google "Flybussen Stavanger" or something similar (Figure 33). The top result is a link to Flybussen.no's home page. However, all participants chose to click the second search result, with the header "Stavanger Lufthavn – Flybussen.no". The link opened to a tertiary page on the website which displayed an image of the bus's route and "News" (Figure 34). Several participants were confused about what to do next and clicked out of the website completely to google more or find alternative transportation options. There

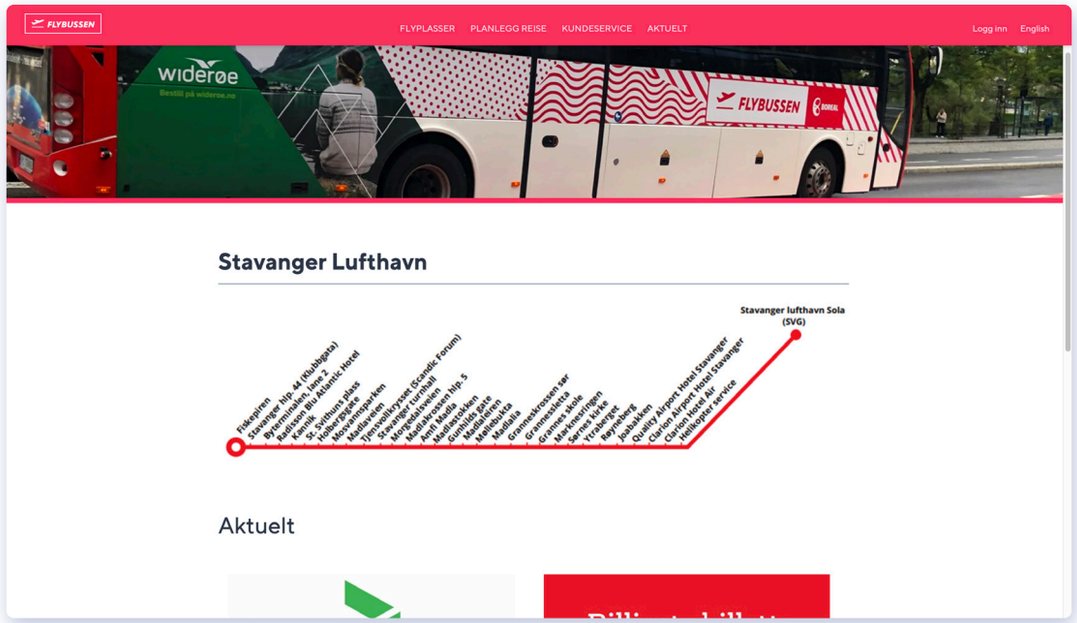


Figure 34: The subpage on flybussen.no where many of the participants landed after their Google search.

"correct" thing to do, would have been to click the logo in the top left corner. This would have brought them to the home page where they would be able to purchase tickets.

One participant who did click the logo and was brought to the home page, didn't understand that's what they had done as they were under the impression that the "Stavanger Lufthavn" page was the home page, and remarked how weird it was.

This problem is remarkably similar to the problem with AtB's ticket machines which may (or may not) be difficult to find. The same arguments apply regarding the relevance of the issue. The problem does not exist looking at the website as a closed system, which is what I set out to do. However, by testing the website within a larger context, I was able to identify at least one internal usability problem, being that when many users enter the website for the first time they don't land on the home page.

Search

From getting to the home page an onward everyone set out to purchase a ticket.

Several participants faced problems with the search and poor search suggestions and expressed an expectation for search results similar to how e.g.,

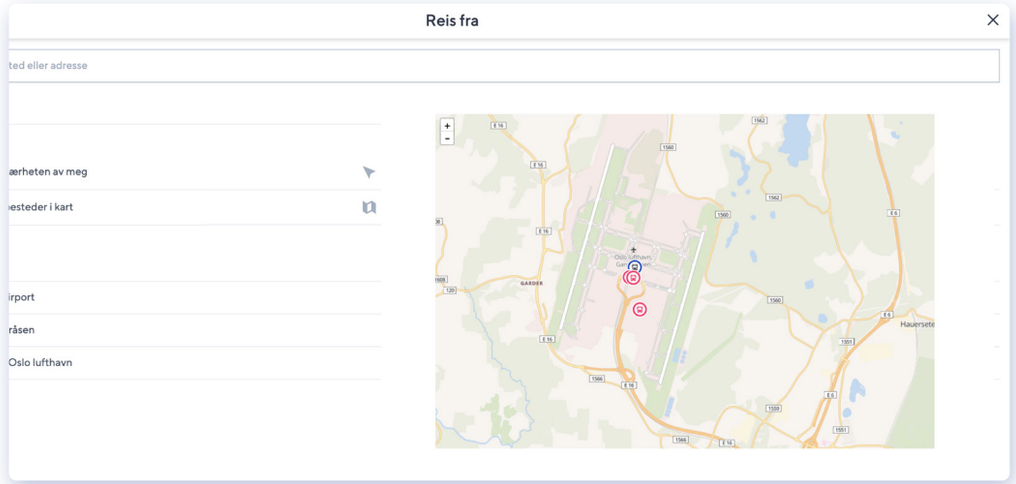


Figure 35: The map that I included on the search page was not well received by the participants.

Entur's search works, where search words don't have to be as exact, and the database contains locations other than bus stops. Some needed to go back to Google to figure out what they should be searching for instead. One participant used the image shown in Figure 34. This uncovered the fact that the image is not up to date and does not display the bus stop closest to the university. It was also difficult to find because they had no idea where they had seen it the first time.

The live map

None of the participants were able to find the traffic announcements and live map on the original website. However, when asked about how they would handle the situation in real life, they all stated other strategies for getting information that do not involve using the website. Such strategies included calling for help from their friend waiting for them at Oslo S, getting a taxi or Flytoget, or going inside to look for assistance from an employee.

Many stated that it did not make sense to look for this information on the website because they would be too stressed or simply didn't trust the website to upload this kind of information correctly. This suggests that the live map is not needed at all.

When asked about it the icons on the live map (after me guiding them there), two participants interpreted the blue pins to be inactive buses, while one participant interpreted the pink pins to be Flybussen buses and the blue pins to be buses from other companies.

Results for the prototype

Confusions about the prototype itself

Two of the participants were not familiar with the concept of digital prototypes. This caused them to expect the prototype to behave as if it was a "real" website and therefore judged it harshly when it couldn't perform thereafter. For that reason, I chose to discard some of the comments made from these participants when I deemed it obvious that it came from a place of not understanding what a prototype is.

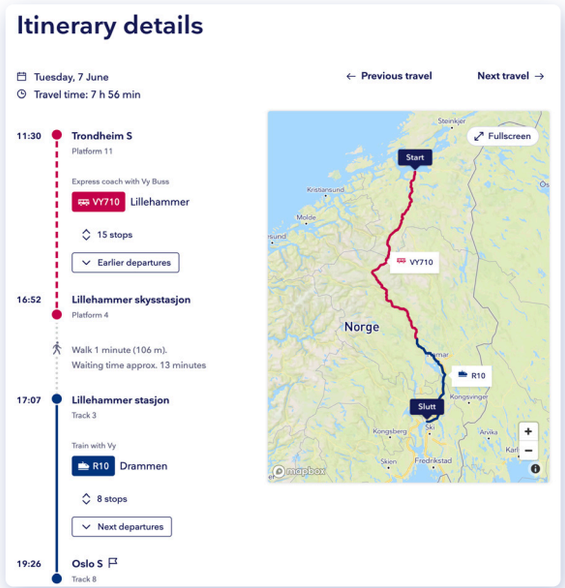


Figure 36: The itinerary on Entur not only provides the entire route end to end. It also displays a map of the selected journey.

The search page map

No participants were the least interested the map on the search page, and most didn't even acknowledge it. Asked about it further, they saw no use for the features the map provided. Several users expressed a wish for the map to display the selected route, one participant giving Entur as an example.

These results suggest that I was probably too quick with this design fix and missed the mark on what users actually want and would use.

Removal of "find trip" page

All participants were clear about wanting to know when their bus would be departing. By removing the "find trip" page I had also removed information about the bus schedule. The information does, however, partly exist but it's buried very deep on the website on each airport's respective news page.

(Few participants commented on the existence of the "find trip" page on the original website as they did not see the text saying that the tickets are valid all day. The one participant who did see it, on the other hand, was clearly annoyed about having to pick a departure when it wasn't needed.)

Miscellaneous

All particpants were able to find the live map quite easiliy. Only one participant wanted to click something else first in the hamburger menu.

All participants easily figured out the small "ÆØÅ" keyboard.

Two users did not enter their email before clicking "Place order," but they immediately noticed their error when the error message and bright red stroke appeared. This suggests that the email input field still isn't visible enough, but it is also proven that the error message works.

No users found the map features as none of them chose to interact with the map.

Identified usability problems

From the results of the usability tests on the original website I composed a list of all usability problems I could identify. The list does not give an indication of how severe each problem is and how many times they occurred in the tests.

- It needs to be obvious where to click to buy a ticket on every page on the website.
- It needs to be obvious to the user where they are on the website.
- The search is unsatisfactory.
  - Some search words don't result in relevant search results.
  - It's difficult to find out which bus stops exist.
  - It's impossible to find out which bus stop is closest to the user's final destination.
- The live map is way too hidden.
- The meaning of the icons on the live map are unclear.
- The phone number for customer service is too hidden.
- The "find trip" page is superfluous.
- Arrival/departure information should be visible on the home page.
- Traffic announcements are not to be trusted.
- The website doesn't help the user plan their journey all the way to their final destination.
- The passenger selection menu doesn't display prices.
- Global menu has bad naming.

Next steps

If the project were to continue and the design were to be implemented, there are a long list of tasks that I would recommend to do.

Redefining the context of use

I am in a way no closer to narrowing down the context of use with Flybussen's website, than I was after evaluating AtB's ticket machines. No one in my group of test participants had taken, or even heard of Flybussen.no prior. This could suggest that they are not within the real user group.

However, through their answers to my general questions about travel habits, I could infer that for most of them convenience was key when choosing transportation to or from the airport.

According to my own anectdotal evidence from having taken the bus from Værnes airport many times, I have seen that a large portion of the passengers (maybe even most) buy their tickets on the bus, even though it's more expensive. To me, this suggests that they were ready to jump on whichever form of public transport that was closest to the airport door. It is therefore reasonable to assume that the test participants are likely to do the same thing. This is supported by several statements from the participants themselves.

One participant swore to Google Maps for planning. During the test, Google Maps did recomment Flybussen as the best route from Stavanger Airport to Stavanger University, which was enough for them to have chosen this route. Another participant prefered to figure it out as they go. As Flybussen is the sole airport bus from most (if not all) airports on their list, it is not unthinkable that seveal of the participants would jump onto this bus if there was no other apparent options.

For all of these reasons I am leaning towards counting all the participants as primary users of Flybussen even though they have never used it before. This is a good start, but it should be investigated further.

Based on all the information I have gathered I redefine the context of use as so:

1. There are still two stakeholders: Flybussen and everyone (see definition on page 5).
2. The user characteristics I will still be considering are anything that my definition of everyone covers. I define the website to be a walk-up-and-use system, and I will thus keep in mind that users can be completely inexperienced with the service or similar services.
3. The main user goals include the following tasks: purchasing a ticket, journey planning, reaching customer support, and finding out whether the service applies to them (in terms of price, airport, bus stops etc.).
4. The user tasks can be performed anywhere and anytime. Ticket purchasing either happens at home in advance where the users may be more likely to use a PC, or on the go on their phones. In both cases the user already knows which transportation method to use. Some may also use the website to buy a ticket at the airport when they see the bus with their own eyes. I believe journey planning and researching travel options also happens mostly at home in advance.

Specifying user requirments

Through my general questions about travel habits, I was able to start gathering an understanding of what the users want and need, still assuming that the test participants are likely users of Flybussen.

Most users will likely be using their phones, which is something Flybussen has already designed their website mainly for. If the users are buying a ticket either on the way to or at the bus stop, it is a requirement that the task should be quick and easy with as few steps as possible.

For users who are trying to plan their journey, the desired information should be easily available and digestible.

Similarly, customer service should be easy to find for a user who is stressed out and in need of immediate help.

Next next steps

With both the updated context of use and user requirements in mind, redeisgn tasks should be prioritised and new design sulutions should be created before evaluating the design again.

On a less formal note, I would first and foremost recommend focusing on the search as this is something every test participant struggled with. In the next iteration, I believe it would be a good idea to gather ideas from similar services, such as Entur or Vy, which were both mentioned several times during the tests.

To focus on the search, I consider it a given that a full developement team be involved in the process. During the previous round of redesigns I struggled making decisions on usability problems I wasn't qualified to determine the technical feasability of. This problem is still present in the next iteration of redesigns.

Final discussion and conclusion

After closing both the cases of AtB's ticket machines, and two rounds with Flybussen.no, I have gained an increased understanding of the application of usability evaluation methods and how they relate to ISO 9241-210:2019. In this last chapter I will discuss in more details what these lessons entailed, and how I assessed the effectiveness of the methods in the context of walk-up-and use systems.

Small experiments

My process throughout the project consisted largely of trial and error. Some things worked and others didn't, and some things only worked in certain contexts.

Testing with myself

idea for a self-test was mine and mine alone, and I think it has been demonstrated to me why it's not recommented in any literature that I have been able to find.

In order for me to design sensible usability tests to evaluate the ticket machines I needed to learn about how it worked. Through my self-test I did find possible usability problems and I did get ideas for how I should be testing it with users. The way I went about it most definitely taught me about the workings of the machine, but I believe I went a little bit overboard and got myself too emotionally involved. My view on the usability of the machine was most definitely colored by my own feelings about it, and it unconsciously made me very vulnerable to confirmation bias.

I won't say that my usability tests were doomed after performing the self-test, but it did make it difficult for me to separate myself as a user and as a the designer

Looking back, I don't think I would perform a scenario the way I did. I would rather simply try the machine and explore its functionalities. It could even be a fun idea to borrow some tricks from heuristic evaluation in order to be structured and methodological in my approach.

Heuristic evaluation vs usability testing

One reason why I chose to evaluate the original website again with usability testing was because I was curious about the difference in results I would get. The way I applied these methods definitely uncovered clear differences in the results.

I was able to gather a much higher number of usability problems through the sessions with heuristic evaluation than usability testing. There are several reasons this could be the case. The evaluators are usability experts and will find flaws with the website that won't affect many, or even most users, such as universal design principles.

Their job in the exercise is also different from the test participants. The evaluators were actively searching for usability problems, while the test participants were focused on accomplishing their tasks. By being goal oriented, the test participants were only interacting with dialogue elements which they believed would bring them closer to their goal. The evaluators on the other

hand, were to a larger degree testing out different dialogue element just for the sake of testing them, and were as a result finding many more usability problems than they would otherwise.

Another big difference was my role as a facilitator during the execution of both methods. In the case of usability testing, usability problems were identified by me analysing the participants' actions, statements, and particularly their struggles. It was therefore up to me alone to interpret the data and turn them into a list of usability problems. In contrast, with heuristic evaluation this job was completely out of my hands and in the hands of the evaluators. With the perspectives of four different people, I am not surprised that their combined effort resulted in a higher number of usability problems.

Additionally, I did not use any specific criteria for deciding whether or not something should count as a usability problem, which may have made me more critical of what should and should not count. With a specific list of usability heuristics, this was never a question for the evaluators.

Every participant in the usability tests found usability problem 57, which was the top rated usability problem by the evaluators. Still, none of the evaluators found it. This finding suggests that the focus of the evaluators was less goal oriented. The test participants were provided with a specific time and day to travel, while it was completely up to the evaluators to decide what information to enter into the input fields.

Lessons learned

If I were to redo the heuristic evaluation and the usability tests (on Flybussen.no) with the goal of uncovering more usability mistakes, there are many changes I could make.

The application of the method described in Nielsen and Mack (1994) includes a lecture on the specific scenario the evaluators were to go through during the evaluation. I dropped this part because it is a walk-up-and-use system and I therefore deemed it obvious what the website is meant for. Looking back, a specific scenario may have helped the evaluators be more goal oriented and to a larger degree see the website from a user's point of view. I believe this would have resulted in, at least some, evaluators finding problem 57.

On the other hand, by letting the evaluators roam around freely they may be able to find usability problems that I otherwise wouldn't have steered them towards.

As for the usability tests, a larger and more diverse pool of participants would have probably uncovered a different set of usability problems. By testing the website with someone who is visually impaired, I may have uncovered e.g., problem 43 (contrast too low on the "Find trip and prices button), not to mention invisible usability problems related to screen readers.

With larger websites such as Flybussen.no, I would need to make many smaller, specific tasks in addition to the big broad ones. With the tasks I gave the participants I was not able to, for example, test the "Q and A" page and how the order of the topics affect its usability.

Economic differences

Both Nielsen and Mack (1994) and ISO (2019) claim usability inspection, and especially heuristic evaluation in the prior author's case, to be a cheap and easy alternative to usability testing. This is an alleged advantage of the method that I was not able to observe. I had no "proper" compensation to offer any of the participants in neither the heuristic evaluation nor usability tests, and recruitment, planning and scheduling efforts were approximately the same in both cases. In reality, I was out by about 400 NOK more after the ticket machine tests having to pay 41 NOK in fares each time in addition to the chocolates I gave away, while I paid nothing for the heuristic evaluations .

Usability testing as a tool to understanding the context-of-use

In the case of AtB's ticket machines – the fact that I was not able to define the context of use any further based upon my self-test and subsequent usability tests, proves to me that the methodology was not appropriate for this purpose in this context.

It is clear to me now that I do need some knowledge about the system in order to make assumptions about it. Although I went about it with the goal of evaluating the usability of the machine only, and not worry so much about being correct with the context of use and the user requirements, these results show the importance of doing these first steps in the HCD process properly.

The bigger picture

Limitations of usability testing

Usability evaluation is not everything

As suggested in the reflections section of the AtB chapter, the method of usability testing alone will not necessarily fix underlying problems or create value. This applies to all usability evaluation methods. If I am right about the ticket machines being a dying service, then I would be wasting AtB's money by insisting they update they ticket machine design to make it a little bit more user friendly. It doesn't matter how great the user experience is if nobody uses it.

I therefore believe it's important to be mindful of the fact that user testing is not the answer to everything. It does not attract customers, make things pretty or increase revenue. It is simply a tool used with the goal of increasing the usability of a product or system, and the aforementioned effects may be a result of increased usability due to said usability testing.

The key is what is tested, why, and how the results are used. I want to argue that an overarching understanding of the current user base and the complete system in which the subsystem resides is crucial for keeping up with never ending societal changes. Therefore, continuous testing through the lifetime of the system is important as new problems can arise. This is supported by ISO (2019).

Usability testing and walk-up-and-use systems

Inside factors

When watching people try and plan a trip from Stavanger Airport to Stavanger University more than one participant found a route on Google Maps where Flybussen was the recommended mean of transport. However, none of them realised it because of the naming of the bus (FB5). This demonstrates to me that "inside" factors of Flybussen, such as their bus naming, may have an influence on the "outside." In this particular case it probably results in many lost sales because people can't figure out what FB5 means.

Another example of the same kind of problem was presented when several participants entered the website through a tertiary page. It is a possibility that Flybussen can't do much to change the Google search results, but they can help the user gain a sense of where they are on the website when they inevitably enter the website via an unexpected link.

These are not a big usability problem using the website from the inside. But coming from the outside and trying to get can be a usability nightmare. These are problems that would not have been uncovered if I had not created an open scenario.

Looking outwards instead of inwards

I created a strict and narrow scope for myself. My goal was to look at the walk-up-and-use systems as closed systems, with the argument that it doesn't matter why a user wanting to use your design or how they got there, because the user experience should be good no matter what.

It was very hard for me to stick to this principle and ended up with scenarios in my usability tests that covered more than the closed system isolated. In the case of both AtB's ticket machines and Flybussen.no I created scenarios where the participants were not told about the system that was being tested. I am happy I did so, because even proving it to be fruitless would be valuable.

On the contrary, I believe I was able to prove that there is value in looking outwards instead of inwards when evaluating a design. By making open tasks with endless possibilities of things going wrong, I am able to observe habits, user patterns, and problem solving methods that I would never think of myself. Additionally, I find that by having an outwards view of the system, and seeing the system as a subsystem within bigger systems, helps put the design in perspective and raises the question of whether there is any value in the design at all, not only now, but in the near and distant future.

Conclusion

In this project I was able to evaluate the usability of two walk-up-and-use systems, AtB's ticket machines and Flybussen.no. Based on said evaluations I was able to provide design recommendations and design fixes. These exercises gave me a good basis for assessing their effectiveness in my chosen context, given my own way of applying the methods.

Through this project I have learned two things:

1. **It pays off to look outwards when evaluating a design.** By always keeping in mind that no design exists in isolation, it possible to gain a deeper understanding of how a system works and thereby why things may go wrong.
2. **Apply usability evaluation where appropriate and don't neglect or underestimate the importance of other steps in the HCD process.** Without a good understanding of the context of use and well defined user requirements, a usability evaluation may not be able to reach its potential fruitfulness.

References

Garza, V. (2020). Metro ticket machines in Oslo to be removed and replaced by card readers in 2021 https://norwaytoday.info/news/metro-ticket-machines-in-oslo-to-be-removed-and-replaced-by-card-readers-in-2021/

Gibbons, S. (2021). 5 Prioritization Methods in UX Roadmapping. https://www.nngroup.com/articles/prioritization-methods/

ISO. (2019). Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems (ISO 9241-210). In: International Organization for Standardization, Geneva.

Karat, C.-M. (1994). A Comparison of User Interface Evaluation Methods. In Usability Inspection Methods. Katherine Schowalter.

Krug, S. (2014). Don't make me think, revisited : a common sense approach to web usability (3rd ed. ed.). New Riders.

Martin, B., & Hanington, B. (2012). Universal methods of design : 100 ways to research complex problems, develop innovative ideas, and design effective solutions. Rockport Publishers.

Moran, K. (2017). The Aesthetic-Usability Effect. Nielsen Normal Group. https://www.nngroup.com/articles/aesthetic-usability-effect/

Moran, K. (2019). Usability Testing 101. https://www.nngroup.com/articles/usability-testing-101/

Nielsen, J. (1993). Usability engineering. Academic Press.

Nielsen, J. (1994). 10 Usability Heuristics for User Interface Design. https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J. (2005). Putting A/B Testing in Its Place. https://www.nngroup.com/articles/putting-ab-testing-in-its-place/

Nielsen, J., & Mack, R. L. (1994). Usability inspection methods. John Wiley.

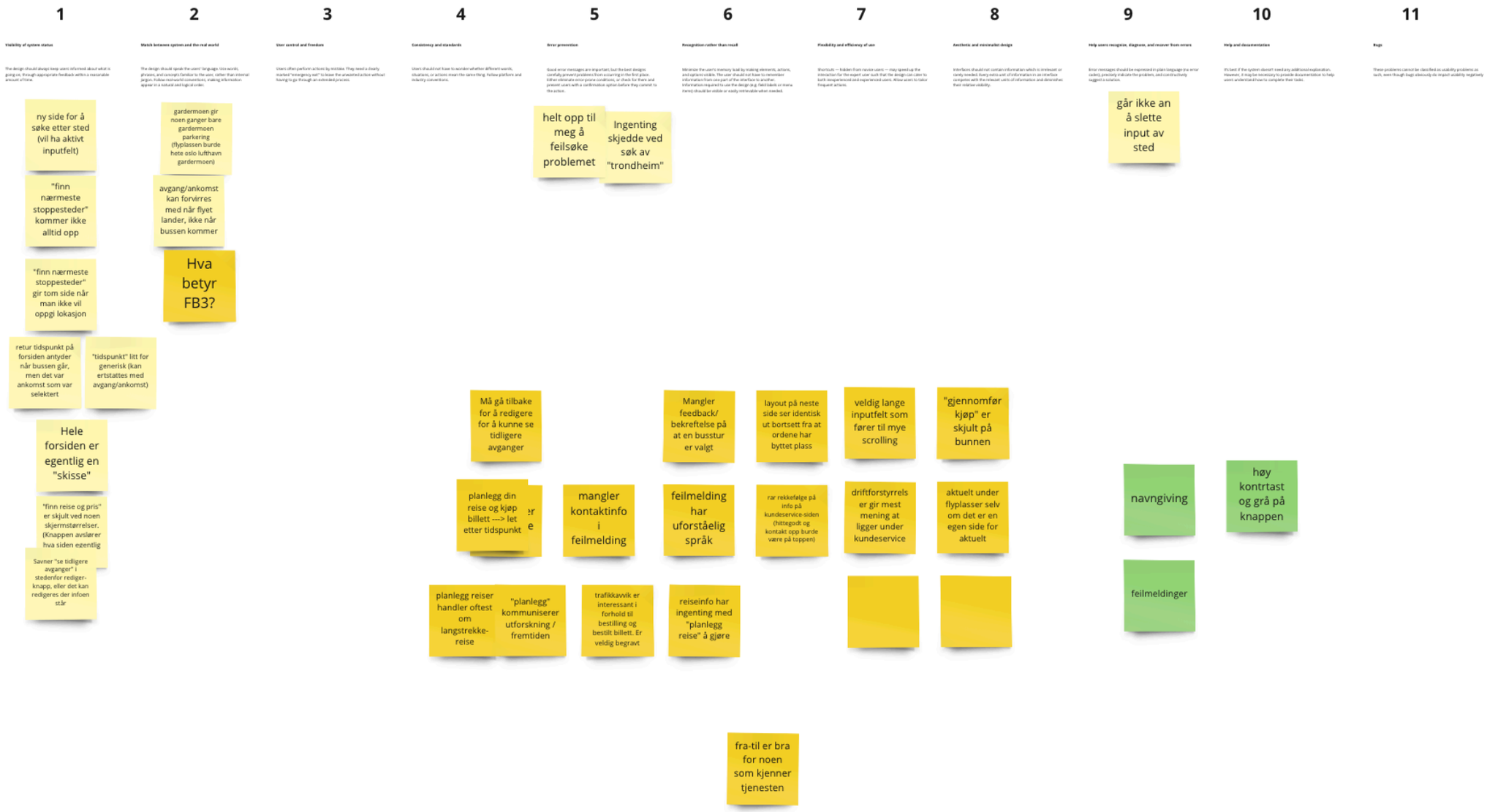
Norman, D. A., & Norman, D. A. (2013). The design of everyday things (Revised and expanded edition. ed.). Basic Books.

PressAssociation. (2017). Fifth of rail passengers buy wrong ticket from machines, says regulator. The Guardian. https://www.theguardian.com/uk-news/2017/feb/10/fifth-of-rail-passengers-buy-wrong-ticket-from-machines-regulator-paid-too-much

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. Journal of experimental social psychology, 13(3), 279-301. https://doi.org/10.1016/0022-1031(77)90049-X

Appendix

Appendix 1


1
Visibility of system status
The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.
ny side for å søke etter sted (vil ha aktivt inputfelt)
"finn nærmeste stoppesteder" kommer ikke alltid opp
"finn nærmeste stoppesteder" gir tom side når man ikke vil oppgi lokasjon
retur tidspunkt på forsiden antyder når bussen går, men det var ankomst som var selektert
"tidspunkt" litt for generisk (kan ertstattes med avgang/ankomst)
Hele forsiden er egentlig en "skisse"
"finn reise og pris" er skjult ved noen skjermstørrelser. (Knappen avslører hva siden egentlig
Savner "se tidligere avganger" i stedenfor rediger-knapp, eller det kan redigeres der infoen står
2
Match between system and the real world
The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.
gardermoen gir noen ganger bare gardermoen parkering (flyplassen burde hete oslo lufthavn gardermoen)
avgang/ankomst kan forvirres med når flyet lander, ikke når bussen kommer
Hva betyr FB3?
3
User control and freedom
Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.
4
Consistency and standards
Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.
5
Error prevention
Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action.
helt opp til meg å feilsøke problemet
Ingenting skjedde ved søk av "trondheim"
6
Recognition rather than recall
Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.
7
Flexibility and efficiency of use
Shortcuts — hidden from novice users — may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
8
Aesthetic and minimalist design
Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.
9
Help users recognize, diagnose, and recover from errors
Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.
går ikke an å slette input av sted
10
Help and documentation
It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.
11
Bugs
These problems cannot be classified as usability problems as such, even though bugs obviously do impact usability negatively
Må gå tilbake for å redigere for å kunne se tidligere avganger
Mangler feedback/ bekreftelse på at en busstur er valgt
layout på neste side ser identisk ut bortsett fra at ordene har byttet plass
veldig lange inputfelt som fører til mye scrolling
"gjennomfør kjøp" er skjult på bunnen
planlegg din reise og kjøp billett ---> let etter tidspunkt
mangler kontaktinfo i feilmelding
feilmelding har uforståelig språk
rar rekkefølge på info på kundeservice-siden (hittegodt og kontakt opp burde være på toppen)
driftforstyrrels er gir mest mening at ligger under kundeservice
aktuelt under flyplasser selv om det er en egen side for aktuelt
navngiving
høy kontrtast og grå på knappen
feilmeldinger
planlegg reiser handler oftest om langstrekke-reise
"planlegg" kommuniserer utforskning / fremtiden
trafikkavvik er interessant i forhold til bestilling og bestilt billett. Er veldig begravt
reiseinfo har ingenting med "planlegg reise" å gjøre
fra-til er bra for noen som kjenner tjenesten

1

Visibility of system status

The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.

2

Match between system and the real world

The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.

from: expected dropdown list

doesn't have correct keyboard to write Ås

offers return time before I set out

3

User control and freedom

Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.

came back to same place after error

4

Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.

oslo bussterminal to oslo airport

many Oslo airports

oslo central station suggests oslo bussterminal

why separate screens? losing context

5

Error prevention

Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action.

6

Recognition rather than recall

Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.

suggests home address (edge)

7

Flexibility and efficiency of use

Shortcuts — hidden from novice users — may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

from: expected dropdown list

time selection

can't see return times

is not taking account of return

8

Aesthetic and minimalist design

Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.

9

Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.

website doesnt work

rubbish error checking

10

Help and documentation

It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.

offering to edit your personal information (edge problem)

11

Bugs

These problems cannot be classified as usability problems as such, even though bugs obviously do impact usability negatively

is stuck on the departure selection

conceptual model of how to buy a ticket

two different dialogue styles: microsoft and flybussen. Is confusing

TRYING TO UNDERSTAND WHY BUS IS LATE

website is not giving a clue as to where to find information

defaults to "plan your trip"

I have to know which bus company I have booked with

English site is showing in norwegian

Found the live map

where is the airport??

map is very small. Not gonna see that on the phone

not allowing me to enter a location

what are the red and the blue buses?

need to know a lot about the bus system to understand the map

no predictive time of arrival

they expect you to read the instruction

1

Visibility of system status

The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.

doesn't say when checkout actually occurs

can't see if I. have picket departure or arrival on front page

Need to pick a specific bus (did not see info on top)

map was only a picture

2

Match between system and the real world

The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.

map is not helpful for finding a specific bus

3

User control and freedom

Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.

cannot type å

already selected time so expected to be able to take any bus or get a specific suggestion

4

Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.

prices at checkout are in norwegian

map info in norwegian

5

Error prevention

Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action.

6

Recognition rather than recall

Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.

7

Flexibility and efficiency of use

Shortcuts — hidden from novice users — may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8

Aesthetic and minimalist design

Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.

date picker is not appropriate for wide screens

9

Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.

10

Help and documentation

It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.

don't know what fb5 means

info icon doesn't work

11

Bugs

These problems cannot be classified as usability problems as such, even though bugs obviously do impact usability negatively

lacks a desktop version. Things are too wide and far apart

would like to have table available digitally

Customer service: all sections are presented the same way regardless of importance

phone number is hidden

plan your trip is actually just ticket information

time picker should show current time

can find bus table digitally but it's presented as a purchase feature

select departure should have an indicator saying if the bus is on time

1

Visibility of system status

The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.

little info in home page

unexpected into on home page

2

Match between system and the real world

The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.

ticket information instead of letting me buy a ticket

prices without currency marker

the bus should have had a time estimate

3

User control and freedom

Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.

asks for location to find stops on map

"Where is the bus?" is very hidden

4

Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.

couldnt identify language setting on a smaller screen

not everything on the page is translated

inconsistencies in languages

back to norwegian when I go back to edit

translation issues on payment page

pins are marked but they are not translated

functionalities not existing in both languages

5

Error prevention

Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action.

"are you sure?"

6

Recognition rather than recall

Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.

going back to English results in informatino being deleted

main meny shows no information on how to buy a ticket

7

Flexibility and efficiency of use

Shortcuts — hidden from novice users — may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

wall of text on certain pages

should be a quick way to get live updates (traffic into and real-time maps)

"Where is the bus?" is very hidden

8

Aesthetic and minimalist design

Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.

wall of text on certain pages

information box has too much white space

9

Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.

blank list of traffic announcement

10

Help and documentation

It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.

phone number a little bit hidden

no way to say that you havent found what youre looking for in the QA

11

Bugs

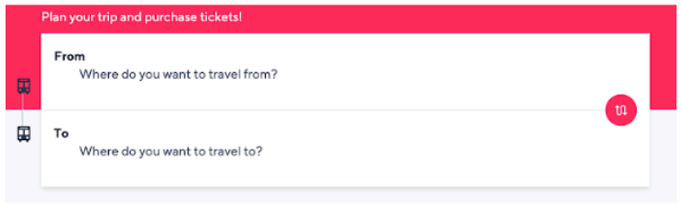
These problems cannot be classified as usability problems as such, even though bugs obviously do impact usability negatively

Appendix 2

1. There is no way to search for places or bus stops containing the letters Æ, Ø or Å.

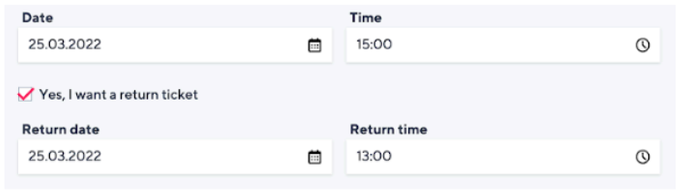
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

2. Expected a dropdown list or active input field when clicking the first input field instead of being sent to a separate page.



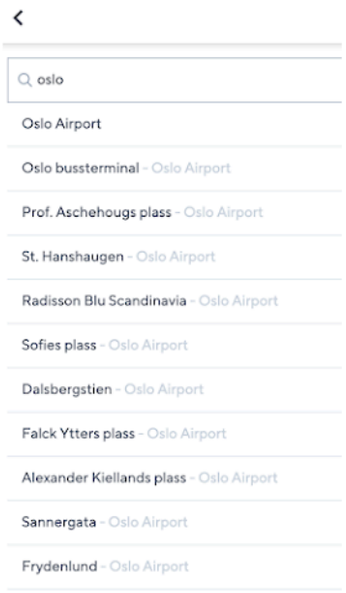
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

3. The suggested return time is sometimes earlier than the selected departure time.



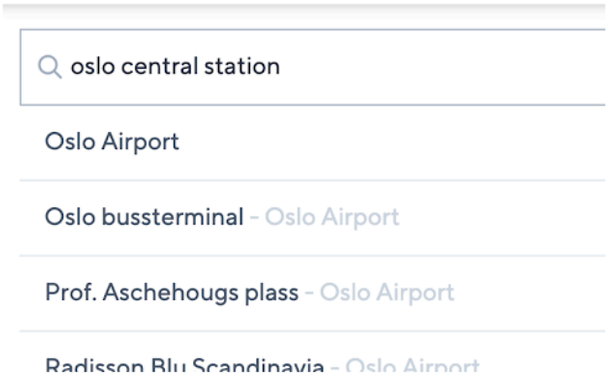
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

4. It's unclear what the grey text means, making it look like there are many Oslo Airports or many stops near Oslo Airport (in the example below).



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

5. Searching for "Oslo Central Station" gives "Oslo Airport" (incorrect) as the first suggestion and "Oslo bussterminal" (correct) as the second suggestion.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

6. There are too many separate screens, causing the user to lose context.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

7. The user has to select travel times twice: First on the front page, then once more after clicking the "Find trip and price" button.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

8. On the "Airports -> Oslo Airport" page the user has to know and remember which bus company they have booked with.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

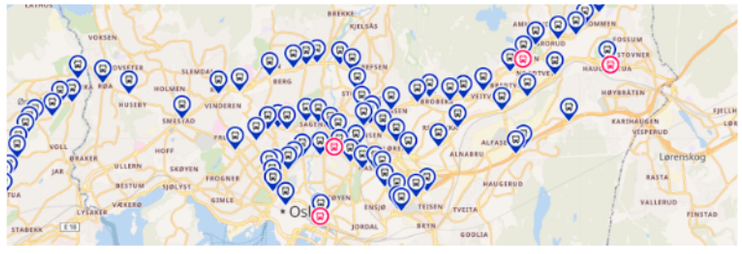
9. Norwegian text on English page ("Gå til side").

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

10. The website is not giving a clue as to where to find information about why the bus is late.

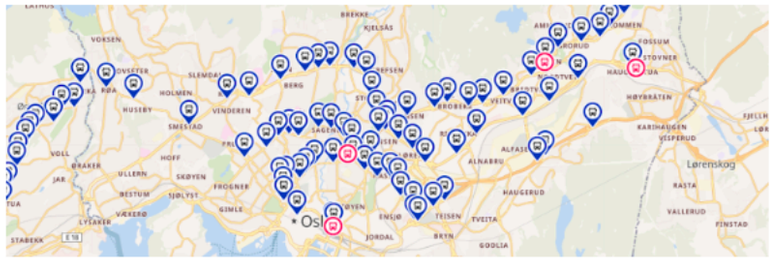
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

11. Bus stops are not searchable on the real-time map.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

12. Buses are not searchable on the real-time map.

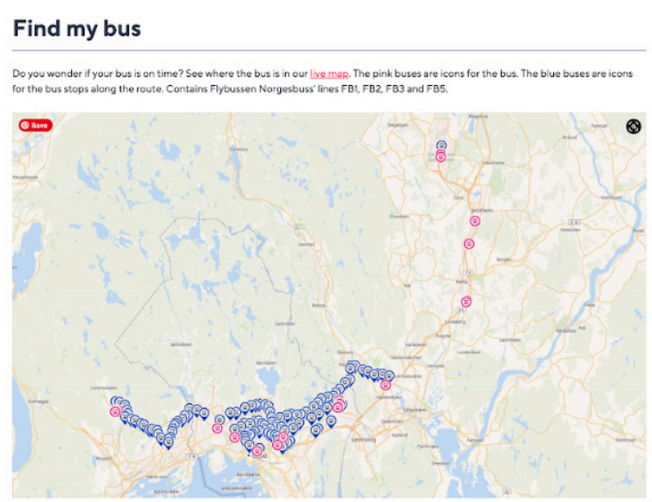


- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

13. The real-time map is too small for a phone.

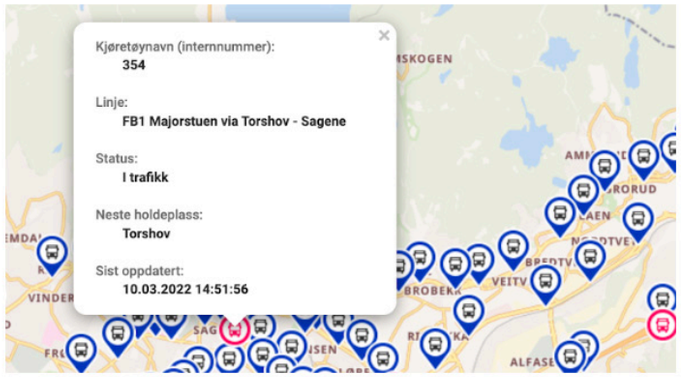
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

14. The image below looks like a real-time map but is only a static image.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

15. There is no predicted time of arrival for the buses on the real-time map



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

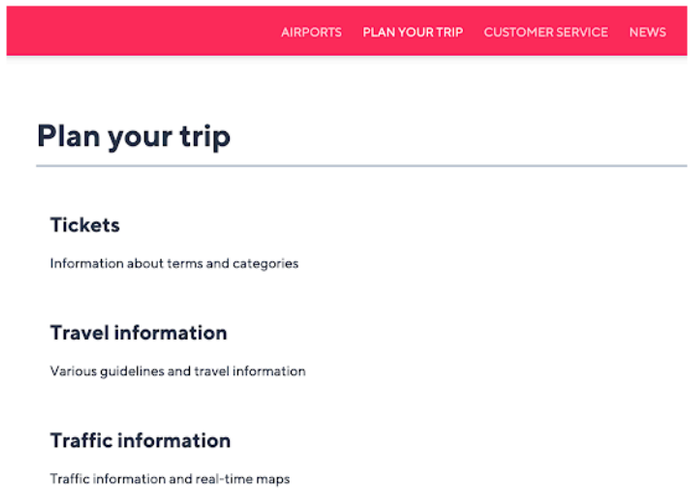
16. Info on the real-time map has not been translated to English.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

17. It's unclear how the blue and red pins are different unless the user has read the instructions before entering the map.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

18. The naming of the "Plan your trip" page communicates poorly what the page actually is.

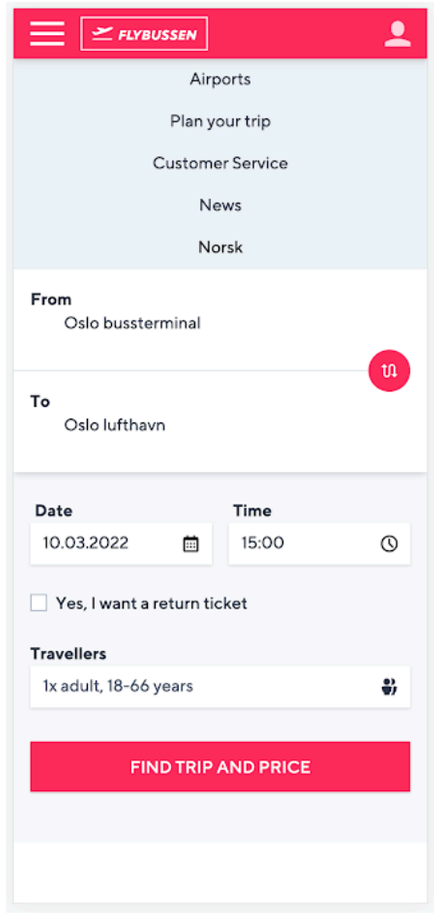


- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

19. There is no info on which currency the price is in.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

20. It's difficult to identify the language setting on small screens.

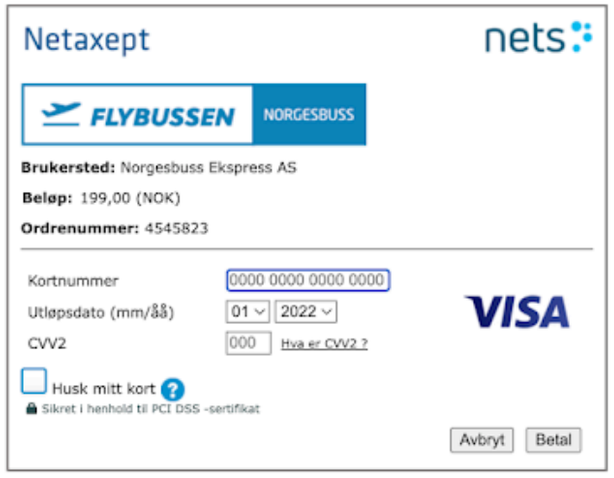


- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

21. The edit (pencil) button changes the language back to Norwegian.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

22. Payment page is not available in English



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

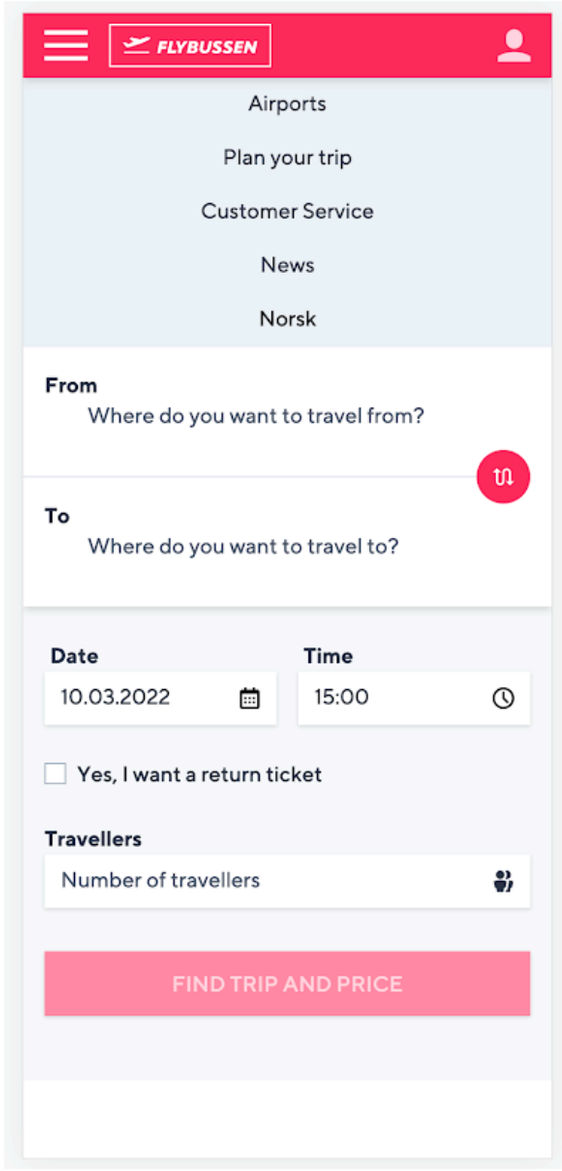
23. Some links are not available in both languages. (Only one of the links in the image below exists on the English page)

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

24. Input information is deleted when switching between languages.

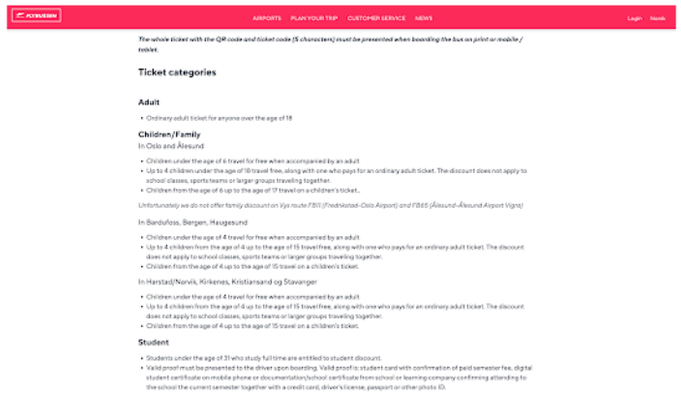
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

25. The hamburger menu (on small screens) shows no info on how to buy a ticket.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

26. Some pages are just a wall of text.

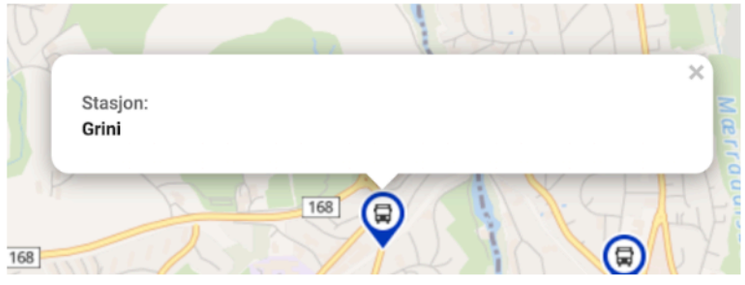


- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

27. The website is missing a quick way to find live updates.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

28. The info boxes have too much whitespace, which may be communicating that information has failed to load.

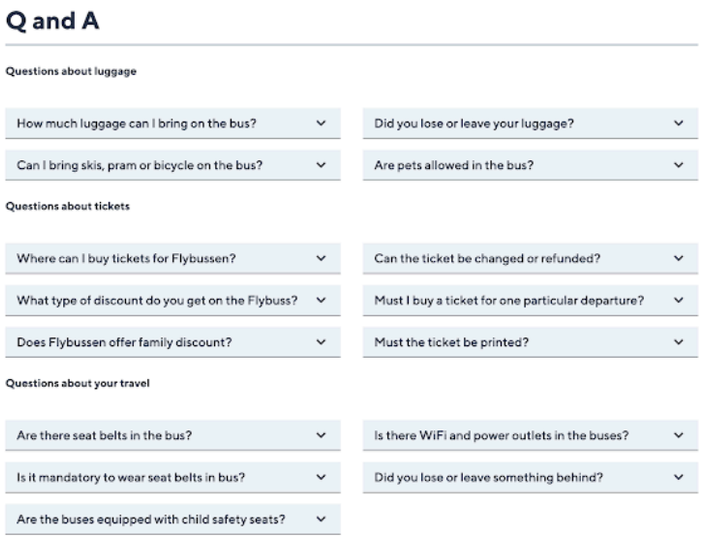


- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

29. The list of traffic announcements is blank, which may be communicating that information has failed to load.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

30. There is no direct way for the user to say that they haven't found what they are looking for in the Q&A.

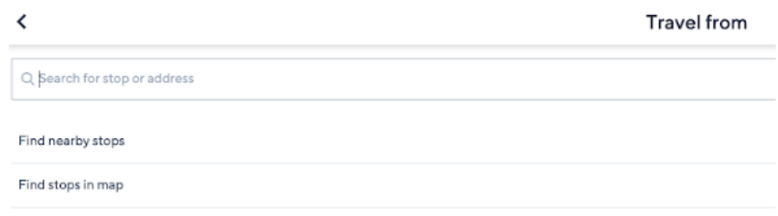


- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

31. Phone number is too hidden.

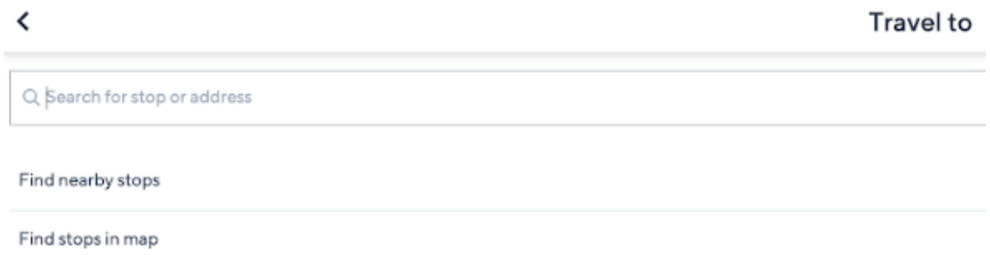
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

32. The "Find nearby stops" feature only shows a blank page.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

33. The "Find stops in map" feature shows a blank page if location sharing is turned off.

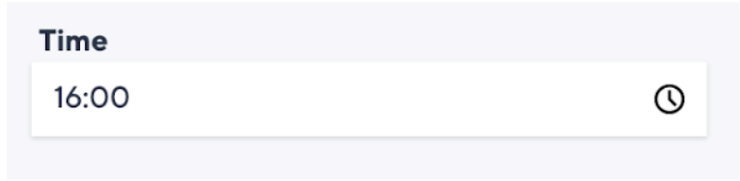


- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

34. When the screen is blank, it is completely up to the user to troubleshoot the problem.

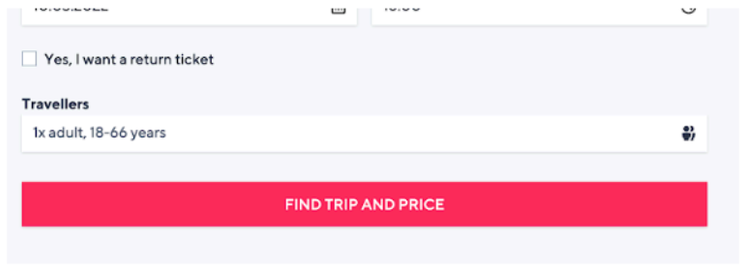
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

35. Unclear if travel time on the home page is selected for arrival or departure.



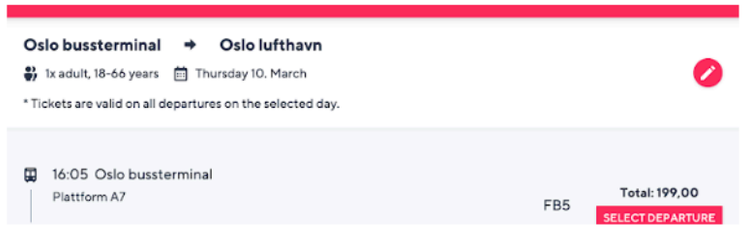
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

36. The "Find trip and price" button is hidden on smaller screens until the user scrolls down.



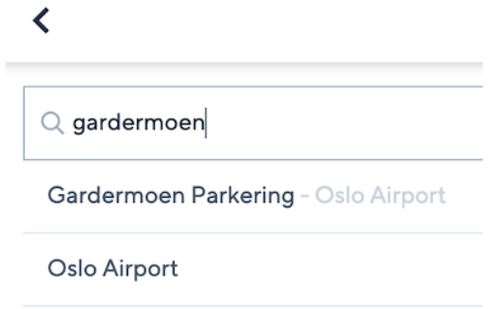
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

37. The edit feature is tedious if all the user wants is to see earlier departures.



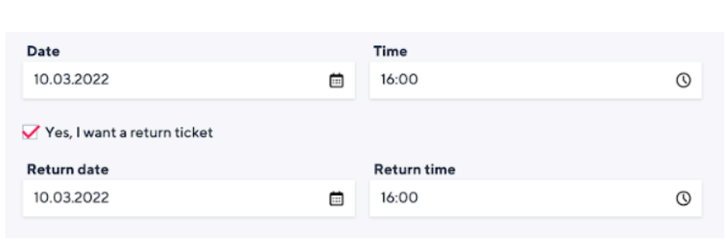
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

38. It's unclear from the search suggestions that "Oslo Airport" is the same airport as "Gardermoen"



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

39. Departure/arrival time can be confused with when the plane lands



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

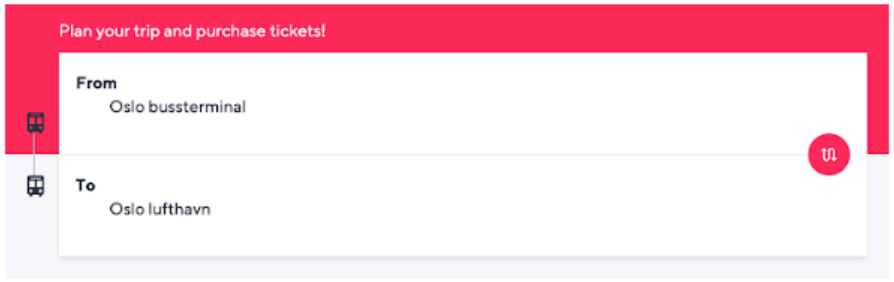
40. It's unclear to the user what FB5 means.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

41. Absolutely nothing happens when the user searches for places that don't exist in the database.

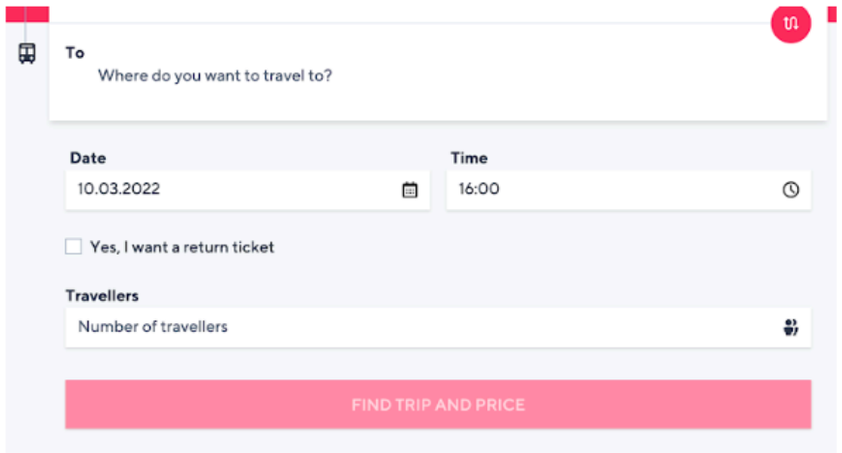
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

42. It's not possible to delete text inputs.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

43. The contrast is too low on the "Find trip and prices" button.

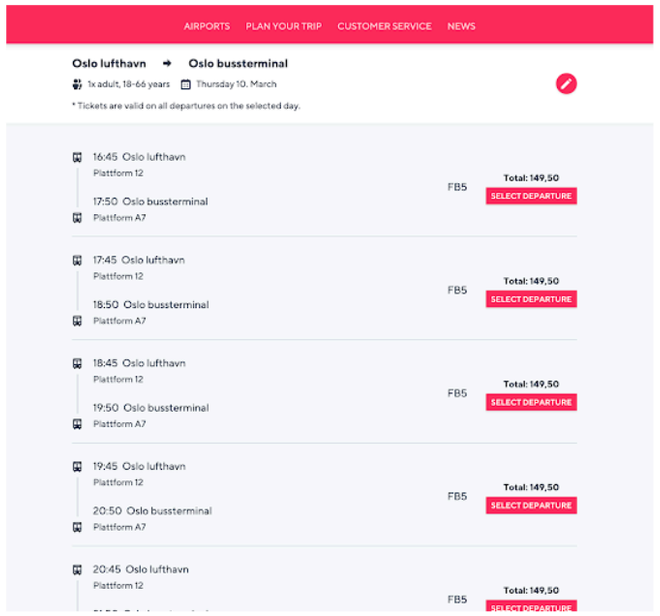


- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

44. Lack of feedback when a departure is selected.

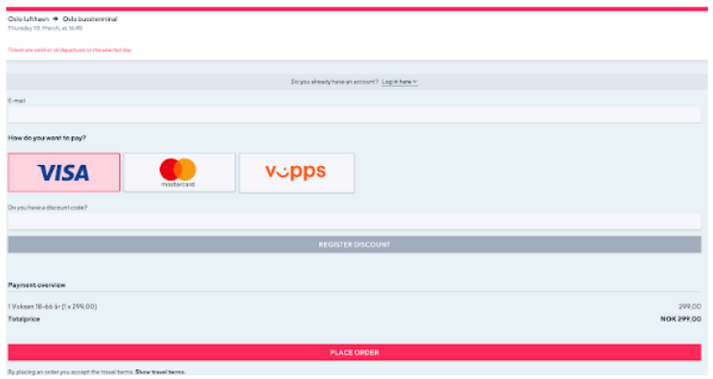
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

45. If the user wants a return ticket the layout on the two departure selection pages are almost identical.



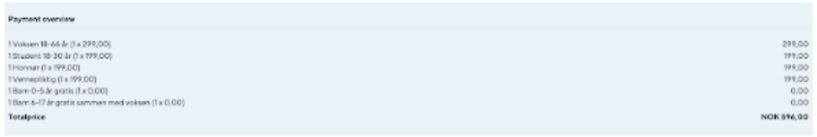
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

46. Many of the input fields, buttons etc are too wide resulting in a lot of scrolling that would have been avoided with better use of space.



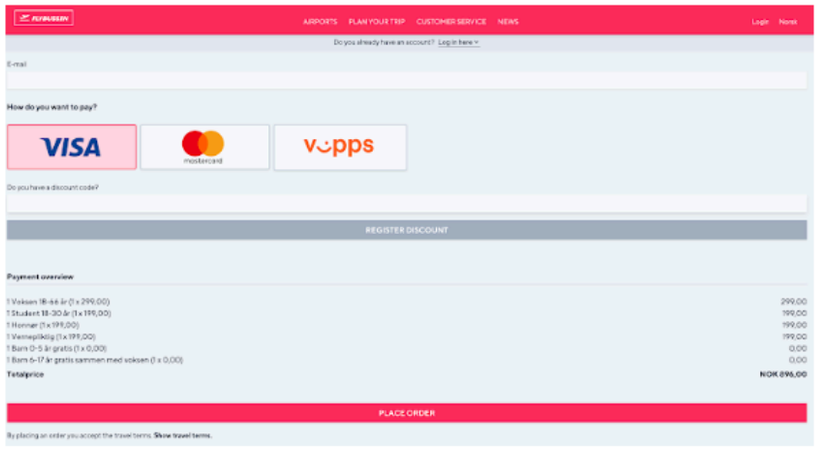
- I don't agree that this is a usability problem at all
- Cosmetic problem only – need not be fixed unless extra time is available on project
- Minor usability problem – fixing this should be given low priority
- Major usability problem – important to fix, so should be given high priority

47. The text is too far apart making the prices to the right hard to spot and hard to see which text is on the same line.



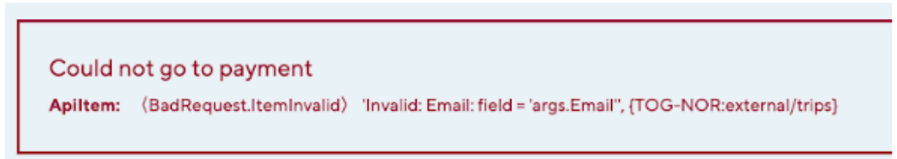
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

48. The "Place order" button is hidden at the bottom of the page outside the initial viewport.



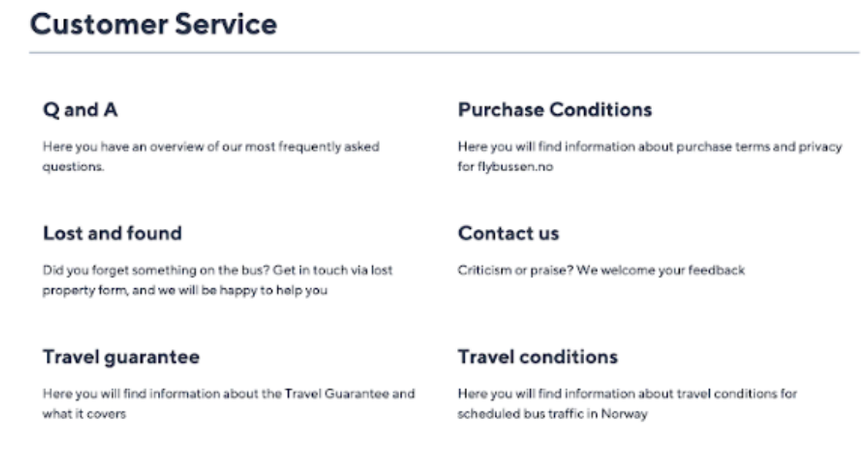
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

49. Bad error message when the user enters an invalid email at the checkout page.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

50. The different sections on the "Customer service" page are not ordered according to importance.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

51. There are news on the "Airports" page under most airports even though there is a separate news page.

Bardufoss Airport

News

Buy your ticket online and save time and money!

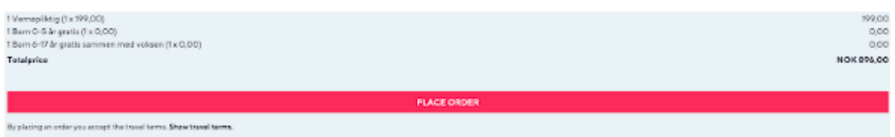
Purchase your ticket in advance on flybussen.no and save both time and money. If you buy a return ticket you save even more!

Travel discounts on Flybussen Bardufoss

Are you a student, travelling with children, or a senior? Flybussen Bardufoss offers favourable discounts.

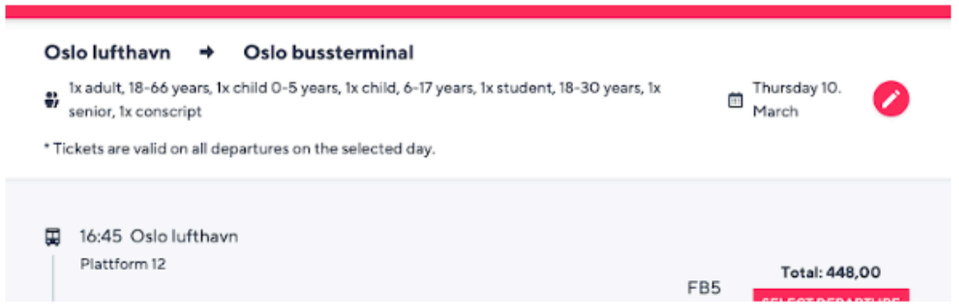
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

52. There is no available info on when the user is charged after clicking "Place order."



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

53. Info about tickets being valid all day is too difficult to spot.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

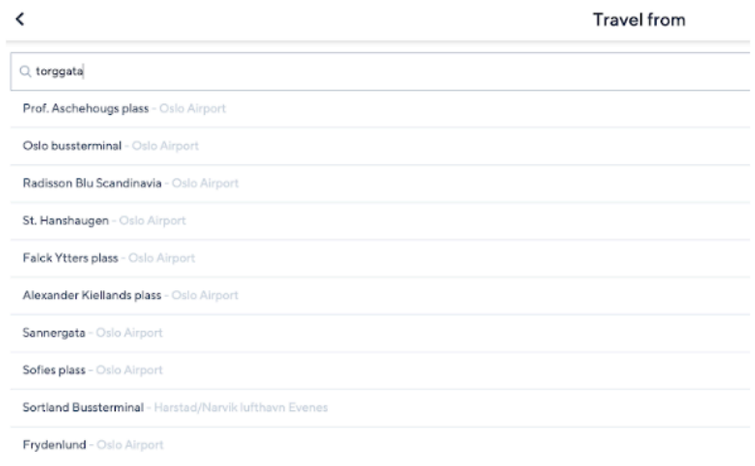
54. The prices at checkout are only available in Norwegian.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

55. Info icon is not functional.

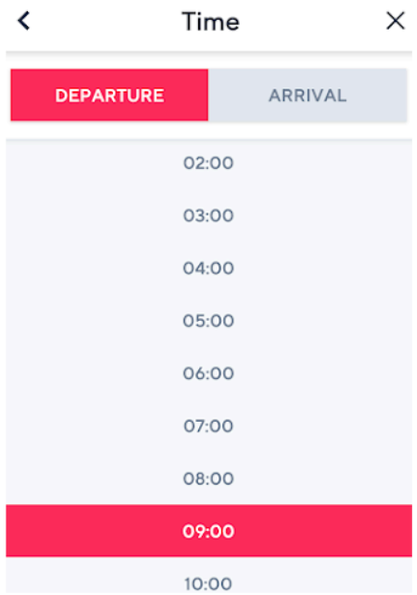
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

56. It's not clear how the search suggestions are sorted, making it unclear which bus stop is closest in distance.



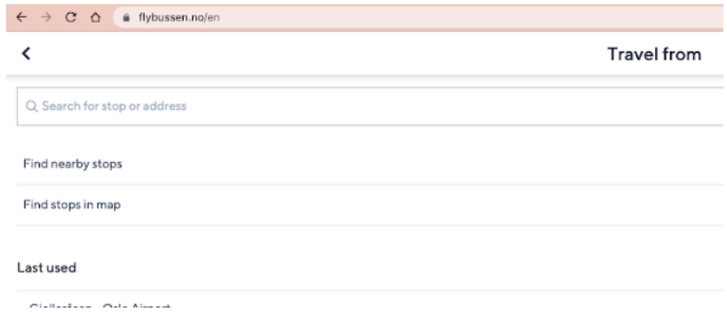
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

57. The website does not remember the user selecting "Arrivals" instead of "Departures," and will then only show bus journeys departing AFTER the user wants to arrive.



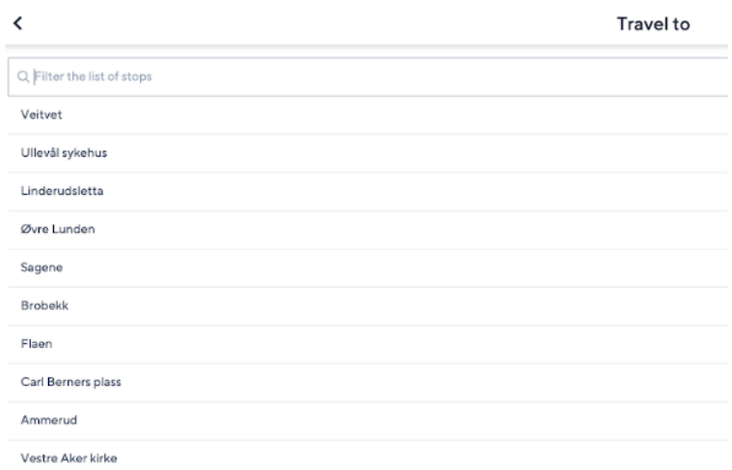
- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

58. The browser back button does not take the user to the previous page from any of the input type of pages.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

59. The "find stops in map" features are not always available.



- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

60. There is no way to check prices without going all the way to the checkout page.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

61. It's not possible to choose travellers before choosing places to travel to and from.

- ◯ I don't agree that this is a usability problem at all
- ◯ Cosmetic problem only – need not be fixed unless extra time is available on project
- ◯ Minor usability problem – fixing this should be given low priority
- ◯ Major usability problem – important to fix, so should be given high priority

Do you have any feedback/comments/ideas/thoughts regarding the method, survey, usability problems or my project?

Appendix 3

57. The website does not remember the user selecting "Arrivals" instead of "Departures," and will then only show bus journeys departing AFTER the user wants to arrive.

1. There is no way to search for places or bus stops containing the letters Æ, Ø or Å.

34. When the screen is blank, it is completely up to the user to troubleshoot the problem.

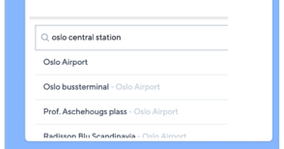
35. Unclear if travel time on the home page is selected for arrival or departure.

41. Absolutely nothing happens when the user searches for places that don't exist in the database.

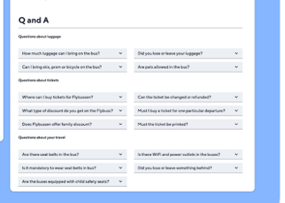
22. Payment page is not available in English

32. The "Find nearby stops" feature only shows a blank page.

5. Searching for "Oslo Central Station" gives "Oslo Airport" (incorrect) as the first suggestion and "Oslo bussterminal" (correct) as the second suggestion.



30. There is no direct way for the user to say that they haven't found what they are looking for in the Q&A.



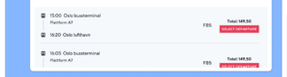
10. The website is not giving a clue as to where to find information about why the bus is late.

23. Some links are not available in both languages. (Only one of the links in the image below exists on the English page)

24. Input information is deleted when switching between languages.

42. It's not possible to delete text inputs.

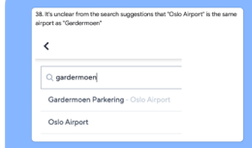
7. The user has to select travel times twice: First on the front page, then once more after clicking the "Find trip and price" button.



36. The "Find trip and price" button is hidden on smaller screens until the user scrolls down.

60. There is no way to check prices without going all the way to the checkout page.

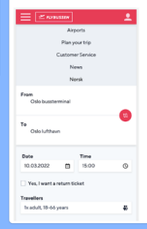
38. It's unclear from the search suggestions that "Oslo Airport" is the same airport as "Gardermoen"



44. Lack of feedback when a departure is selected.

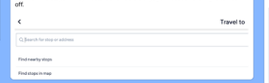
3. The suggested return time is sometimes earlier than the selected departure time.

20. It's difficult to identify the language setting on small screens.

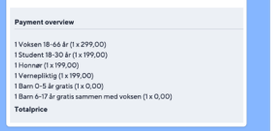


21. The edit (pencil) button changes the language back to Norwegian.

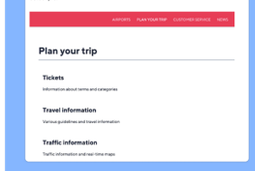
33. The "Find stops in map" feature shows a blank page if location sharing is turned off.



54. The prices at checkout are only available in Norwegian.



18. The naming of the "Plan your trip" page communicates poorly what the page actually is.



19. There is no info on which currency the price is in.

37. The edit feature is tedious if all the user wants is to see earlier departures.

49. Bad error message when the user enters an invalid email at the checkout page.

59. The "find stops in map" features are not always available.

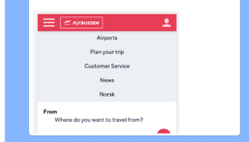
48. The "Place order" button is hidden at the bottom of the page outside the initial viewport.

9. Norwegian text on English page ("Gå til side").

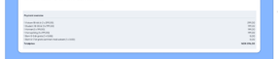
13. The real-time map is too small for a phone.

27. The website is missing a quick way to find live updates.

25. The hamburger menu (on small screens) shows no info on how to buy a ticket.



47. The text is too far apart making the prices to the right hard to spot and hard to see which text is on the same line.

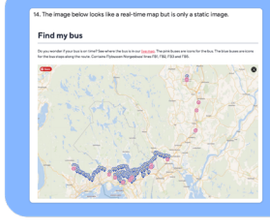


55. Info icon is not functional.

61. It's not possible to choose travellers before choosing places to travel to and from.

56. It's not clear how the search suggestions are sorted, making it unclear which bus stop is closest in distance.

14. The image below looks like a real-time map but is only a static image.



52. There is no available info on when the user is charged after clicking "Place order."

53. Info about tickets being valid all day is too difficult to spot.

Appendix 4

The introduction script below was based on a script by Steve Krug. By his own wishes, I will provide the link to the website where you can download the script file: https://sensible.com/download-files/

- chit chat
- my name is Elise and I'm writing my master's in Industrial Design, hence this usability test today.
- Before we start I have some information for you, and I'm going to read to make sure that I cover everything.
- I'm asking people to try using two versions of a website that I'm sort of working on so that I can see whether it works as intended. One is a prototype built in Figma, while the other is a fully implemented website.
- The entire session should take about 30 minutes.
- I want to make it very clear that I am testing the website/prototype and not you, so don't worry about hurting my feelings. I want to improve the website, so I need to hear your honest reactions.
- As you use the website/prototype, I want to you try to think out loud. Tell me about any assumptions you have, what you expect to be seeing, what you want to do etc.
- If you have any questions as we go along, just ask them, but I might not be able to answer them since I'm interested in seeing how you would do without my assistance.
- I'm not going to be recording anything, so I might ask you to pause what you are doing sometimes to write notes and ask you some follow up questions.
- That means I won't be collecting any personal information, so you'll be completely anonymous.
- Any questions before we start?