



Norwegian University of  
Science and Technology

# Digital feedback channels as platforms for solving accessibility problems

**Bjørg Andrea Lysbakken**

Master in Interaction Design

Submission date: June 2017

Supervisor: Frode Volden, ID

Co-supervisor: Ole E. Wattne, ID

Norwegian University of Science and Technology  
Department of Design



## **Preface**

This thesis marks the completion of my master in Interaction Design from NTNU in Gjøvik. The goal of the study is to provide new knowledge useful for the development towards more inclusive ICT.

I would like to thank MediaLT for encouraging the study, especially Morten Tollefsen, who provided initial directions for the thesis topic and was a helpful advisor during the process. Furthermore, I want to express my gratitude to my supervisors Frode Volden and Ole E. Wattne, for investing time and interest in the project. Their guidance throughout the course of working on the thesis has been of great value. I also wish to thank my family and boyfriend for all their support. Last but not least, I want to give big thanks to the three persons interviewed in the study, for their participation and willingness to share openly about personal experiences.

I hope this thesis can bring more attention to an important topic and inspire others to take an interest in how digital solutions can be improved to include a broader range of people, regardless of their personal abilities and needs.

**Bjørg Andrea Lysbakken**

Gjøvik, May 30, 2017

## **Abstract**

Society is facing continuous challenges regarding accessibility in information and communication technology (ICT), including mobile applications. Solutions fail to accommodate the diverse abilities and requirements of different users, depriving many people of possibilities to take full advantage of digital services and products.

Based on the present state of ICT there is a need to explore new methods for how to develop widely inclusive solutions. Digital feedback channels represent an opportunity for communication between end users and companies to solve issues with products and services. However, research seems to be lacking on the use of such channels for accessibility purposes.

This thesis investigates how effectively two digital feedback channels, emails and App Store reviews, function as platforms for solving accessibility issues in mobile applications. More specifically, the study focuses on problems for blind persons using the screen reader VoiceOver to interact with mobile content. A field experiment was conducted, in which a sample of iOS applications from Norwegian top lists was evaluated for the purpose of finding accessibility problems. Feedback concerning 50 problems in 50 separate applications was sent in via emails or App Store reviews. The applications were re-tested after 30 days to document the effect. As part of interviews performed, the study also looked into habits, preferences, and experiences of persons in the user group, with regards to sending feedback about accessibility problems through digital channels.

Findings from the field experiment indicate that sending feedback through emails and App Store reviews currently has a minimal effect on solving VoiceOver related accessibility problems in mobile applications. As the sample was limited to a selection from the most popular iOS applications in

Norway within certain categories, the results are considered likely to be more generalizable to this specific subpopulation than to the general population of mobile applications. Factors such as time, type and severity of problems and feedback formats, formulation and quantity could potentially influence the results and therefore be subject to further research. Additional findings from the interviews suggest that there are challenges which should be further investigated regarding how persons in the user group experience to employ digital feedback channels for communication about accessibility problems.

Collaboration between end users and companies through digital feedback channels such as emails and App Store reviews could be part of the solution for achieving increased accessibility in mobile apps and other ICT solutions. Further studies should be done on the potential of using these channels for such purposes, and how to utilize this potential in more effective ways.

# Contents

<b>Preface</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Contents</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Topics covered by the project	1
1.2 Keywords	1
1.3 Problem description	1
1.4 Justification, motivation and benefits	2
1.5 Research questions	3
1.5.1 Explanation of terms	3
1.6 Contributions	4
<b>2 Theory and background</b>	<b>6</b>
2.1 User diversity	6
2.2 Assistive devices and technologies	7
2.2.1 Screen readers	7
2.3 Accessibility	8
2.3.1 ICT accessibility	8
2.3.2 Mobile accessibility	10
2.3.3 Methods for accessibility evaluation	12
2.4 Digital feedback channels	14
<b>3 Research design and methodology</b>	<b>16</b>
3.1 Evaluation of applications	16
3.1.1 Inclusion criteria	16
3.1.2 User group	17
3.1.3 Sampling and assignment strategies	17
3.1.4 Evaluation method	18
3.1.5 Trial evaluations	21
3.2 Reporting accessibility problems	22

3.2.1	Feedback channels	22
3.2.2	Reporting cycle	23
3.2.3	Feedback format	23
3.3	Documentation of response and post-evaluations	26
3.4	Notice	27
3.5	Interviews	27
3.5.1	Participant profiles	27
3.5.2	Interview type	27
3.5.3	Informed consent	28
3.5.4	Anonymity and confidentiality	28
<b>4</b>	<b>Results</b>	<b>29</b>
4.1	Accessibility problems	29
4.2	Effect of sending feedback	34
4.3	Insights from interviews	36
4.3.1	Accessibility problems encountered	36
4.3.2	Feedback habits and preferences	38
4.3.3	Feedback effect	41
4.3.4	User experiences	43
<b>5</b>	<b>Discussion</b>	<b>46</b>
5.1	VoiceOver related accessibility problems in mobile apps	46
5.1.1	Types of problems	46
5.1.2	Problem occurrence	47
5.1.3	Screening technique versus evaluations with real users	47
5.1.4	Defining accessibility	49
5.1.5	Status of accessibility in applications	49
5.1.6	Differences in problem perception with and without visual support	50
5.2	Digital feedback channels as platforms for solving accessibility problems	51
5.2.1	Feedback effect	51
5.2.2	Time	54

5.2.3 Type and severity of problems	54
5.2.4 Feedback formulation	55
5.2.5 Feedback formats	56
5.2.6 Feedback quantity	57
5.2.7 Channels	57
5.3 Reasons behind the seemingly low effectiveness	58
5.4 User perspectives	59
5.4.1 Feedback habits and preferences	60
5.4.2 User experiences	61
5.5 Implications of the results	62
<b>6 Conclusions</b>	<b>64</b>
<b>Bibliography</b>	<b>66</b>
<b>Appendix: Interview guide (translated to English)</b>	<b>71</b>



As stated by the World Health Organization (1), ICT accessibility is needed in order for people to participate fully in society.

How can we ensure this?

# 1 Introduction

The following chapter provides an introduction to the main topics covered by the thesis, the problem area of study and the motivation behind the project. Furthermore, it will present the research questions and intended contributions.

## 1.1 Topics covered by the project

This thesis studies issues related to accessibility within the field of information and communication technology (ICT), more specifically mobile applications. The primary focus is on methods for discovering and communicating accessibility issues. The project takes a closer look at how digital feedback channels, such as App Store reviews and emails, can be used to address accessibility problems in ICT solutions. In this particular study, attention is on accessibility for persons who are blind, relating to the use of assistive technology.

## 1.2 Keywords

Inclusive design, Accessibility, User feedback, Customer reviews, Online reviews, Email, App Store, ICT, Mobile applications, Smartphone, Apple iOS, Blind users, Assistive technology, Screen readers, VoiceOver

## 1.3 Problem description

In recent years there has seemingly been an increasing focus on accessibility within the field of ICT. Yet, literature suggests that we have a way to go when it comes to accessibility of digital products and services, including mobile applications (2-5). Solutions fail to accommodate the diverse abilities and requirements of different users, depriving many people of possibilities to take full advantage of digital services and products.

Based on the present state of ICT there is a need to explore new methods for how to develop widely inclusive solutions (6). Digital feedback channels represent an opportunity for communication between end users and companies to solve issues with products and services. These channels give users platforms for notifying companies about problems encountered, and companies the chance to receive feedback from a broad range of users. A number of studies revolve around user feedback in digital channels. However, research seems to be lacking on the use of such solutions for accessibility purposes.

To investigate how effectively digital feedback channels currently work for communicating and solving accessibility problems in mobile applications, a field experiment was conducted, along with interviews. Two channels, App Store reviews and emails, were studied in the field experiment. The research focused on accessibility for persons who are blind and are using the screen reader VoiceOver to interact with mobile content.

#### **1.4 Justification, motivation and benefits**

The importance of developing more inclusive ICT increases as we are becoming more and more dependent on ICT-based products and services to receive important information and perform tasks necessary for participating in society (1). Products and services related to education, transportation, finance, business and other essential areas are now offered through mobile applications and other digital solutions.

Building knowledge about the potential of using digital feedback channels as tools for communicating accessibility information can not only benefit users but also companies providing digital products and services. This can possibly be an efficient, cost-effective method for companies to get input from a wide range of users to improve accessibility of solutions. Increased accessibility will give more users the opportunity to employ desired products and services

without struggles. Making it possible for a higher number of people to use the solutions can profit companies by increasing their customer base.

## **1.5 Research questions**

The thesis explores the following research questions:

Main questions

1. Which effect does user feedback through emails and App Store reviews have on solving VoiceOver related accessibility problems in mobile applications?
2. What works most effectively, emails or App Store reviews?

Sub-question

3. What works most effectively, feedback through a public or a private channel?

As part of interviews conducted, the study also looked into habits, preferences, and experiences of persons using mobile applications via VoiceOver due to blindness, with regards to giving feedback about accessibility problems through digital channels.

### **1.5.1 Explanation of terms**

This section describes what key terms refers to in the context of the study.

Digital feedback channel: An option provided in a digital product or service which allows users to give a feedback, in the form of text, points, images, videos or other means.

Public feedback channel: A digital feedback channel where feedback given is visible for other users.

Private feedback channel: A digital feedback channel where feedback given is not visible for other users.

Mobile application: A computer program which can be installed on mobile devices such as smartphones and tablets.

VoiceOver: A screen reader built into iPhones and other Apple products (see further description in section 2.2.1)

VoiceOver related accessibility problem: An incident of information or functionality provided in a mobile application being inaccessible through VoiceOver via audio output or touch input (see further description in section 3.1.4).

Effect: A correction of the mobile application linked to the reported problem.

## **1.6 Contributions**

The thesis aims to contribute with:

- a. Insight into different types of VoiceOver related accessibility problems occurring in mobile applications.
- b. Indications of how results obtained from using a screening technique for accessibility evaluation compare with experiences of real users.
- c. Indications of how effectively the two digital feedback channels investigated, emails and App Store reviews, currently function for solving accessibility problems in mobile applications. There is seemingly no previous research on how user feedback in digital channels regarding accessibility issues is utilized by companies in the development of ICT solutions. This study will contribute with new

knowledge in this area.

- d. Insight into habits, preferences and experiences of end users who are blind, with regards to sending feedback in digital channels about VoiceOver related accessibility problems encountered in mobile applications.
- e. Directions for further research on the potential of using digital feedback channels in the development of accessible mobile applications.

## 2 Theory and background

This chapter will give an insight into key concepts and the current state of research relevant to the project. As part of the research process, literature searches were performed in the online databases SpringerLink, IEEE Xplore, ACM Digital Library, ScienceDirect and Sage Journals.

### 2.1 User diversity

Every individual has a personal mix of abilities, needs, habits and preferences which affect how we use products and services. Many people experience some form of disability. This is the case for over 1 billion people, around 15% of the world's population, according to the *World Report on Disability* (1), published by the World Health Organization (WHO) in 2011. WHO states that everyone is likely to experience disability either directly or have a family member who experiences difficulties in functioning at some point in life (7).

Disability is a complex, ambiguous, evolving concept. In recent decades, there has been a transition from an individual, medical perspective of disability, towards the understanding that people are disabled by society and environmental factors as well as by their bodies (1). The World Health Organization declare that "A person's environment has a huge impact on the experience and extent of disability. Inaccessible environments create disability by creating barriers to participation and inclusion" (1, p. 4).

The *International Classification of Functioning, Disability and Health* (ICF), WHO's framework for health and disability, categorize problems with human functioning in three interconnected areas: impairments, activity limitations and participation restriction (1). The ICF regards *disability* as difficulties encountered in any of these three areas, arising from the interaction between individuals with health conditions and contextual factors (personal and environmental). Environmental factors include products, technology, natural

and built environments, support and relationships, attitudes, services, systems and policies. Motivation and self-esteem are examples of what the framework suggests as personal factors. WHO points out that based on the ICT definition, disability is a result of interaction, not a personal attribute. The organization further explain that the framework does not categorize people with disabilities as a separate group, but views disability as a continuum where it is not a matter of *yes or no*, but rather *more or less*.

According to the *World Report on Disability* (1), people's experience of disability can vary greatly. As stated in the report, disability encompasses a wide specter of different health conditions which can be visible or invisible, temporary or long term, static, episodic or degenerating, painful or inconsequential. Impairments can be sensory, physical, mental, or intellectual. WHO declare that generalizations about "people with disabilities" can be misleading, as "persons with disabilities have diverse personal factors with differences in gender, age, socioeconomic status, sexuality, ethnicity, or cultural heritage" (1, p. 8). The organization stresses that two individuals with the same impairment might have very different experiences and needs.

## **2.2 Assistive devices and technologies**

The primary purpose of so-called *assistive* devices and technologies is to "maintain or improve an individual's functioning and independence to facilitate participation and to enhance overall well-being" (8). Examples of such devices and technologies include wheelchairs, hearing aids, visual aids, prostheses, and computer software and hardware that increase mobility, hearing, vision, or communication capacities (8). The World Health Organization (9) states that over 1 billion people in the world need 1 or more assistive products.

### **2.2.1 Screen readers**



Screen readers are a form of assistive technology. These are software programs that convey content on the screen through a speech synthesizer or refreshable braille display so that users can hear the information spoken aloud, or read it tactually (10). Through functionality provided by the software, users can also navigate around the screen (10). This technology enables people who are blind or have visual impairments to access and interact with content in electronic products and services without being dependent on visual representations. Newer Android and iOS devices have screen readers built into the operating systems. Android's screen reader program is called TalkBack, and the iOS version VoiceOver.

## **2.3 Accessibility**

*Accessibility* is about giving everyone equal access, whether it is to physical environments, transportation, information and communication, including information and communication technology, or other aspects of society (11, 12). The focus is commonly on access for people who experience disabilities (13). However, it is argued that accessible solutions can also benefit others (11, 14). Accessibility is covered in *The Standard Rules on the Equalization of Opportunities for Persons with Disabilities* and the *Convention on the Rights of Persons with Disabilities* by the United Nations (12, 15). Accessibility is closely related to concepts such as *usability*, *universal design*, *inclusive design* and *design for all* (13).

### **2.3.1 ICT accessibility**

*Electronic accessibility*, or *E-accessibility*, is defined by the World Health Organization as “the ease of use of information and communication technologies (ICTs), such as the Internet, by people with disabilities” (16). W3C propose a more concrete definition, relating to accessibility on the Web: “Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web” (17).

Developments in the field of information and communication technologies in past decades offers new opportunities for participation and inclusion.

According to the World Health Organization

Accessing general information online enables people with disabilities to overcome any potential physical, communication and transport barriers in accessing other sources of information. ICT accessibility is therefore needed for people to participate fully in society. (1, p. 183-184)

Access to information and communication is included in the United Nations *Standard Rules on the Equalization of Opportunities for Persons with Disabilities* (15). The *World Report on Disability* (1) states that accessibility within ICT covers the design and supply of ICT products and services, including computers, telephones, telephony and television. According to the report, ICT accessibility relates to both technology (like control and navigation) and content (like sounds, images, and language produced and delivered by the technology).

An extensive amount of initiatives and studies regarding accessibility of ICT have emerged in recent years, confirming that there are still many issues that should be addressed in order to utilize the opportunities technology provides for increased inclusion. Inaccessibility of information and communication technology is one of the topics addressed in the *World Report on Disability* (1). According to the report ICT devices and systems are often incompatible with assistive devices and technologies, such as hearing aids or screen readers. The World Health Organization stresses that

Given the wide spectrum of ICT products, services, and sectors (commerce, health, education, and so on) a multisectoral and multi-stakeholder approach is required to ensure accessible ICT.

Governments, industry and end-users all have a role in increasing accessibility. (1, p. 186)

### **2.3.2 Mobile accessibility**

E-accessibility includes access to applications and other mobile content. According to W3C, *mobile accessibility* refers to “making websites and applications more accessible to people with disabilities when they are using mobile phones and other devices” (18).

Recent years have shown an increasing importance of mobile devices and apps. Currently, over 5 million apps are available in Apple’s App Store and the Google Play Store combined, and as of September 2016, over 140 billion apps were downloaded from the App Store (19, 20). There are apps available within an extensive list of categories, including business, education, travel, health and fitness, food and drink, music, sports, photo and video, finance, news, social networking and entertainment (21). As previously discussed, the World Report on Disability (1) suggests that more than 15% of the world’s population are experiencing a disability, which means that this will be the case for a big number of app users. Companies such as Google and Apple provide guidelines for helping developers and designers in creating accessible mobile applications (22, 23).

Several studies published in previous years have investigated the current state of mobile and app accessibility (2-4, 24-27). Serra et al. (3) evaluated the accessibility of e-government mobile applications in Brazil using an adapted version of the Web Content Accessibility Guidelines (WCAG) 2.0. Results showed that several elementary accessibility problems were encountered. Based on the findings, the researchers stress the importance of further research in accessibility design and evaluation of mobile apps in order to provide more inclusive access.

Park et al. (4) investigated the needs of persons with visual impairments when accessing mobile content and applications. Four participants were observed while performing tasks on a smartphone and thereafter interviewed. According to the researchers, serious accessibility problems were found, mainly caused by the speed of typing text and inaccessible application design, particularly when the VoiceOver function was used. The researchers propose a set of 10 heuristics for developing accessible mobile applications and underline the need for international guidelines and standards to improve the current mobile environment for people with disabilities.

Leporini et al. (25) studied blind users' interaction with Apple touchscreen devices (iPad, iPhone, and iPod) using VoiceOver. Feedback from 55 blind users was collected via an online survey, and the user interfaces of the devices were evaluated through usability inspections. Based on the results, researchers conclude that Apple devices are basically accessible to blind users through VoiceOver, but that there are issues related to usability.

Wentz et al. (2) conducted a survey to explore the impact of website, software, and mobile app updates on accessibility for blind users. 61% of the respondents reported that they had experienced an accessibility problem after updating a mobile application. Problems described by respondents included that certain parts of applications would not work, buttons were no longer accessible, the screen reader would lose focus on the app, new button labels were not understandable, and that apps that used to be accessible no longer was. The researchers conclude that social inclusion and inclusive design requires way different results than what is currently the reality, and declare that a focus on maintaining sustainable accessibility and usability should be included in future research on inclusive design.

Several other studies have also been conducted, which are not concerned with evaluating the current state of app accessibility, but rather the development of accessible applications, including research on frameworks

and guidelines. The studies discovered through the literature search only cover a certain, often small, selection of apps. More research is apparently needed to get a broader picture of the present state of accessibility in mobile applications. One of the contributions of this thesis is more knowledge about accessibility issues currently occurring in mobile applications.

### **2.3.3 Methods for accessibility evaluation**

A number of methods are proposed and used in accessibility development to evaluate how solutions conform to user requirements. Examples of common methods include *user testing*, *screening techniques*, *subjective assessments* and *conformance/expert reviews* (28). Accessibility evaluations can be conducted manually or by using automatic tools.

Some methods, such as conformance reviews and screening techniques, rely on experts (as opposed to users) to evaluate the state of accessibility. In conformance reviews, solutions are compared against a set of criteria or guidelines. General guidelines, like the *Web Content Accessibility Guidelines* (WCAG) (29), are supposed to cover requirements for a broad range of user groups. However, studies imply that far from all problems experienced by users are covered by guidelines such as the WCAG (30-32). Results from an empirical study of problems encountered by blind users on the Web showed that only 50.4% of the problems experienced by users were covered by success criteria in the Web Content Accessibility Guidelines 2.0 (31).

Several studies and sources highlight the importance of involving real users in accessibility audits to discover important problems and needs (1, 3, 7, 30, 31, 33, 34). The World Health Organization states that “People with disability have unique insights about their disability and situation but have been excluded from the decision-making process about issues that directly affect their lives” and therefore that “persons with disabilities through their representative organizations should be fully consulted and actively involved in all stages of formulating and implementing policies, laws, and services that

relate to them” (7, p. 4). The World Wide Web Consortium (33) claim that while conformance to accessibility standards is important, including users with disabilities in evaluations is beneficial in order to understand how a solution really works for users and to identify issues that are not discovered by conformance evaluation alone. As a counterclaim to the appropriateness of user involvement in accessibility evaluation, Brajnik (28) declare that performing user tests is likely to result in a number of usability problems general for all users, which are irrelevant with respect to accessibility for persons experiencing disabilities.

One of the drawbacks of user testing is that it can be resource demanding and costly compared to other methods, which can limit companies in the number of test sessions conducted and users included (6, 28, 34). However, it is not likely that the essential requirements within all different user groups will be covered by testing only with a small number of people, especially not in the case of developing software such as mobile applications which are intended for a wide range of potential users (35).

The use of several evaluation methods could be a solution for covering a broad range of user needs. However, Baazeem et al. (6) state that combining the traditional methods can be impractical in the long run, due to time, expenses and other constraints, and that as a result, accessibility issues still increase over time, since the solutions are not checked and updated regularly.

According to Baazeem et al. (6), there has been a lack of significant evolution of accessibility evaluation methods in recent years. Based on a study on the advancements in web accessibility evaluation methods for the past years, the researchers conclude that there is a need for further investigation and development of accessibility evaluation methods.

## 2.4 Digital feedback channels

Today end users have the opportunity to voice their opinion about ICT products and services through digital feedback channels such as those provided by online application distribution platforms (app stores), social media (including Facebook and Twitter) and websites. Among other possibilities, users can choose to give feedback through sending an email, posting on the wall of a Facebook group, writing to companies through the Messenger chat or submitting a review in App Store. Feedback can come in the form of ratings or comments using text, images, videos, sound clips, and links.

App store reviews and other digital feedback systems give users channels for notifying companies about user experiences, bugs and desired features and they present businesses with the option to gather feedback from a big number of users. According to Panichella et al. (36), app stores provide useful electronic platforms for productive exchange of app information between users and developers. Palomba et al. (37) state that app store reviews are free and fast mechanisms for crowd feedback that can be used in the development process, and that the reviews can describe various issues exhibited in diverse combinations of devices, screen sizes, operating systems and network conditions, which might not be discovered during development and testing activities.

A number of studies published in previous years revolve around user feedback in digital platforms, mostly regarding reviews in app stores (35-43). Pagano et al. (38) investigated feedback habits and content as well as its impact on the user community, by analyzing over a million user reviews from the Apple App Store. Results showed that the majority of feedback is provided shortly after new releases, with a quick decrease in frequency over time. It was found that feedback can contain multiple topics and that the quality and constructiveness vary widely. The researchers state that part of the feedback is superficial, while other feedback includes useful comments,

bug reports, user experience and feature requests, which can help developers in understanding user needs and extending applications.

Literature suggests that companies can be presented with challenges in processing, analyzing and extracting valuable information from users' feedback messages, connected with the feedback volume and its quality (36, 38, 39, 41, 42). Panichella et al. (36) state that for the most popular apps, the large amount of feedback as well as its unstructured nature and varying quality can make it very challenging to identify useful feedback. According to Guzman et al. (39), the amount of reviews for many apps is too extensive for it to be processed manually. Maalej et al. (41) claim that although studies have shown that user reviews represent a rich source of information for app vendors and developers, the majority of reviews are rather non-informative. Several studies have been concerned with how to prioritize and extract information from user reviews (36, 39, 41-43).

There seems to be a lack of research with regards to communication in digital feedback platforms for accessibility purposes specifically. In a pilot study, Anam et al. (44) analyzed the contents of 173 app reviews from 25 applications, and propose a system to automatically detect accessibility information in reviews and test its polarity. However, to the author's knowledge, no research has yet been done on how digital feedback channels work as platforms for solving accessibility issues in mobile applications and other ICT products and services. The main aim of this thesis is therefore to investigate this further.



## **3 Research design and methodology**

A field experiment was conducted, along with interviews, to investigate how digital feedback channels currently work for addressing accessibility issues in mobile applications, and how end users experience to employ such solutions to communicate about accessibility problems encountered.

The study consisted of four main stages:

1. Evaluation of a mobile application sample for the purpose of finding accessibility problems
2. Reporting of accessibility problems in two separate feedback channels
3. Documentation and testing of the reporting effect
4. Interviews with persons from the user group in focus

### **3.1 Evaluation of applications**

A selection of iOS applications for iPhone was evaluated by the author for the purpose of finding accessibility problems. The problems discovered were later on reported through one of the two feedback channels studied, either reviews in App Store or emails, which will be further explained in section 3.2. The apps were sampled from a total of 1100 applications; the 50 most popular free iOS applications in Norway within 22 categories, based on App Store's top lists for Norway from the 2nd of February 2017.

#### **3.1.1 Inclusion criteria**

Out of the total of 25 categories provided in App Store, the following 22 categories were included: Books, Business, Catalogs, Education, Entertainment, Finance, Food & Drink, Health & Fitness, Lifestyle, Medical, Music, Navigation, News, Photo & Video, Productivity, Reference, Shopping, Social Networking, Sports, Travel, Utilities, and Weather. Game applications were excluded from the study for the reason that these were considered to

have such a different structure than other types of applications that it would require another accessibility evaluation method than the one used for evaluating the rest of the applications. The top list for the App Store category “Games” was therefore not included. Apps from other categories which were connected with games but not actual games were included. In addition to games, the top lists for the categories “Kids” and “Magazines & Newspapers” were also excluded, as these were collections of applications belonging to other categories. To illustrate, the top list for “Kids” was made up of apps from categories such as “Entertainment”, “Games” and “Education”, while the apps in the top list “Music” solely included apps within the category “Music”. Due to economic constraints of the project, only free applications were evaluated.

The study also excluded applications which 1) were not in Norwegian or English language, 2) were rated 17+ in App Store for sexual content and nudity, 3) required a membership that the author did not have access to, or which was paid, 5) turned out to be games even if selected from another category than “Games”, 6) were no longer available in App Store, or 7) did not work.

### **3.1.2 User group**

In this study, the focus was on accessibility problems for blind persons using the screen reader VoiceOver to interact with mobile content.

### **3.1.3 Sampling and assignment strategies**

To eliminate selection bias and ensure that all the apps represented in the 22 top lists had equal probabilities of being chosen for the sample, a *simple random sampling* strategy was used to make the selection of applications from the total population of 1100 apps. A computer-generated *lottery method* was used, in which each of the 1100 applications was assigned a unique number between 1 and 1100 and selected using a random generator in Excel. The selected applications were randomly assigned to one of the two possible feedback channels; review in App Store or email. Applications which did not

have the assigned feedback channel were excluded from the accessibility evaluations in order to avoid evaluating applications and finding problems that could not be reported. This was only the case with applications assigned to email since App Store reviews are available for all applications.

*Sampling without replacement* was employed. If an application was generated for the second time it was not included in the selection in order not to evaluate the same application several times. However, if an application which was excluded because it did not have the assigned channel (email) was sampled again and this time randomly assigned to App Store, the application was included.

#### **3.1.4 Evaluation method**

A *screening technique* was employed to do manual accessibility evaluations of the applications. This testing method involves using an interface in a way that sensory, motor or cognitive capabilities are artificially reduced in order to discover accessibility problems (28). To emulate the user experience for persons in the user group, the apps were evaluated using the screen reader VoiceOver for iPhone. In an attempt to make the screening conditions more similar to the experience of persons who are blind, visual support was removed when performing the evaluations by activating the *screen curtain* function in VoiceOver (45). The phone used for the testing was an iPhone 6 with iOS 10.2.1. English was the chosen phone language. All applications were updated to the newest version available.

When employing screening techniques, one option is to include pre-selected goals which evaluators try to achieve when exploring solutions (28). However, due to the wide variety of usage areas for the apps included in the study, it was not found appropriate to select particular goals, as it would be challenging to find goals applicable to all applications. To create appropriate tasks, the evaluator/author would have to familiarize with the applications. This was also a reason for not choosing to include specific goals since it was

considered important for the author to have as little experience with the applications as possible prior to the evaluations in order to prevent this potentially affecting the evaluation results. A more arbitrary solution was therefore chosen, where the author used the applications in the way that seemed natural and examined which VoiceOver related accessibility problems arose.

For the purpose of this study, the definition of *accessibility problem* was limited to “an incident of information or functionality provided in the mobile application being inaccessible through VoiceOver via touch input or audio output”. This included information conveyed through text, images, videos, colors, and graphical elements. In cases where information or functionality was reachable through touch input and audio output in VoiceOver, but it was hard to find or navigate through because of the lack of a logical navigation order, this was not considered an accessibility problem based on the definition described above.

Problems relating to interactive maps were not included in the study, as these were considered to have such a complex and different structure from other app content that this would require potential adjustments to the evaluation method and a more specific knowledge which the author did not have time to acquire within the project timeframe. Problems concerning ads were also excluded, for the reason that ads were classified as problems with third-party content, not the actual applications.

If and when considered appropriate the following VoiceOver gestures was used for interacting with the applications during the evaluations:

<b>VoiceOver gesture</b>	<b>Action</b>
Drag over the screen	Select and speak each item as you touch it.

One-finger tap	Speak the selected item.
One-finger swipe right or left	Select the next or previous item.
One-finger swipe up or down	This gesture performs different actions depending on the context: On an adjustable element, such as a slider, increment or decrement the value. In a text view, move the insertion point backwards or forwards.
One-finger double tap	This gesture performs one of the following: Activate the selected item. Toggle the selected switch. Unlock the lock screen when the Unlock switch is selected.
One-finger double press—With one finger, perform a double tap. During the second tap, continue to hold your finger against the screen.	Drag an item.
Two-finger tap	Pause reading. Two-finger tap again to resume reading.
Two-finger double tap.	Start and stop the current action.
Two-finger scrub—A Z-shaped gesture.	Go back hierarchically.
Two-finger swipe up	Read all accessible items from the top of the screen.
Two-finger swipe down	Read all accessible items from the current position.
Two-finger pinch open	Select text.

Two-finger pinch close	Deselect text.
Three-finger swipe up or down	Scroll a list or area of the screen.
Three-finger swipe right or left	Navigate to the next or previous page.
Four-finger tap at top or bottom of screen	Select the first or last accessible element on the screen.

Table 1: VoiceOver gestures, adapted from table by Apple Inc. (46).

Full accessibility evaluations of the applications were not performed, but rather evaluations with the purpose of finding one accessibility problem in each application. Every application was evaluated for a maximum of 30 minutes. The first problem found was recorded and reported in (as will be further described in section 3.2). If a problem was not discovered within these 30 minutes, the evaluator moved on to a new application. In order to achieve the same number of problems to report in each of the two channels studied, applications assigned to one channel were evaluated until a problem was found. Then applications assigned to the other channel were evaluated until discovering a problem. This process was repeated until a total of 50 accessibility problems were found (see section 4.1). 51 applications were evaluated in total. The apps were all re-tested two additional times to ensure that the problems discovered were not just one-time incidents.

To further investigate the problems discovered through the evaluations, they were also analyzed with the mobile screen on.

### 3.1.5 Trial evaluations

To practice the method and increase the probability of the author being able to find real accessibility problems, trial evaluations of 20 applications were performed in January 2017. Not practicing the method could cause the author to experience a steep learning curve when performing the actual evaluations,

consequently affecting consistency across the results. Training the author's skills through performing the 20 trial evaluations did not eliminate the risk of a learning curve influencing the results. However, it likely contributed to flattening the curve, minimizing its potential effect. The author sent a description of the first seven problems discovered to an expert in app accessibility, who confirmed that these were all valid accessibility problems. The person in question works for a company specializing in projects and research concerning ICT accessibility.

## **3.2 Reporting accessibility problems**

The total of 50 accessibility problems discovered through the evaluations performed were reported in two chosen feedback channels during a period of 22 days, from the 23rd of February to the 16th of March 2017.

### **3.2.1 Feedback channels**

Two digital feedback channels were studied in the research; App Store reviews and emails from applications. It was a priority to include both a public and a private channel, as comparing these was one of the aims of the study. Review in App Store was chosen as the public feedback channel. All iOS applications have the App Store review option, and these reviews are publicly available. As the private feedback channel, email was selected. The study was limited to emails provided in applications (as opposed to emails mentioned in websites, Facebook pages or other places), to ensure that only emails relevant for feedback about the applications were included. Feedback forms were not classified as emails in this study, even if messages sent through feedback forms could potentially result in emails to the companies receiving them. Both email addresses directly specified in the applications and emails opened in the Mail application in iPhone after clicking on links such as "give feedback" or "report a problem", were included. The Mail application was used for all reporting through email.

### 3.2.2 Reporting cycle

From the 50 accessibility problems discovered, 25 were reported through reviews in App Store pages for the applications, and 25 reported by emails specified in the applications. Ten problems were reported in at a time (five in App Store and five by email) on five separate days.

### 3.2.3 Feedback format

Text was the chosen reporting format. No additional means, such as videos or images, were employed. It was viewed as important for the texts to be similar across the different problem reports, in order to obtain comparable results. The following templates were therefore created and used as a basis for all the feedback messages. As opposed to operating under a fictitious name and account, the author used her own name and email address, as this was considered most ethical.

Template for feedback by email:

Subject: Accessibility problem in the (app name) app

Hi. There is a problem in the (app name) app related to the screen reader VoiceOver for iPhone: (describe problem). This problem can make the app less accessible for persons who operates it via VoiceOver due to a visual impairment, or for other reasons. I hope it will be fixed as soon as possible.

Best regards, Bjørg Lysbakken

Template for feedback in App Store reviews:

Accessibility problem in the app

★★★★☆ Bjørg L.

There is a problem in the (app name) app related to the screen reader



VoiceOver for iPhone: (describe problem). This problem can make the app less accessible for persons who operates it via VoiceOver due to a visual impairment, or for other reasons. I hope it will be fixed as soon as possible.

In App Store reviews it is required to give a star rating in addition to the textual comment. The object of the study was not to investigate the effect different star ratings can have on the result. The same rating (three stars) was therefore given in all the 25 reviews. In iTunes the App Store stars are described as follows: 1 star = I hate it. 2 stars = I don't like it. 3 stars = It's OK. 4 stars = It's good. 5 stars = It's great. Three stars were chosen because this was considered to be the most neutral and appropriate rating. As the reviews submitted in this study were concerning only one problem in the applications, giving a very low rating of one or two stars did not seem fair or ethically correct. Nor did giving one of the best ratings, four or five stars, since a problem was in fact discovered in the applications.

Two examples of finished reporting texts:

Subject: Accessibility problem in the (app name) app

Hi. There is a problem in the (app name) app related to the screen reader VoiceOver for iPhone: In the "Settings" page VoiceOver reads out "button" in the top of the screen, but gives no information about what the button is for. It can therefore be challenging to know the purpose of the button through VoiceOver. This problem can make the app less accessible for persons who operates it via VoiceOver due to a visual impairment, or for other reasons. I hope it will be fixed as soon as possible.

Best regards, Bjørg Lysbakken

Accessibility problem in the app

★★★★☆ Bjørg L.

There is a problem in the (app name) app related to the screen reader VoiceOver for iPhone: In the first page as a new user (the page with “Welcome”, “Login”, “Sign up” et cetera) the page control does not work properly when used through VoiceOver. When trying to switch between the three different pages, VoiceOver reads out “page 1 of 3”, “page 2 of 3” and “page 3 of 3”, but the content is the same for all three pages (it does not change). This problem can make the app less accessible for persons who operates it via VoiceOver due to a visual impairment, or for other reasons. I hope it will be fixed as soon as possible.

In cases where emails which were opened in the Mail application contained additional, automatically generated content such as a subject title or information about the app version, operating system version or phone language, this was included in the message.

Example of an email containing an automatically generated subject title and supplementary information:

Subject: (app name) Feedback

Hi. There is a problem in the (app name) app related to the screen reader VoiceOver for iPhone: In the “Map” page VoiceOver reads out “Image” in the bottom of the screen, but gives no information about the image content. It is therefore not possible to know what is on the image through VoiceOver. This problem can make the app less accessible for persons who operates it via VoiceOver due to a visual impairment, or for other reasons. I hope it will be fixed as soon as possible.

Best regards, Bjørg Lysbakken

App Name: (app name)

App Version: 4.3

Device: iPhone 6s

iOS Version: 10.2.1

Both the evaluation and reporting methods were designed in an attempt to be as realistic and close to the actual experience of the user group in focus as possible. Problems were therefore described in the feedback messages as they were experienced with the screen curtain on (no visual support), not as they were perceived when further analyzed with the screen on.

### **3.3 Documentation of response and post-evaluations**

All responses received from companies on the feedback sent were recorded. App Store does not give companies an opportunity to respond to the review messages. Hence, the documentation only included responses regarding the email feedback.

Each of the 50 applications with accessibility problems discovered was re-evaluated 30 days after reporting the problem, to check if the problem had been solved. Since the problems were reported in bulks of ten on five separate days, the new evaluations were also performed during five days. The problem situation was recreated using the same conditions as in the evaluations; apps were evaluated on the iPhone 6s with operating system version 10.2.1, VoiceOver, and the screen curtain function activated. The applications were updated to the newest version available. Apps which were not updated during the 30-day period were not evaluated.

### **3.4 Notice**

After completing the field experiment, the author made an effort to contact companies responsible for the 50 applications involved to inform about the study. For 25 applications, companies were contacted through the same email addresses as used to report problems in the experiment. For the 25 applications with problems reported via public App Store reviews, the contact options regarded as the most appropriate of the private once found was used. Although the author attempted to choose the most suitable alternatives, it can not be guaranteed that relevant employees or departments were reached in all cases. The notice was not given prior to or during the research, as it was considered that this would be destructive to the interaction the author wished to study in the experiment. Due to ethical considerations, the report does not contain names of the applications studied or companies connected.

### **3.5 Interviews**

To compare results from the field experiment with experiences of persons in the user group, interviews were conducted with three individuals recruited from an interest group for app accessibility on Facebook. The interviewees were contacted via Messenger with information about the study and the requests for participation.

#### **3.5.1 Participant profiles**

The participants ranged from the age of 33 to 57 and were all using VoiceOver due to blindness. Each had experience in employing digital channels to send feedback about VoiceOver related problems in mobile applications.

#### **3.5.2 Interview type**

The interview design can be classified as a mix between *structured* and *semi-structured*. The interviews were all based on the same set of pre-defined questions (see Appendix: Interview guide) as is common in structured

interviews. However, if additional questions and relevant information arose during the interviews, this was also recorded. A combination of open-ended and close-ended questions was used. The order of the questions was not strictly followed when this did not seem most appropriate.

The interviews were conducted via telephone. The conversations were recorded, using the program QuickTime Player on a laptop.

### **3.5.3 Informed consent**

Prior to the interviews, the subjects were given information about the research project, interview conditions, recordings, treatment and presentation of data collected, and their rights as research subjects, including their right to withdraw at any time (see Appendix: Interview guide). All subjects gave oral consent to participation before the interviews were conducted.

### **3.5.4 Anonymity and confidentiality**

The author's best efforts have been used to preserve anonymity and confidentiality of the participants in the presentation of information collected during the interviews. Descriptions of subjects include only the information considered most essential and does not contain personal information which could be directly linked to the participants, such as their name or contact information.

## 4 Results

The following sections present results from the accessibility evaluations, field experiment, and interviews conducted in the study.

### 4.1 Accessibility problems

From the total of 51 applications evaluated, a VoiceOver related accessibility problem was found in 50 (98 %) of the applications within the time limit of 30 minutes. The following table shows the different types of problems discovered and the number of problems found of each type.

<b>Id</b>	<b>Problem description</b>	<b>Amount</b>
1	No audio information about that an item is clickable (a “button” or “link”).	21
2	No audio information about a button’s purpose. Only “button” is read out by VoiceOver.	11
3	The page control does not work. Pages do not change.	7
4	No audio information about an image’s content. Only “image” is read out by VoiceOver.	3
5	Content is not reachable. It is not possible to access an item on the screen through VoiceOver.	3
6	VoiceOver gets stuck on certain content and it is not possible to proceed.	1
7	No audio information about an item, except a sound to signify that it is there.	1
8	A link does not work. Nothing appear to happen when clicking on it.	1
9	A video does not work. Nothing appear to happen when clicking on it.	1

10	A button does not work. Nothing appear to happen when clicking on it.	1
----	---	---

Table 2: Accessibility problems discovered through the screening evaluations (without visual support).

The most frequently discovered problem was missing information in VoiceOver about that an item was a link or button, which made it challenging to know that the item was clickable. Another commonly encountered problem was VoiceOver reading out “button” or “dimmed button” but no information about the button’s purpose. It was thereby difficult to determine what the button was for, without having to click on it and guess based on the resulting information. Page controls not functioning properly was also a recurring problem. When attempting to switch between different pages by swiping up or down with one finger (the VoiceOver gesture for changing pages in page controls) VoiceOver read out different page numbers, like “page one of three”, “page two of three”, “page three of three”, but the page content did not change.

For three apps the problems discovered related to images. No description of the image content was provided in VoiceOver; only “image” was read out. Due to the lack of description, it was not possible to know what was depicted in the images.

Three problems were concerned with information not being reachable via VoiceOver. In these cases, the evaluator understood that certain content was supposed to be on a page based on the context and other information provided, but it was not possible to access the content using any of the VoiceOver gestures.

In one of the apps evaluated VoiceOver got stuck at a certain point after registering a new account. It was not possible to proceed by clicking on any

items. The app was therefore no longer usable. In another application, one of the items did not have any audio description apart from a sound which indicated that there was an item there. Three of the problems found related to either a link, video or button not appearing to work when clicking on it.

The next table show how the problems discovered were perceived when analyzed without the screen curtain on.

<b>Id</b>	<b>Problem description</b>	<b>Amount</b>
1	No audio information about that an item is clickable (a “button” or “link”).	21
2	No audio information about a button’s purpose. Only “button” is read out by VoiceOver.	7
3	The page control does not work. Pages do not change.	7
4	False content (an item that is not visible on the page) is read out by VoiceOver.	6
5	Content is not reachable. It is not possible to access an item on the screen through VoiceOver.	3
6	No audio information about an image’s content. Only “image” is read out by VoiceOver.	2
7	VoiceOver gets stuck on certain content and it is not possible to proceed.	1
8	No audio information about an item, except a sound to signify that it is there.	1
9	A link does not work. Nothing happens when clicking on it.	1
10	No audio information about a video’s content.	1

Table 3: Accessibility problems as identified through the analysis (with visual support).



Through the analysis of the 50 problems with the screen on (screen curtain off), it was discovered that seven of them were experienced as different problems with and without visual support. Four of the problems identified as “No audio information about a button's purpose” in the evaluations, turned out to be “false” content not visible on the pages, which was read out by VoiceOver as “button”. This was also the case with one of the problems of “No audio information about an image's content” and the problem identified as “A button does not work”. In the first instance, VoiceOver read out “image” which made it seem like there was an image on the page which was lacking audio description. The second problem was classified as a button not working as nothing appeared to happen when the author clicked on what was read out as “button” through VoiceOver. However, in both these incidents, the evaluations with visual support showed that the real problem was that elements not intended to be presented in the pages (as these were not visible on the screen) got picked up and read out by VoiceOver.

In the case of the problem classified as “A video does not work” analysis with the screen on revealed that the video was in fact working. It was a silent video constantly running in the page background, and the actual problem was that no video description was provided in VoiceOver; only “Video” was read out.

In some of the incidents where false information was read out in the applications, the reason appeared to be that VoiceOver picked up content from underlying pages. In other incidents, the problem was seemingly that invisible elements used to build the page layout were read out by VoiceOver.

The two examples below demonstrate how a problem can be experienced differently with and without the screen curtain on. When VoiceOver is activated a black line marks the item currently selected, as shown in the illustrations.

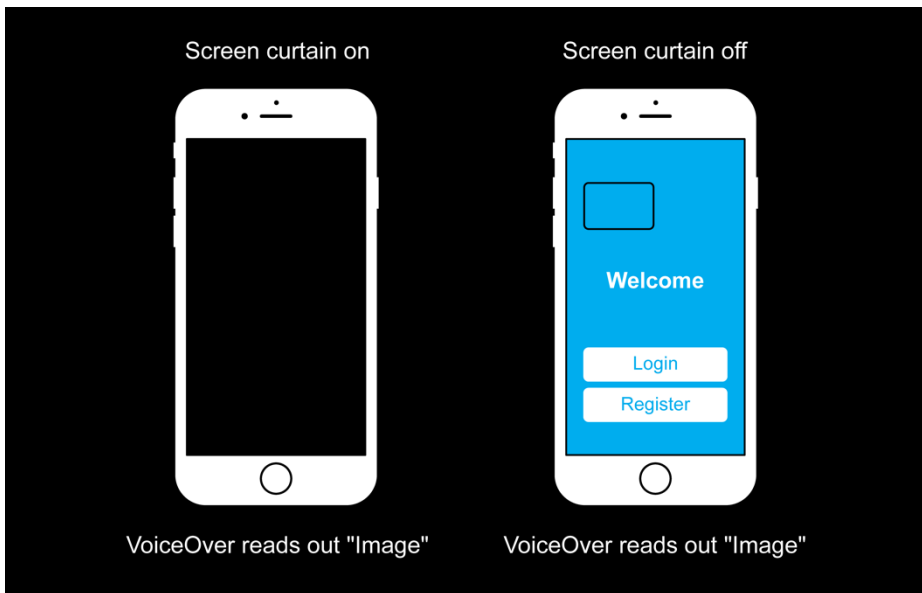


Figure 1: Illustration of a VoiceOver related accessibility problem with the screen curtain on and off. Example 1.

Figure 1 shows how a problem which is seemingly a lack of information about the content of an image when experienced with the screen curtain on, is, in fact, a problem of content not actually visible on the screen being read out by VoiceOver. In this case, VoiceOver seems to select and read out an image from a page under the current one.

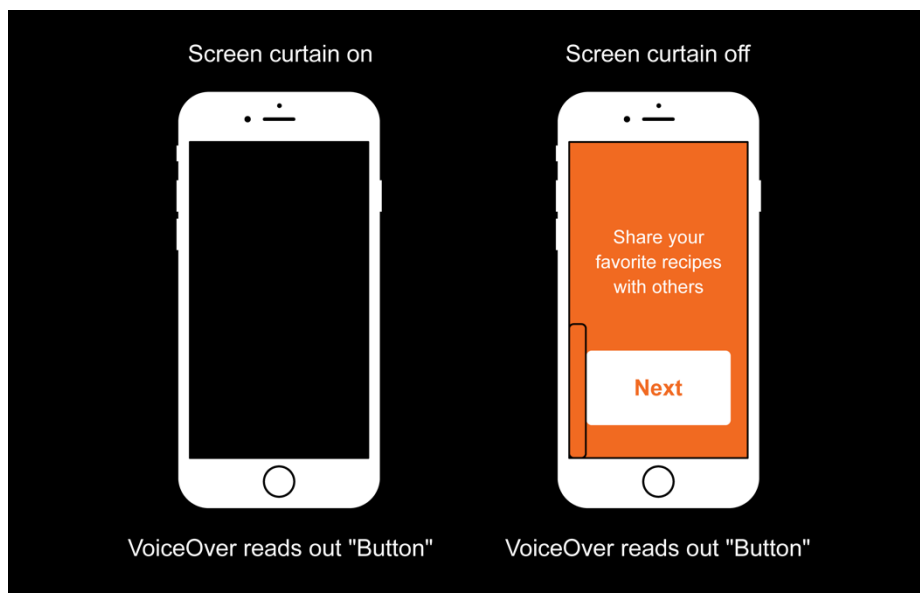


Figure 2: Illustration of a VoiceOver related accessibility problem with the screen curtain on and off. Example 2.

Figure 2 pictures a similar case as in the previous example. The problem which can be experienced as a lack of information about a button's purpose when the screen curtain is on turns out to be that information which is not presented on the screen is read out by VoiceOver. However, the item does not seem to be from an underlying page in this incident, but rather an element used to form the page layout.

## 4.2 Effect of sending feedback

From the 50 problems reported in through App Store reviews and emails, one (2%) of the problems were fixed when re-testing the applications after 30 days, while the remaining 49 problems (98%) were still occurring. However, in the one case where the problem was solved, this did not seem to be caused by the author's feedback, based on the company's response to the email. Through email correspondence between the author and app company, it was explained that the firm was already working on a new app version where the

issue reported was not present. This new version was published within the time of re-testing the application.

22 (44%) of the apps were updated with new versions at least once during the 30-day period from reporting the problems to re-testing the applications, based on App Store's version history for each application.

<b>Number of new versions</b>	<b>Applications</b>
0	28 (56%)
1	9 (18%)
2	3 (6%)
3	5 (10%)
4	2 (4%)
5	3 (6%)

Table 4: New versions of the applications released during the 30-day period between reporting and re-testing.

For the 25 emails sent with feedback about VoiceOver problems, 16 (64%) received a response. In 14 cases (56%), companies wrote that the issue would be either considered, fixed or forwarded to another person/department. In one response it was written that the issue did not concern the company contacted and that somebody else should be contacted instead. In another response, the company declared that they were not developing the application any further.

In 15 cases (60%), the companies did not encourage a further conversation by asking questions in their responses, while in one instance (4%) the company questioned the author about the problem and there was a correspondence back and forth.

Based on the author’s judgments some of the responses appeared to be automatic, computer-generated, and others written by humans. No exact number is presented, as it was in several cases challenging to determine if the response was automatic or manual.

### 4.3 Insights from interviews

In the interviews, the participants were asked about their experience and thoughts concerning VoiceOver related problems in mobile applications and sending feedback about such problems in digital platforms like App Store, Facebook, email, websites, Twitter and feedback options within applications. All direct quotes are translated by the author from Norwegian to English.

#### 4.3.1 Accessibility problems encountered

The ten types of VoiceOver related accessibility problems discovered through the app evaluations were presented to the participants, and they were asked if they could remember to have experienced any of them when using applications. Results are displayed in the following table.

Id	Problem description	Participants’ responses
1	No audio information about that an item is clickable (a “button” or “link”).	Experienced by all 3
2	No audio information about a button’s purpose. Only “button” is read out by VoiceOver.	Experienced by all 3
3	The page control does not work. Pages do not change.	Experienced by all 3
4	No audio information about an image’s content. Only “image” is read out by VoiceOver.	Experienced by all 3

5	Content is not reachable. It is not possible to access an item on the screen through VoiceOver.	Experienced by 1 Unclear answer from 2
6	VoiceOver gets stuck on certain content and it is not possible to proceed.	Experienced by 2 1 do not remember
7	No audio information about an item, except a sound to signify that it is there.	Experienced by 1 Unclear answer from 1 1 do not remember
8	A link does not work. Nothing appear to happen when clicking on it.	Experienced by all 3
9	A video does not work. Nothing appear to happen when clicking on it.	Experienced by 1 Unclear answer from 2
10	A button does not work. Nothing appear to happen when clicking on it.	Experienced by all 3

Table 5: Participants' experiences with accessibility problems discovered through the screening evaluations.

Each participant confirmed to have encountered most of the problems. In some cases, it was not clear to the interviewer/author if the interviewer and participant shared the same understanding of the problem described and if the participant had experienced the issue, based on the answer. This was classified as an "unclear answer".

Two participants said that they have experienced problem 1 and 2 often. One of them expressed that problem 2 is really annoying. One participant claimed to very commonly encounter problem 3 when installing new applications. The participant who had encountered problem 9 said that this issue was more prevalent earlier than it is nowadays. For problem 8 and 10, one participant pointed out that it can be difficult to know for sure when using VoiceOver if

links or buttons do not work, or if the problem is that new content appears in unexpected places making it challenging to find it.

One of the participants shared about other VoiceOver related problems encountered in addition to the once listed above:

- Wrong marking of elements. For example, that a button is marked as a headline.
- VoiceOver focus or pages changing without doing anything to influence it. It can seem like the text is just flickering and that everything happens by its own will.
- No opportunity to go back (like a back button or link) when using VoiceOver.
- Videos that are slow when used via VoiceOver.

According to the participant, problems with accessibility and screen readers has been apparent for decades. “I have been doing this since the end of the 80s with a screen reader and PC, and it has mostly been the same issues throughout the years”.

#### **4.3.2 Feedback habits and preferences**

“I started with sending feedback on applications quite early because I saw that it was necessary in order to get anywhere with gaining better user experiences myself”.

One of the participants stated to have recently started with sending feedback about VoiceOver related problems in applications. Another one has been reporting problems for about half a year and the third participant for around two years.

The participants were asked which feedback channels they have employed when reporting problems. One of the subjects has used emails, feedback

forms in apps, Facebook and App Store. The emails were usually found in the applications, through links for giving feedback. In Facebook, the participant has used both public wall posts and private messages, and said that this depends on the goal. "If I want to be a bit invisible, I use Messenger, but if it is an app one should be able to expect more from I can choose to use the wall". The participant mentioned that getting support from others is also a reason for sometimes deciding to use the public wall, and said that Facebook can work as a pillory where one can really make an impact.

The participant has stopped with using App Store for feedback about problems, and said that this has to do with the lack of opportunity for the developer to give a response to the feedback, the way the communication is regulated by Apple by for instance providing country dependent reviews, and that feedback messages disappear as soon as a new version is released, so that one has to look up messages for the previous versions. "I don't think that App Store works as a feedback channel in any way, actually". Now the participant rather uses App Store if getting very fond of an application and wanting to recommend it to others by giving a high score and writing that the app is accessible with VoiceOver.

The participant has not used feedback options in Twitter or websites connected with applications and expressed to think that the word limit in Twitter is not sufficient to properly explain accessibility problems, and that feedback via websites is bothersome and not direct enough, compared to concrete emails or Facebook pages.

The second participant expressed difficulty in remembering places used for sending feedback but said that in one case the feedback function within the application was used. The participant has not used Twitter or Facebook, and stated that the worst with Facebook must be that it is public. "I prefer a bit of discretion, even if one could maybe be part of a pressure group if doing it on Facebook". The participant adds that being part of a pressure group argues in



favor for using Facebook for feedback.

The last participant has reported problems via wall posts in Facebook pages and options provided in applications such as feedback forms and emails, as well as emails from other sources. The participant mentioned that it worked okay using the feedback solutions offered in Google apps, where one fills in predefined fields and explains the problem, and additional information such as which app version is used is automatically added to the message. The participant has not sent feedback via Twitter or websites and stated not to have managed to use App Store for giving feedback because of experiencing difficulties in finding the place for it, and also in using the feedback solution itself through VoiceOver.

All three participants said that they have only been using text in the feedback sent, not videos or screenshots, except for one person having used screenshots once. One subject stated to have no experience in making videos and that it would be a benefit to supplement the texts with videos, but that it is a challenge to overcome learning the technique. Another person explained about being uncomfortable with using videos. "I have to make it a bit simple because I am not comfortable with recording videos myself since I do not see what I am recording, so that results in only text".

The last participant declared not to have any idea about how to make videos but said that it would be very impactful and that it is a plan to investigate how to do it. The participant has learned how to take screenshots and states that technically this could, therefore, have been used, but that it has felt okay just using words to describe the problems. One of the other participants said that if using screenshots one would probably have to verify the screenshots with a person who can see them, to avoid getting suspected of not having a clue of what you sent feedback about, which could make things even worse.

On the question of how they think it works to explain problems using only text,

two participants stated that this depends on the problem. One of them said that it varies based on how in-depth the participant addresses the problem and if the participant is communicating with the developer or not. The participant has sometimes written directly with the developer and said that in these cases it has been very okay to use text. The other interviewee stated that the nature of the problem is a factor, and that if explaining where a problem is it can maybe be good to use visual means, but in incidents where something happens as a result of doing particular things in the application it is easier to explain this with text.

The participants were asked if they have ever sent feedback about the same problem several times. One participant said no, and two yes. However, one was not completely sure. The other participant told about sending feedback regarding the same application on a regular basis since 2015 and said that if the issues will not be fixed by the next update, the participant is thinking to bring it to either Facebook or the media.

#### **4.3.3 Feedback effect**

The participants' experiences regarding responses and effect gained from sending feedback have been different. The person who has recently started with sending feedback said that the response has been good. The participant has sent in feedback about two problems so far and stated that in both cases they were fixed. Regarding several people reporting the same issues, the participant said that one makes a stronger case in many ways if being a pressure group. The participant stated that things spread fast online and that it can really be of damage if companies do not choose to do something about the issues.

The participant who have reported problems for about two years also said to have experienced that some of the issues reported in got fixed, in the best cases within a week, but stated that in general less has been achieved than what the participant has given feedback on. The participant pointed out that it

is difficult to know if a problem is corrected because of the feedback sent in, or if this is a coincidence, “but of course when you get offered to be a beta tester, then you feel that you have accomplished something at least”.

The participant talked about one incident where corrections were made already in the next update of an application, while in another case, which the participant describes as the worst incident encountered, the message from the developer was that this is not something they prioritize, period. One company answered that they were not sure if they could use so much time for this because there are so many things they have to use time for. The participant has mostly gotten a form of response to the feedback sent.

When asked where the problems that got fixed were reported, the participant said “at least not App Store”, and that this has to do with the premises App Store has set. The participant explained that one of the problems that got solved was reported via the feedback option in the application.

According to the participant who has sent feedback for about half a year, the problems reported in have extremely rarely been solved. The participant said to recall only one time where a problem was fixed but stated that there were probably more people sending feedback about this issue.

On the question of whether the participant often gets a response to the feedback about VoiceOver problems in applications, the participant said that this depends on the definition of feedback. “The only response you get is that they have received and passed them on, and then it is silent. So, I would not call it a response really”. The participant explained that for 90% of the feedback there has been a reply, saying that the message was received and forwarded to the right person, but that the participant has never been put in contact with developers, or got asked any questions back.

#### **4.3.4 User experiences**

“I can’t bear to give feedback all the time, on everything. You get to a point where it’s enough”. The participants shared their feelings around using applications and sending feedback about issues encountered.

One of the interviewees talked about often feeling let down when using applications with VoiceOver. “That is a bit of the medal’s backside of using VoiceOver because you so often get disappointed. You so often meet things that do not work. Things could have been so easy, but they are still not”. The participant also said to be very positively surprised for those apps that actually facilitate VoiceOver, where you see that someone has thought about accessibility. The participant gave an example of an application installed about two weeks ago which worked very well, and said that in this case the participant considered to send feedback on the app because it was great, since receiving positive feedback could potentially make the company continue the good trend and think more about VoiceOver in future versions. The participant said that it can be smart to give feedback both on things that are bad and good since the good things can be forgotten in the middle of it all.

The participant talked about both negative and positive experiences connected with sending in feedback about problems. In one incident the participant stated to feel down prioritized by the company based on the response, while in another case it was a good experience since the app was corrected shortly after. The participant expressed frustration about sending a lot of feedback and said that at some point it is enough. “You do get a bit tired of giving feedback on things that don’t work all the time because then you are kind of taking on a very negative position in life that I am not that comfortable with really”. “You work very thoroughly. You often use a lot of time, and get nothing back”.

Another interviewee said that the motivation behind sending feedback is the wish to correct things that do not work so that it will get better for other people

and oneself eventually. The participant talked about getting anxious when the feedback sent in is not addressed. "You get a bit impatient when you don't get a response, and nothing happens". The participant said that one gets tired of giving feedback when usually not receiving a proper response that shows that the right people have gotten and understood what was communicated. "You do get tired and then maybe you drop the whole app". The participant shared an example of a case like this when choosing to stop using an application because of too many VoiceOver related problems.

The third participant talked about experiencing insecurities connected with sending feedback, regarding where the participant stood professionally to be able to explain the problems, and what to expect. The participant said that it was a big step starting to do this and that joining a group for app accessibility on Facebook gave the participant a push to try. "Then I dared to do it". The interviewee said that it gets easier to do something oneself when becoming part of a contact network like this Facebook group. The participant talked about feeling frightened the first time of formulating a feedback message, and also that it can feel scary when writing it in English.

The participant said that in both cases of sending feedback the experience has been positive and that once the participant had reported the first problem, the thought of reporting again seemed less daunting. "When you first have done it once then it is like okay, this went good and I was heard. Of course I can do this again". The person stated that in cases where one does not need to have lots of knowledge it is very okay and unscary to send feedback.

"I am so fond of that iPhone of mine". The participant talked about experiencing a new inspiration and curiosity after getting an iPhone, which the participant did not have previously because then it was only scary. According to the participant, the "button-anxiety" disappeared. The participant said that because of feeling so happy about all the things that are possible now with iPhones which were not possible earlier, the participant has previously just

left problems be. “It gives us so many opportunities, and it is even more out there”, “One almost get overwhelmed by all the possibilities that suddenly have opened up that the other things have only been left be, the things that have not worked so good. But of course, with the opportunity of giving feedback then it can only get even better”.

## **5 Discussion**

How effectively digital feedback channels can be used as platforms for solving accessibility problems depends on both the end users, who report issues encountered, and the companies, which address feedback received. This research looked into how companies respond to feedback sent in through digital channels, and how end users in the user group experience to employ these channels for notifying companies about accessibility problems. The following chapter will discuss results obtained from the experiment and interviews conducted, implications of the findings, limitations of the study and directions for further research.

### **5.1 VoiceOver related accessibility problems in mobile apps**

Results from the accessibility evaluations (see section 4.1) give insight into some of the accessibility problems users who employ VoiceOver can encounter in today's mobile apps. Since full evaluations were not performed, the results do not give a complete picture of existing VoiceOver problems, but rather highlight a few of the issues which can occur.

#### **5.1.1 Types of problems**

Ten different types of problems were identified through the evaluations with the screen curtain on (see Table 2 in section 4.1). Some problems had to do with information not being accessible, for example, that there was no audio information provided in VoiceOver about items being clickable, the purpose of buttons, or the content of images (problem 1, 2 and 4). Other problems discovered were related to inaccessibility of functionality provided in the applications, such as page controls, links or buttons not working, or that VoiceOver got stuck on particular content (problem 3, 6, 8 and 10). It was discovered that some of the problems were experienced differently when examined with visual support, which will be discussed in a following section.

### **5.1.2 Problem occurrence**

Due to the chosen evaluation design, the results can not give clear indications for how frequently the types of problems discovered appear in applications. Although certain kinds are more highly represented in the results, this does not necessarily mean that these are more common than problems found more rarely. The study was limited to include only the first problem discovered in each application. The apps could, therefore, contain additional problems which were not covered by the results. It is considered likely that certain types of problems can be more easily discovered than others when performing accessibility evaluations. This could also influence which type of problems were found, and the distribution between them. The problem of buttons missing audio description beyond just “button” for instance (problem 2 in Table 2), could potentially be simpler to discover than problems such as links not working (problem 8), where the evaluator had to try clicking on the link and exploring the page for results before realizing that the link did not work.

The problems detected were often from the initial pages presented in the applications, since only the first issue discovered was included. Certain problems might be more common in these particular pages, which could be another explanation for why they were found in a high number of applications. One example is the problem of page controls not working (problem 3). The initial pages users are introduced to when employing applications for the first time often contain page controls, which could be a reason why this was one of the problems discovered most frequently through the evaluations.

### **5.1.3 Screening technique versus evaluations with real users**

The purpose of the evaluations was to discover VoiceOver related accessibility problems which persons in the user group can experience when using applications. The fact that a screening technique was employed, as opposed to evaluating the applications with persons in the user group, is a limitation of the study with a potential impact on the validity of the evaluation results. The screening conditions were selected to be as realistic and similar



as possible to the experience of using applications for persons in the user group. However, the only way to be truly certain that results obtained are accurate for persons in the user group in question is to involve actual users. It has therefore been recommended to combine screening tests with evaluations performed by real users (34). According to Henry (34) screening tests can produce inaccurate or incomplete results, particularly when concerning assistive technologies. Henry claims that screen readers have a reasonably high learning curve and that problems discovered through screening evaluations can, in reality, be cases of evaluators' lacking skills in using screen readers. She also states that "People with disabilities are likely to do much better using a product than people who are using screening techniques because they are used to interacting with products with their functional limitation" (34).

In this research, trial evaluations were performed to minimize the chance of a steep learning curve affecting the results obtained in the actual evaluations. Although studying VoiceOver and practicing the screening technique in advance, the author's experience with using VoiceOver was still limited when performing the evaluations. It was, therefore, considered likely to be most comparable to the experience of novice VoiceOver users within the user group. For this reason, results could potentially be less valid for more experienced VoiceOver users.

To investigate if the accessibility problems discovered through the screening evaluations are legitimate problems experienced by persons in the user group, the three interview participants were asked if they have ever encountered any of the ten problem types when using mobile applications. All the problem types were confirmed by at least one participant, and some by two or three participants (see Table 5 in section 4.3.1). This suggests that the screening evaluations produced valid results. However, the type of problems discovered and the distribution among them could still have been different if performing evaluations with real users. As discussed in section 2.1, users

who share the same impairment might have very different experiences and needs. Conducting evaluations with a number of persons in the user group could therefore give an even more realistic picture of the current situation. There is also a chance of misunderstandings occurring between the interviewer and participants regarding the problems presented, which could lead to inaccurate confirmations. In cases where the interviewer doubted if the participant shared the same understanding of the problem described based on the answer given, this was classified as an “unclear answer”.

#### **5.1.4 Defining accessibility**

The study was limited to a particular definition of *accessibility problem*, presented in section 3.1.4. Results from the accessibility evaluations could, therefore, be different if using other definitions. Some of the problems found may not be classified as accessibility problems based on other definitions, and additional problems could be included which were not covered by the chosen definition.

#### **5.1.5 Status of accessibility in applications**

A problem was discovered in 50 out of the 51 applications evaluated in the study. This is a strong indication that VoiceOver problems are quite common in apps. Since full evaluations were not performed, the results can not provide insights into the total state of accessibility in applications with regards to VoiceOver. However, the findings do suggest that there is at least one VoiceOver related accessibility problem (as defined in section 3.1.4) in the vast majority of iOS applications. The sample of applications was created based on a random, unbiased selection (see section 3.1.1), which should strengthen the generalizability of the results by increasing the likelihood of the sample being representative of the chosen subpopulation. It is considered plausible that the results are more generalizable to popular iOS mobile applications in Norway within particular categories than to the overall population of mobile applications, as the app sample was drawn from this subpopulation (see section 3.1). The generalizability of the results could

potentially be improved by extending the research with evaluations on a larger sample.

In one of the applications, a problem was not detected. This does not necessarily suggest that there were no VoiceOver related problems in the application, but rather that a problem was not discovered within the limited evaluation time of 30 minutes.

The evaluator experienced it as easy and quick to uncover a problem in most of the applications examined. The majority of problems were found after only a few minutes of evaluation. In many cases, the evaluator did not need to go further than to the initial screen in the application before encountering a problem. This makes it seem probable that many applications, particularly within the subpopulation, will contain more than one VoiceOver related accessibility problem. However, to get a clear picture of the general state of VoiceOver accessibility in applications, more comprehensive accessibility evaluations with a high number of applications should be conducted.

#### **5.1.6 Differences in problem perception with and without visual support**

The problems identified through the accessibility evaluations were further analyzed with the screen curtain off, to compare how problems can be experienced without and with visual support. It was discovered that some of the problems appeared different when examining them visually, then when experiencing them solely through audio and touch (see section 4.1). In five cases where the problem was seemingly a lack of information, either about the purpose of a button or content of an image, the button or image turned out to be “false” content not visible on the pages, which was nevertheless picked up and read out by VoiceOver. This was also the case with a problem identified as a button not working. In another incident where the problem appeared to be that a video did not work though VoiceOver, analysis with the

screen curtain off revealed that the video was in fact working and that the actual problem was the lack of a video description.

The results indicate that there can in some cases be a difference in how VoiceOver related problems are perceived by persons experiencing them with and without visual support. This could potentially have an impact on the effectiveness of communication between end users in the user group and companies when addressing accessibility issues through digital feedback channels. Section 5.3 further discuss this matter. It is important to stress also in this context that results were obtained through evaluations performed by the author, not by persons in the user group. Evaluations with users could potentially give different results. This issue should, therefore, be investigated further with individuals from the user group involved.

## **5.2 Digital feedback channels as platforms for solving accessibility problems**

The main purpose of the study was to investigate how well digital feedback channels currently function as platforms for solving accessibility problems. Findings from the field experiment conducted may give indications of what effect user feedback through emails and App Store reviews have on solving VoiceOver related accessibility problems in mobile applications. Do companies address the feedback? Are problems usually fixed? Additionally, interview results provide insight into how feedback through digital channels has worked for persons in the user group.

### **5.2.1 Feedback effect**

From the 50 problems reported through App Store reviews and emails, only one problem was solved when evaluating the applications again after 30 days, while the remaining 49 problems were still present. The feedback for the solved problem was sent in via email. As explained in section 4.2, it seems improbable that the problem was corrected as a consequence of the

author's feedback, based on what was discovered through the email correspondence between the author and app company.

The results indicate that sending feedback through emails and App Store reviews currently has a minimal effect on solving VoiceOver related accessibility problems in mobile applications. This implies that companies are choosing not to address and correct such problems when receiving feedback about them through either of these channels. Factors like time, type and severity of problems, and feedback formats, formulation and quantity could potentially influence the results. This will be further discussed in following sections.

Since the experiment was conducted in a real-world environment, as opposed to a controlled lab setting, it was exposed to external factors. This could possibly compromise the reliability and internal validity of the findings. It would be challenging to find strong evidence of a causal relationship between the feedback sent in and problems being solved. Problems could be corrected due to other circumstances, such as feedback from additional users, or that companies were already working on solving the issues in question. As the results showed only one solved problem, this issue is not of big concern with regards to the findings of this particular study. In the case of the solved problem, it is considered unlikely that this was caused by the feedback sent in, as discussed in a previous section. However, this has a limited impact on the overall findings and conclusions.

As discussed in section 5.1.6, the random sampling strategy employed in the study should strengthen generalizability of the results. So should the fact that the research was performed in a naturalistic setting rather than a lab environment as it is considered plausible that the results more realistically reflect the phenomenon studied. Like with findings from the accessibility evaluations, the results may be more generalizable to the chosen subpopulation of popular iOS mobile applications in Norway within certain

categories (see section 3.1), then to the overall population of mobile applications. Extending the study to include a higher number of problem reports for a larger sample of mobile applications could improve the generalizability further.

According to interview results (see section 4.3.3), all three participants have experienced one or more incidents where VoiceOver related accessibility problems in mobile applications were solved after they sent feedback on the problems through digital channels. The results suggest that feedback in digital channels can sometimes have an effect on problems being fixed. It is, however, difficult to determine whether problems were solved as a result of the feedback submitted, or due to other circumstances, as discussed earlier in this section.

One participant stated that problems the participant have sent feedback on through digital channels have extremely rarely been solved, which support the findings from the field experiment. However, if comparing results from the experiment and interviews, the participants have experienced more problems being fixed after sending feedback than what the experiment results would indicate. This could potentially have to do with some of the factors discussed below, such as the type and severity of the problems reported, the way feedback messages were formulated, how many others sent in feedback about the same issues, and which channels were employed (users talked about digital feedback channels in general, not only App Store reviews and emails).

Two of the research questions this study intended to investigate was “What works most effectively, emails or App Store reviews?” and “What works most effectively, feedback through a public or a private channel?”. As the results of the field experiment conducted showed a marginal difference between the effectiveness of sending feedback in the two channels, these questions cannot be answered based on the findings. Results do not give indications on

which of the two channels works more effectively for solving VoiceOver related accessibility problems in mobile applications. What it seems to show is that neither of the channels works efficiently for this purpose.

### **5.2.2 Time**

The experiment was limited to a timeframe of 30 days between reporting problems and re-testing the applications to check whether problems were solved. More problems could potentially be corrected after a longer period. 44% of the applications were updated with new versions at least once during the 30-day period (see Table 4 in section 4.2), but this does not necessarily mean that these companies chose not to fix the problems. It could be prioritized for later versions. Solving problems might be a long process in some companies, due to reasons such as outsourcing the development of the application, or the amount of other problems which should be fixed. Companies could potentially have extensive lists of bugs reported in by users, making the process of correcting them time-consuming.

If judging by companies' responses to the feedback sent in by email, more problems will be addressed eventually. In 14 cases (56%), companies wrote that the issue would be either considered, fixed or forwarded to another person/department. However, even though companies write that the problems will be looked into or addressed, this might not be the case in practice.

### **5.2.3 Type and severity of problems**

For the purpose of this study 50 VoiceOver related accessibility problems in mobile applications, which can be categorized into ten different problem types (see Table 2 in section 4.1), were reported in through the two channels investigated. As full accessibility evaluations were not performed, it is considered probable that the results do not cover all different types of VoiceOver related problems persons in the user group encounter in mobile applications. Type and severity of problems could potentially influence the

feedback effect. Companies might give problems which are considered more critical higher priority than problems appearing less significant. Certain problems could also be easier to fix, possibly resulting in more problems solved of these kinds.

This study did not rank or prioritize problems based on type or severity, as the first problem discovered in each application was the one included. It may be the case that including other VoiceOver related accessibility problems, for instance only problems which could be considered critical, would have an impact on the results. This might be one of the main reasons why there is a difference between results of the experiment and the interviews in terms of the amount of solved problems. Interview participants may have sent feedback about problems which were more substantial or critical, which could be the reason why more problems were solved.

#### **5.2.4 Feedback formulation**

This study did not investigate how different ways of formulating feedback messages could potentially influence the results. The messages were limited to one person's style of writing and the template used. It is considered likely that users' messages will vary in style of writing, length and level of detail, which could impact companies' choice to address problems or not. Some users can, for instance, submit brief, unspecific messages like "This app doesn't work" or "The app is not accessible for VoiceOver", while others may send longer messages where problems are explained in detail. In cases where messages are vague or unclear, it could be challenging for companies to identify and correct the problems, while other messages might be more successful in conveying the issues, which could make it easier and quicker to address them. Messages might also be more or less convincing to companies, based on users' technical expertise and writing style.

Differences in the author's and interview participants' feedback formulation



could also be a potential reason why interview participants experienced that more problems were addressed than what the experiment results would suggest.

### **5.2.5 Feedback formats**

The feedback messages used in the study were solely text-based. Employing additional means such as videos or screenshots in messages could be more effective in terms of problems being solved. It might be easier for companies to understand and find problems if they are documented on video or screenshots, leading to more problems getting fixed.

As discussed in section 5.1.6, results from the accessibility evaluations conducted indicate that there can in some cases be a difference in how VoiceOver related problems are perceived by persons experiencing them with and without visual support. Videos or screenshots could contribute to avoiding misunderstandings that might occur between end users and companies in the communication about problems. However, there is a question of how many persons in the user group are comfortable and willing to employ these solutions.

None of the three participants interviewed had been using videos or screenshots in their feedback messages, apart from one participant employing screenshots once (see section 4.3.2). One participant said that it is a challenge to learn the technique of making videos. Another person explained about not being comfortable using videos because of not seeing what is recorded. A participant also mentioned that if using screenshots one would probably have to verify them with a person who can see before sending them. The interview results do not give an extensive picture of user habits and preferences within the user group regarding feedback formats, but they do show that not all users are choosing to employ videos and screenshots in

their feedback messages.

### **5.2.6 Feedback quantity**

Whether or not problems get solved may depend on the amount of attention they receive. If companies get a large amount of feedback about a particular issue, they might feel more pressure to solve it than if a small number of people notified about the problem. The amount of feedback could also be connected with the severity of the problems. In this study, one feedback message was sent in for each problem. It could be the case that more problems would be solved if they were reported in a higher number of times. Along with other factors discussed in earlier sections, the amount of feedback could be another potential reason why more problems were corrected in interview participants' experience than what was shown in the field experiment results. In addition to the feedback submitted by the participants, the problems which were solved can also have been reported in by a number of other people.

### **5.2.7 Channels**

This research investigated two digital feedback channels: App Store reviews and emails. Several other feedback solutions exist, which are potentially working more effectively as platforms for solving VoiceOver related accessibility issues in applications. This includes feedback options offered in social media, such as wall posts on companies' Facebook pages or messages through the Messenger chat, and tweets or private messages to companies on Twitter. Companies' web pages can also provide users with different options, like feedback forms, direct chats and forums.

The choice of channel could also be a possible explanation for why there was a deviation between the findings from the field study and interviews. Interview participants were talking about their experiences in using digital feedback channels in general, not only App Store reviews and emails. The reason why

more problems were seemingly solved as a result of the feedback submitted could, therefore, be that they were reported in through other channels than App Store reviews and emails. The status of how other channels currently work for solving accessibility issues, and how their potential can be utilized in the development of more inclusive ICT solutions, is interesting subjects for further research.

### **5.3 Reasons behind the seemingly low effectiveness**

As discussed in previous sections there are several factors which could potentially influence the results if the experiment would be conducted differently. However, the current findings do indicate that companies are in most cases not addressing feedback users send in through App Store reviews or emails about VoiceOver related accessibility problems. If this is actually the case, *why* are companies choosing to ignore issues raised by users through these channels? Could it be due to economy? Priorities? A lack of knowledge? Challenges in the communication with end users?

Results from the accessibility evaluations suggest that problems can in some cases be perceived differently if experiencing them with and without visual support (see section 4.1), which could cause confusion between end users in the user group and company employees when communicating about problems. In cases where problems are perceived differently by company employees when investigating the issues, then what was explained by users in the feedback, it might be difficult to understand and address the problems. This might be a reason why some problems do not get corrected.

The amount of feedback received may also affect companies' ability to solve problems. According to literature (discussed in section 2.6), businesses can receive extensive amounts of user feedback through platforms such as app stores, which can make it a challenging task to process it. Other explanations could be that company employees do not have sufficient knowledge to

address the problems, or that companies do not prioritize accessibility problems for reasons such as the belief that it will not benefit business. Companies might also be poor at utilizing user feedback in digital channels in general, not only feedback about accessibility issues.

This study can not give answers to the question of why accessibility issues voiced by users through App Store reviews or emails are seemingly not addressed by companies. Further studies should be done to investigate which reasons and potential challenges are holding companies back from attending to the problems, and how this could be changed.

## **5.4 User perspectives**

How successfully digital feedback channels can be used to solve accessibility issues not only depend on companies' response to problems reported in, but also end users' willingness to invest time and effort in sending feedback, and use the feedback solutions provided. Users personal preferences could determine their choice of feedback solutions, and their experience when using these solutions might impact whether or not they wish to continue with sending feedback.

In addition to investigating how companies address accessibility feedback received through digital channels, this study looked into the interaction in these channels from users' perspective. Results from the interviews conducted give insights into habits, preferences, and experiences of three persons in the user group, connected with sending feedback about VoiceOver related accessibility issues through digital channels (see section 4.3). The results do not provide a general picture of the situation for persons in the user group, but rather draw attention to experiences of some individuals, which can give valuable understanding for further work on the use of digital feedback channels as tools for solving accessibility problems.

#### **5.4.1 Feedback habits and preferences**

According to interview results, the participants have employed a number of different feedback channels, including options provided in Facebook and App Store, solutions from applications, such as emails and feedback forms, and emails from other sources. The results also show that there are channels which the participants have not used and that they can sometimes have particular reasons for choosing not to employ certain solutions. According to the participants, none of them have used Twitter for sending feedback. One participant said that App Store have not been employed, and another participant claimed to have stopped using it for feedback purposes. One participant said not to have used Facebook and two that they have not used websites.

It appears that participants' choices of which channels to use (and not to use) when sending feedback have been based on different reasons, including:

- challenges or bother in employing the channel
- limitations of the channel, such as a word limit or lack of opportunity for response
- issues with the way the channel is organized, for instance, that feedback messages are filtered by country or application version
- preferences for using a public or private channel

As for feedback formats, all three participants stated that they have only used text in their feedback, apart from one participant using screenshots once.

Based on the results, some of the reasons why participants have not employed videos or screenshots are seemingly connected with:

- a lack of experience and knowledge in making videos
- challenges in learning the technique of creating them
- discomfort in using videos, because of not seeing what is recorded
- contentment with using just text, as opposed to including screenshots

- discomfort in sending screenshots without first verifying them with a person who can see

Results give the impression that some of the participants would wish to use videos or screenshots, or would consider it, but that reasons described above have caused them not to up to this point. One participant stated that it would be a benefit to supplement the text with videos. Another participant said that it would be very impactful to use videos and that it is a plan to look into how to do it. One participant said that depending on the nature of the problem it may be good to use visual means.

The results show that some users can choose not to employ certain feedback formats and channels when notifying companies about accessibility issues and that these choices can be based on various reasons. It could be important for businesses to consider end users' habits and preferences when deciding on which feedback solutions to offer, in order to provide options that a large number of users wish to employ. To gain a broader picture of feedback habits and preferences of persons in the user group, as well as other relevant user groups, further research should be conducted.

#### **5.4.2 User experiences**

Participants have seemingly had both positive and negative experiences connected with the use of digital feedback channels. Two participants talked about positive experiences. One of them expressed that one incident was a good experience since the app was corrected shortly after the participant sent feedback. Another participant said that both cases of sending feedback have been positive, which has motivated the participant to continue with doing so. However, the participant also told about insecurities experienced when first starting to send feedback, with regards to being qualified enough to explain the problems, and what to expect.

Two of the other participants talked about negative experiences and emotions connected with sending feedback. One participant described the worst experience encountered, where the participant felt down prioritized when reading the company's response. The same person expressed frustration regarding sending a lot of feedback. The participant stated not to be comfortable with the negative focus of sending feedback about problems, and that the participant often invests a lot of time and effort without it having an impact. The second interviewee talked about feeling negative emotions due to the lack of effect of the feedback and proper responses from companies. Both participants said that they are getting tired of sending feedback.

Results suggest that some persons in the user group have experienced negative incidents and emotions connected with sending feedback in digital channels regarding accessibility issues, which could potentially influence their wish to continue with giving feedback in the future. Results also show that one of the participants was initially feeling insecure about starting to send feedback. Although the fears did not stop this particular user from sending the first feedback, it might potentially prevent other persons in the user group from doing so, if they experience the same fears. The findings suggest that more research should be done on how users experience to employ digital channels for sending feedback about accessibility problems, how this can potentially influence their habits, and how feedback processes can be improved to avoid bad user experiences such as those discovered.

## **5.5 Implications of the results**

Previous studies indicate that there are currently issues when it comes to the accessibility of mobile applications (as discussed in section 2.3.2), which is supported by findings from the interviews and accessibility evaluations conducted in this study. There is an apparent need to find solutions for how to develop more inclusive applications, to give a broader range of users the chance to take full advantage of all the opportunities applications provide.

Digital feedback channels such as App Store reviews and emails represent an opportunity for collaboration between end users and companies to solve accessibility problems, including problems relating to applications' compliance with assistive technologies such as the VoiceOver screen reader. However, findings from the field experiment suggest that in the majority of cases companies are not addressing the feedback received through App Store and by email regarding VoiceOver related accessibility issues. If the indications are accurate, this implies that there is a potential which is currently not being utilized by companies. User feedback in digital channels could provide companies with a valuable source of information about how applications work for users and what should be improved in order to enhance the accessibility. This could potentially be a cost-effective solution for getting input from a large number of users, to supplement methods used in earlier stages of the application development.

The present study can not give answers as to *how* big the potential of using digital channels for solving accessibility issues is. For that research should be done on aspects such as how many users employ these channels for sending feedback about accessibility problems and how useful this feedback is for companies in solving them. However, the three persons interviewed are examples of people in the user group who are employing digital channels for sending this kind of feedback.

Companies' apparent lack of initiative in solving VoiceOver related issues reported by users through App Store reviews and emails might also exist when it comes to other digital feedback channels and other types of accessibility issues experienced by users. This is certainly a matter for further investigation.



## 6 Conclusions

Results from the field experiment conducted indicate that sending feedback through emails and App Store reviews currently has a minimal effect on solving VoiceOver related accessibility problems in mobile applications. As the sample was limited to a selection from the most popular iOS applications in Norway within certain categories, the results are considered likely to be more generalizable to this specific subpopulation than to the general population of mobile applications. Factors such as time, type and severity of problems, and feedback formats, formulation and quantity could potentially influence the results and therefore be subject to further research.

The results showed a marginal difference between the effectiveness of using emails and App Store reviews. Which of the two feedback channels works more effectively can not be determined based on the findings. The same applies to the question of whether a public or private channel is more effective.

According to interview results, persons in the user group have experienced incidents where VoiceOver related accessibility problems in mobile applications were fixed after sending feedback through digital channels. This implies that communication in such channels can at times work successfully to solve accessibility problems. It is, however, difficult to determine whether problems were solved as a result of the feedback submitted, or due to other circumstances.

Additional findings from the interviews suggest that there are challenges which should be further investigated regarding how persons in the user group experience to employ digital feedback channels for communicating with companies about accessibility problems. Two interview participants talked about having negative experiences and emotions connected with sending

feedback on accessibility issues, for reasons such as not receiving proper responses, and lack of effect in terms of problems being addressed. One participant also talked about feeling insecure when first starting to send feedback, with regards to being qualified enough to explain the problems, and what to expect. The study does not address the question of how commonly the issues discovered are experienced among persons in the user group. However, the findings raise awareness about problems experienced by some individuals in the user group, which can give valuable understanding for further work on the use of digital feedback channels in communicating and solving accessibility problems.

The experiment results indicate that companies rarely solve VoiceOver related problems in mobile applications which are reported in through emails or App Store reviews. Questions arise as to why companies are not seemingly addressing feedback received through these channels. Could this be due to economy? Priorities? A lack of knowledge? Challenges in processing the amount of feedback? Future research can hopefully give answers to these questions.

Results from the interviews and accessibility evaluations conducted in the study reinforce what previous research indicate; that there are currently issues when it comes to the accessibility of mobile applications. Collaboration between end users and companies through digital feedback channels such as emails and App Store reviews could be part of the solution for achieving increased accessibility in mobile applications and other ICT solutions. Further studies should be done on the potential of using these channels for such purposes, and how to utilize this potential in more effective ways.

## Bibliography

1. World Health Organization. World report on disability. 2011 [cited 2017 May 24]. Available from: [http://www.who.int/disabilities/world\\_report/2011/en/](http://www.who.int/disabilities/world_report/2011/en/).
2. Wentz B, Lazar J. Exploring the Impact of Inaccessible Redesign and Updates. In: Langdon P, Lazar J, Heylighen A, Dong H, editors. Designing Around People: CWUAAT 2016. Cham: Springer International Publishing; 2016. p. 3-12.
3. Serra LC, Carvalho LP, Ferreira LP, Vaz JBS, Freire AP. Accessibility Evaluation of E-Government Mobile Applications in Brazil. *Procedia Computer Science*. 2015;67:348-57.
4. Park K, Goh T, So H-J. Toward accessible mobile application design: developing mobile application accessibility guidelines for people with visual impairment. *Proceedings of HCI Korea*; Seoul, Republic of Korea. 2729491: Hanbit Media, Inc.; 2014. p. 31-8.
5. Grussenmeyer W, Folmer E. Accessible Touchscreen Technology for People with Visual Impairments: A Survey. *ACM Trans Access Comput*. 2017;9(2):1-31.
6. Baazeem IS, Al-Khalifa HS. Advancements in web accessibility evaluation methods: how far are we? *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services*; Brussels, Belgium. 2843850: ACM; 2015. p. 1-5.
7. World Health Organization. WHO global disability action plan 2014-2021. 2015 [cited 2017 May 24]. Available from: <http://www.who.int/disabilities/actionplan/en/>.
8. World Health Organization. Assistive devices and technologies. 2017 [cited 2017 May 22]. Available from: <http://www.who.int/disabilities/technology/en/>.
9. World Health Organization. Assistive technology. 2016 [cited 2017 May 22]. Available from: <http://www.who.int/mediacentre/factsheets/assistive-technology/en/>.
10. American Foundation for the Blind. Screen Reading Technology and Refreshable Braille Displays. 2017 [cited 2017 May 22]. Available from: <http://www.afb.org/info/living-with-vision-loss/using-technology/assistive-technology-videos/screen-reading-technology/1235>.

11. United Nations. Accessibility: A guiding principle of the Convention. 2007 [cited 2017 May 21]. Available from: <http://www.un.org/esa/socdev/enable/disacc.htm>.
12. United Nations. Convention on the Rights of Persons with Disabilities. 2007 [cited 2017 May 24]. Available from: <http://www.un.org/esa/socdev/enable/rights/convtexte.htm#optprotocol>.
13. W3C. Accessibility, Usability, and Inclusion: Related Aspects of a Web for All. 2016 [cited 2017 May 21]. Available from: <https://www.w3.org/WAI/intro/usable>.
14. W3C. ACCESSIBILITY. 2016 [cited 2017 May 21]. Available from: <https://www.w3.org/standards/webdesign/accessibility>.
15. United Nations. The Standard Rules on the Equalization of Opportunities for Persons with Disabilities. 2007 [cited 2017 May 24]. Available from: <http://www.un.org/esa/socdev/enable/dissre00.htm>.
16. World Health Organization. What is e-accessibility? 2013 [cited 2017 May 21]. Available from: <http://www.who.int/features/qa/50/en/>.
17. W3C. Introduction to Web Accessibility. 2005 [cited 2017 May 21]. Available from: <https://www.w3.org/WAI/intro/accessibility.php>.
18. W3C. Mobile Accessibility. 2016 [cited 2017 May 24]. Available from: <https://www.w3.org/WAI/mobile/>.
19. Statista. Cumulative number of apps downloaded from the Apple App Store from July 2008 to September 2016 (in billions). 2016 [cited 2017 May 24]. Available from: <https://www.statista.com/statistics/263794/number-of-downloads-from-the-apple-app-store/>.
20. Statista. Number of apps available in leading app stores as of March 2017. 2017 [cited 2017 May 24]. Available from: <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/>.
21. Statista. Most popular Apple App Store categories in March 2017, by share of available apps. 2017 [cited 2017 May 24]. Available from: <https://www.statista.com/statistics/270291/popular-categories-in-the-app-store/>.
22. Android Developers. Accessibility.[cited 2017 May 24]. Available from: <https://developer.android.com/design/patterns/accessibility.html>.
23. Apple Developer. Accessibility for Developers. 2017 [cited 2017 May 23]. Available from: <https://developer.apple.com/accessibility/>.
24. Dale Ø, Drivenes T, Schulz T, Tollefsen M. Tilgjengelighetsevaluering av sosiale medie app'er på smarttelefon.[cited 2017 May 23]. Available from: <http://www.medialt.no/dokumenter/956.aspx>.

25. Leporini B, Buzzi MC, Buzzi M. Interacting with mobile devices via VoiceOver: usability and accessibility issues. Proceedings of the 24th Australian Computer-Human Interaction Conference; Melbourne, Australia. 2414591: ACM; 2012. p. 339-48.
26. Díaz-Bossini J-M, Moreno L. Accessibility to Mobile Interfaces for Older People. *Procedia Computer Science*. 2014;27:57-66.
27. Díaz-Bossini J-M, Moreno L, Martínez P. Towards Mobile Accessibility for Older People: A User Centered Evaluation. In: Stephanidis C, Antona M, editors. *Universal Access in Human-Computer Interaction Aging and Assistive Environments: 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part III*. Cham: Springer International Publishing; 2014. p. 58-68.
28. Brajnik G. Beyond Conformance: The Role of Accessibility Evaluation Methods. In: Hartmann S, Zhou X, Kirchberg M, editors. *Web Information Systems Engineering – WISE 2008 Workshops: WISE 2008 International Workshops, Auckland, New Zealand, September 1-4, 2008 Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 63-80.
29. W3C. *Web Content Accessibility Guidelines (WCAG) 2.0*. 2008 [cited 2017 May 24]. Available from: <https://www.w3.org/TR/WCAG20/>.
30. Rømen D, Svanæs D. Evaluating web site accessibility: validating the WAI guidelines through usability testing with disabled users. Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges; Lund, Sweden. 1463238: ACM; 2008. p. 535-8.
31. Power C, Freire A, Petrie H, Swallow D. Guidelines are only half of the story: accessibility problems encountered by blind users on the web. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Austin, Texas, USA. 2207736: ACM; 2012. p. 433-42.
32. Disability Rights Commission. *The Web Access and Inclusion for Disabled People: A formal investigation conducted by the Disability Rights Commission*. 2004 [cited 2017 May 24]. Available from: [http://collections.europarchive.org/tna/20070129145721/drc-gb.org/library/formal\\_investigation\\_report\\_w.aspx](http://collections.europarchive.org/tna/20070129145721/drc-gb.org/library/formal_investigation_report_w.aspx).
33. W3C. *Involving Users in Evaluating Web Accessibility*. 2010 [cited 2017 May 23]. Available from: <https://www.w3.org/WAI/eval/users.html>.
34. Henry SL. *Accessibility in User-Centered Design: Screening Techniques*. 2003-2007 [cited 2017 May 17]. Available from: <http://www.uiaccess.com/accessucd/screening.html>.

35. Licorish SA, Tahir A, Bosu MF, MacDonell SG, editors. On Satisfying the Android OS Community: User Feedback Still Central to Developers' Portfolios. 2015 24th Australasian Software Engineering Conference; 2015 Sept. 28 2015-Oct. 1 2015.
36. Panichella S, Sorbo AD, Guzman E, Visaggio CA, Canfora G, Gall HC, editors. How can i improve my app? Classifying user reviews for software maintenance and evolution. 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME); 2015 Sept. 29 2015-Oct. 1 2015.
37. Palomba F, Linares-Vásquez M, Bavota G, Oliveto R, Penta MD, Poshyvanyk D, et al., editors. User reviews matter! Tracking crowdsourced reviews to support evolution of successful apps. 2015 IEEE International Conference on Software Maintenance and Evolution (ICSME); 2015 Sept. 29 2015-Oct. 1 2015.
38. Pagano D, Maalej W, editors. User feedback in the appstore: An empirical study. 2013 21st IEEE International Requirements Engineering Conference (RE); 2013 15-19 July 2013.
39. Guzman E, Maalej W, editors. How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Reviews. 2014 IEEE 22nd International Requirements Engineering Conference (RE); 2014 25-29 Aug. 2014.
40. Hedegaard S, Simonsen JG. Extracting usability and user experience information from online user reviews. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Paris, France. 2481286: ACM; 2013. p. 2089-98.
41. Maalej W, Kurtanović Z, Nabil H, Stanik C. On the automatic classification of app reviews. Requirements Engineering. 2016;21(3):311-31.
42. Carreño LVG, Winbladh K, editors. Analysis of user comments: An approach for software requirements evolution. 2013 35th International Conference on Software Engineering (ICSE); 2013 18-26 May 2013.
43. Chen N, Lin J, Hoi SCH, Xiao X, Zhang B. AR-miner: mining informative reviews for developers from mobile app marketplace. Proceedings of the 36th International Conference on Software Engineering; Hyderabad, India. 2568263: ACM; 2014. p. 767-78.
44. Anam AI, Yeasin M. Accessibility in smartphone applications: what do we learn from reviews? Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility; Bellevue, Washington. 2513421: ACM; 2013. p. 1-2.
45. Apple Inc. Accessibility. 2017 [cited 2017 May 17]. Available from: <https://www.apple.com/accessibility/iphone/vision/>.

46. Apple Inc. Test Accessibility on Your Device with VoiceOver 2013 [cited 2017 May 17]. Available from:  
<https://developer.apple.com/library/content/technotes/TestingAccessibilityOfiOSApps/TestAccessibilityonYourDevicewithVoiceOver/TestAccessibilityonYourDevicewithVoiceOver.html>.

## **Appendix: Interview guide (translated to English)**

Type: Structured/semi-structured, telephone interview

Duration: 20-30 minutes

### **Phase 1: Introduction**

Informal talk

### **Phase 2: Information and oral consent**

Explain about the project and interview conditions:

I am a student within the master program for interaction design at NTNU in Gjøvik, and is currently working on a master thesis project where I investigate how digital feedback channels, such as App Store reviews and emails, work for communicating and solving accessibility problems in applications. More specifically the project is concerned with feedback on accessibility problems related to the screen reader VoiceOver for iPhone. As part of the study, I am conducting interviews with persons who are using apps via VoiceOver due to impaired vision, and who has experience with sending in feedback on accessibility problems in the apps.

The interview takes about 20 to 30 minutes and will revolve around your experiences and thoughts around sending feedback. All personal information like your name and contact information will be treated confidentially, and will not be presented in the report. In the report, it will be written that you are of adult age, that you use VoiceOver and why, and that you have experience with sending in feedback about accessibility problems in applications relating to VoiceOver. It will also say that all interview participants are recruited from an interest group for app accessibility on Facebook.



The interview will be recorded so that it will be easier for me to obtain all the answers. The recording will only be used as an alternative to taking notes. No one else then me will be listening to it, and the recording will be deleted during May 2017 at the latest. If you do not wish for there to be recordings, that is also okay. Participation in the interview is completely voluntarily, and you can withdraw at any time if you wish to.

Does this sound okay?

Is something not clear, or do you have any questions before we start?

Notify about starting recordings, and start it.

### **Phase 3: Interview**

1. How old are you?
2. Can you tell a bit about why you are using VoiceOver?
3. What kind of experience do you have with sending feedback on VoiceOver problems in applications?
4. Where have you sent in the feedback? Examples: App Store, Facebook, Twitter, feedback options in the app, email, websites.
  - 4.1. Which feedback alternatives have you used? Examples: Facebook: Wall post, chat. Twitter: Tweet, private message. In the app: Feedback form, email, chat. App Store: Reviews. Websites: Feedback form, chat.
  - 4.2. Are there any places you do not prefer to send feedback? App Store, Facebook, Twitter, feedback options in the app, email, websites.
  - 4.3. A particular reason for this?

5. What kind of means have you used in the feedback you have sent?  
Examples: Text, screenshots, videos, a mix.
  - 5.1. Are there any means you do not prefer to use? Examples: Text, screenshots, videos.
  - 5.2. A particular reason for this?
6. Has it happened that you reported the same problem several times?
7. How did you experience sending in feedback?
8. How has the feedback worked? Have any of the problems been fixed?
  - 8.1. Can you remember where you sent in the feedback that worked?  
Examples: App Store, Facebook, Twitter, feedback options in the app, email, websites.
  - 8.2. Do you remember if you used text, video, screenshots, or a mix in the feedback that worked?
  - 8.3. Approximately how long did it take before the problems were solved?
  - 8.4. Did you have to report in the problems several times before they got solved?
  - 8.5. Do you know if there were others that reported in the problems which were solved?
  - 8.6. Are there certain feedback channels where it has never or rarely worked to send feedback?

9. Have you often received response on the feedback you have sent in?

10. How have you experienced these responses?

If time:

11. Can you remember some of the VoiceOver problems you have encountered in apps?

12. Have you experienced any of the following things?

- No information in VoiceOver about that something is clickable (that it is a button or link).
- No information in VoiceOver what buttons are for, only “button”, “dimmed button” or something similar is read out.
- Content that completely lack description in VoiceOver, apart from a sound to signify that it is something there.
- Page controls that do not work. When you try to change between different pages VoiceOver reads out “page one of three”, “page two of three” et cetera, but the content does not change.
- No information in VoiceOver about the content of images, only “image” is read out.
- Content can not be reached through VoiceOver.
- VoiceOver gets stuck on a certain point, and it is not possible to proceed.
- Videos that do not work when you click on them (nothing happens).
- Links that do not work when you click on them.
- Buttons that do not work when you click on them.

**Phase 4: Round off**

Ask if there is anything the participant would like to add, and what the participant think about the interview itself.

Give thanks for the participation.

Note: All participants were contacted after the interviews to ask for consent to the author publishing their ages in the report.