

Designing a Conversational Interface for Guiding Online Peer Support Conversations Between Adults with ADHD



Master Thesis in Information Science

Author:

Oda Elise Nordberg

Advisors:

Professor Dag Elgesem

Professor Frode Guribye

June 2019

Acknowledgement

First of all, I would like to thank my supervisors, Frode Guribye and Dag Elgesem, for their support, motivation and expertise throughout this thesis.

I would like to express my gratitude to the members of the INTROMAT ADHD case for letting me be a part of the project, and for all their knowledge and time. So thank you, Astri J. Lundervold, Robin Kenter, Daniel André Jensen, Emilie Nordby, Anne Gro Parneman, and Suresh Kumar Mukhiya.

I am grateful for all participants who have been a part of this project, both in design workshops and in Wizard of Oz trials. You have all contributed with important information and insights for this thesis.

Further, I would like to express my gratitude to Ingvild Fiskerstrand – my own peer support through five years of studies at the University of Bergen. The value of having someone to relate to is indescribable.

Lastly, I would like to thank Kim and my family for their endless support and motivations. Especially in the times I needed it the most.

Thank you.

Table of Content

List of Figures	IX
List of Tables	XI
Chapter 1 Introduction	1
1.1 Research Questions	3
1.2 Structure of Thesis.....	4
Chapter 2 Background and Related Work.....	5
2.1 Human Computer Interaction.....	5
2.1.1 User Experience (UX)	6
2.2 Attention Deficit/Hyperactivity Disorder (ADHD)	7
2.2.1 Diagnosis.....	7
2.2.2 Treatment.....	8
2.2.3 Adult ADHD and Mental Health	8
2.2.4 Resource-Demand Imbalance in Adults With ADHD	9
2.3 Peer Support.....	10
2.3.1 Online Peer Support.....	12
2.3.2 Benefits of Online Peer Support	12
2.3.3 Risks of Online Peer Support	14
2.3.4 Guiding in Online Peer Support	15
2.4 Conversational Interfaces: Chatbots.....	15
2.4.1 Potential Areas of Use	16
2.4.2 Woebot	17
2.5 Design Considerations.....	18
2.5.1 Designing for Adults with ADHD	18
2.5.2 Designing for Mental Health.....	19
2.5.3 Designing Chatbots.....	19
2.6 Chapter Summary	22

Chapter 3 Methodology.....	23
3.1 Research Through Design as a Method for Interaction Design Research in HCI	23
3.1.1 Problem Identification	24
3.1.2 Evaluation	24
3.1.3 Why Use RtD?	25
3.2 Interaction Design - Beyond the Artifact.....	25
3.2.1 From User-Centered to Participatory Design	26
3.2.2 Personas.....	27
3.3 Ethical Approval	27
3.4 Prototyping	28
3.4.1 Design Workshops.....	29
3.4.2 Conceptual Design	29
3.5 Evaluation.....	30
3.5.1 Wizard of Oz.....	30
3.5.2 Sampling of Participants.....	31
3.5.3 Online Questionnaire: Measuring User Experience.....	32
3.5.4 Focus Group.....	32
3.6 Chapter Summary	33
 Chapter 4 Conceptual Design and Establishing Requirements	 35
4.1 Tools	36
4.1.1 Adobe XD	36
4.1.2 Slack.....	36
4.2 First Iteration	36
4.2.1 First Meeting: Discussion of Possibilities	37
4.2.2 Design Workshop: Specifying Requirements.....	37
4.2.3 Summary.....	41
4.3 Second Iteration	41
4.3.1 Personas.....	41
4.3.2 Advancement of the Idea	42
4.3.3 Design Workshop: A Peer Support Chatbot	42
4.3.4 Summary.....	45
4.4 Third Iteration	45
4.4.1 Design of the Chatbot	45
4.4.2 Design Workshop: Improving Chatbot Script.....	47

4.4.3 Summary	49
4.5 Chapter Summary	49
 Chapter 5 Wizard of Oz Trials.....	 51
5.1 First Wizard of Oz Trial: Evaluation of Scripts and Proof of Concept.....	51
5.1.1 Results: The Wizard's Notes and Chat Logs.....	53
5.1.2 Results: Data from the Chats.....	56
5.1.3 Results: Online Questionnaires.....	56
5.1.4 Summary	59
5.2 Adjustments and Specifications	60
5.3 The Second Wizard of Oz Trial: Evaluation with Adults with ADHD.....	61
5.3.1 Results: The Wizard's Notes and Chat Logs.....	62
5.3.2 Results: Data from the Chats.....	64
5.3.3 Results: Online Questionnaires.....	65
5.3.4 Results: Focus Group	66
5.3.5 Summary	69
5.4 Chapter Summary	69
 Chapter 6 Discussion	 71
6.1 The Potential of Using a Chatbot for Peer Support.....	71
6.1.1 Design Choices – The Emergence of Terabot.....	72
6.1.2 The Design of Terabot vs. Design Considerations	75
6.2 Research Questions	78
6.2.1 Designing Guided Chat-based Peer Support	78
6.2.2 Designing Guiding Conversational Interfaces	79
6.2.3 The Wizard of Oz Method Used in the Design of Conversational Interfaces.....	79
6.3 Research Limitations	81
6.4 Chapter Summary	81
 Chapter 7 Conclusion.....	 83
7.1 Future Work.....	84
 Bibliography	 85

Appendix A – Online Questionnaire.....	92
Appendix B - Personas.....	94
Appendix C – Chatbot Script	95
Appendix D - NSD	101
Appendix E – NORSRII Abstract.....	103

List of Figures

Figure 2.1: ADHD subtypes and their core symptoms (Williams, n.d)

Figure 2.2: A diagram for the conceptual Resource-Demand Imbalance predicted for people with ADHD.

Figure 2.3: “A typology of chatbots with four example chatbot purposes located within the typology dimensions” (Følstad, Skjuve and Brandtzaeg, 2018)

Figure 4.1: A visualization of requirement choice regarding purpose of peer support.

Figure 4.2: A visualization of the requirement choice regarding group composition.

Figure 4.3: The left chat demonstrates a fully guided chat, while the two right ones demonstrate a semi-guided chat.

Figure 4.4: A visualization of three different user profiles with various degree of personal information. The far left represent a fully anonymous profile, the middle a quasi-anonymous profile, and the far left a fully open profile.

Figure 4.5: Persona 1 – Woman, 35 years old with ADHD Combined Type.

Figure 4.6: Picture a) shows a possibility for how the Chatbot could start a conversation.

Picture b) display how a chatbot guided peer conversation could work. Picture c) demonstrate how the chatbot could facilitate a group dynamic. Picture d) visualizes how the chatbot promote the conversation to revolve around the self-help program.

Figure 4.7: The profile picture used for the chatbot, Terabot.

Figure 5.1: A screenshot of the chat environment from the wizards’ perspective.

Figure 5.2: *A diagram illustrating the level of connectedness in ‘Chat 1’ and ‘Chat 2’*

Figure 5.3: *An illustration of a possible framework for the chatbot conversation*

List of Tables

Table 3.1: An overview of the different phases, purposes, methods, samples and data collected throughout this thesis.

Table 5.1: The durations of ‘Chat 1’ and ‘Chat 2’ measure in minutes for each group.

Table 5.2: Average scores from the online questionnaire the participants answered after each chat.

Table 5.3: The durations of ‘Chat 1’ and ‘Chat 2’ measure in minutes for each group.

Table 5.4: Average scores from the online questionnaire the participants answered after both chats in the alternative WoOz trial.

Chapter 1

Introduction

Most people experience troubles with their concentration, impulsivity and energy level. These troubles can for instance be demonstrated by difficulties keeping focus when tired, having so much fun that one lose track of time, or suddenly get the desire to do something new and exciting. If this happens every now and then there is no need to worry, but for some people this is a major part of everyday life, and affects their life in a negative way. Attention Deficit/Hyperactivity Disorder, better known as ADHD, is recognized by its three main characteristics: inattentiveness, hyperactivity, and impulsivity (Hannås, 2019). In Norway it is estimated that 3-5% of children and 2,5% of adults qualify for this diagnosis (ADHD Norge, 2016c; Helsedirektoratet, 2016a).

For a long time ADHD was believed to be a condition only affecting children, and when diagnosed the child affected would grow out of the diagnosis before entering adulthood. This has been disproved by recent research, which indicated that 2/3 of all children with ADHD still qualify for the diagnosis as adults (Øie, n.d.; ADHD Norge, 2016a). In 1994 the criteria for the diagnosis was changed in order for adults to be diagnosed and receive treatment (Hannås, 2019). Even though ADHD became classified as a chronic disease (Pliszka, 2007; Atkinson and Hollis, 2010; Kooij, 2010) the support services for adults with ADHD are inadequate. In the recent years it has been expressed an increased demand for non-pharmaceutical treatment options for adults with ADHD (NCCMH, 2009; Intromat, 2016b).

One non-pharmaceutical treatment option for adults with ADHD is self-help programs. This thesis is done as a part of INTROMAT (2016a), which is a project with the aim of improving mental health by the use of innovative technology. Within INTROMAT a subproject is developing an online self-help program for adults with ADHD inspired by Goal Management Training (GMT), an executive functioning intervention designed to help participants improve their organizational skills and achieve goals (Levine *et al.*, 2011). In the online self-help program the participants will get access to information, techniques and exercises aiming at making everyday life easier (Intromat, 2016). As a supplementary support system to the

online self-help program, the INTROMAT team expressed a desire to explore the possibility for peer support technology. In recent years promising research on peer support technology has emerged, especially in relation to mental health and online interventions (e.g. Naslund *et al.*, 2016; O’Leary *et al.*, 2018).

According to Hurvitz (1970, p. 47) “Individuals with the same problems serve as the most effective role models for each other”, meaning that peers have great potential to influence each other in a positive manner. Peer support, the idea that people in similar situations support each other, has been shown to promote treatment engagement and prevent treatment drop out (Alvarez-Jimenez *et al.*, 2014). Recent studies show that peer support can increase confidence, improve mental health and wellbeing, as well as deal with stigma and discrimination (Faulkner and Basset, 2012). There is no doubt that peer support can be beneficial for an online self-help program, but one question is how to design such service. According Baumeister *et al.* (2014), online interventions should always be guided if possible. Having a person guiding these interventions could be time consuming and expensive. Conversational interfaces, often referred to as chatbots, are technologies that provide users with access to data and services through natural language dialogue (Følstad and Brandtzaeg, 2017). Chatbots are cost-effective and easy to use, and have already been used to deliver therapy (Fitzpatrick, Darcy and Vierhile, 2017; Skjuve and Brandzaeg, 2018).

Several researchers have noted that interaction designers and human-computer interaction (HCI) researchers face new challenges with the emergence of conversational interfaces, as their focus has previously been on designing for graphical user interfaces (Følstad and Brandtzaeg, 2017; van Allen, 2018; Q. Yang *et al.*, 2019). With the transition from graphical to conversational user interfaces, one must consider conversations as objects of design and acknowledge the new opportunities and challenges this generate (Følstad and Brandtzaeg, 2017). There is limited research on how to best design for conversational interface experiences. Wizard of Oz (WoOz) trials are characterized by having users interacting with a piece of software as though interacting with a functioning product, but in reality human operator simulates the software’s response. Evaluating possibilities and limitations of the WoOz method when designing chatbots could be beneficial for future chatbot design.

As described above this thesis has several aims:

- Explore the possibilities for peer support technology in relation to an online self-help program for adults with ADHD.
- Try to understand how we can guide online peer support conversations.
- Examine the potential for a conversational interface to guide group chats.
- Review benefits and limitations of using Wizard of Oz trials in the design of a peer support conversational interface.

1.1 Research Questions

The primary research question of this thesis is:

RQ1: How can we design guided chat-based peer support technology accompanying an online self-help program for adults with ADHD?

Based on findings and discussions with professionals, the possibility to design a conversational interface to guide online peer support conversations was articulated. As the research aim got specified to include a conversational interface, two additional research questions emerged:

RQ2: How can we design a conversational interface which role is to guide a group conversation between adults with ADHD?

RQ3: What are the opportunities and limitations of using the Wizard of Oz method when designing conversational interfaces?

1.2 Structure of Thesis

Chapter 1: Presents the problem space for this thesis, along with the research aims and research questions.

Chapter 2: Presents relevant literature on topics related to the research questions.

Chapter 3: Describes methods used in the design and evaluation phase.

Chapter 4: Describes design iteration one to three.

Chapter 5: Describes two iterations of Wizard of Oz trials.

Chapter 6: Discussion of the concept, results, and research questions.

Chapter 7: Conclusion of this thesis. Describes future work.

Chapter 2

Background and Related Work

This chapter focuses on relevant subjects covered in this master thesis. This chapter starts by introducing human-computer interaction (HCI), which is the field of research in this thesis. Further a general overview of ADHD is given, including a more in-depth focus on adult ADHD and mental health and resource-demand imbalance. Additionally, research on peer support and chatbots are presented. The chapter ends with presenting design considerations in connection to ADHD, mental health, and chatbots.

2.1 Human Computer Interaction

Human–Computer Interaction (HCI) is a research field concerned with “...the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them” (Hewett *et al.*, 1992, p. 5). HCI is both an academic discipline that studies technology’s influence on human activity, and a design discipline where the goal is to design technology for maximum effectiveness. HCI can be traced back to the 1950s’, however it was not until the 1980s with the increase in personal computers it really started to take shape and got its’ name. HCI is an interdisciplinary research field that incorporates expertise from several academic areas such as computer science, psychology, sociology, anthropology, cognitive science, and industrial design (Hewett *et al.*, 1992).

Oulasvirta and Hornbæk (2016) have described HCI research as *problem-solving*. This description is built on Laudan’s philosophy of science (Laudan, 1977). Their description of HCI involves the concept of *research problems* as “... inabilities and absences occurring in description; knowledge; or, as often in HCI, constructive solutions”. By this definition a wider selection of possible research problems occur. Research problems can deal with lack of understanding in regards to colors in a prototype, or how a particularly layout on a webpage affect the user experience. Oulasvirta and Hornbæk (2016,) state that the *solutions* to research problems should change the status of the inabilities and/or absence described in the research problem, leading to HCI as *problem solving*.

Building on Laudan's notion of research problems as either *empirical* or *conceptual*, Oulasvirta and Hornbæk added a third category named *constructive*. "Constructive research is aimed at producing understanding about the construction of an interactive artifact for some purpose in human use of computing" (Oulasvirta and Hornbæk, 2016, p. 4958). They further stress it is not the construction or the artifact itself that is of value, but the ideas or principles that manifests. Oulasvirta and Hornbæk (2016) described *empirical research* as "... aimed at creating or elaborating descriptions of real-world phenomena related to human use of computing". In other words, researching a phenomena novel to HCI research. *Conceptual research* problems are non-empirical, meaning that the problems they concern are conceptual and theoretical. Oulasvirta and Hornbæk (2016, p. 4958) described these problems as "aimed at explaining previously unconnected phenomena occurring in interaction". Examples of conceptual problems can be trying to explain empirical phenomena, or define conflicts in design principles.

This thesis contributes with the construction of a prototype (*constructive research*), as well as field trial evaluations of the prototype (*empirical research*). The prototype in this thesis has no psychical form, as it is a concept, requirements and scripts for a peer support chatbot. However, as Følstad and Brandtzaeg (2017) stated; conversations should be considered as prototypes now that conversational interfaces are becoming more common. Additionally, Oulasvirta and Hornbæk (2016) empathized that it is the ideas and principles that manifest that is of value. In this thesis the prototype was evaluated in two Wizard of Oz (WoOz) trials. The aims were to examine the user experience of interacting with a peer support chatbot, explore user behavior in peer support chats guided by a chatbot, and test the opportunities of using WoOz trials to evaluate a scripted chatbot prototype.

2.1.1 User Experience (UX)

User Experience (UX) is a central concept concerned in HCI. The UX of a product concerns how it behaves and is used by people. That is, "... the experience the product creates for the people who use it in the real world" (Garret, 2011, p. 6). Every product that is used by a person provides that person a user experience. Giving the user a satisfying user experience is essential when developing a new product/artifact, but this can be challenging. One can only try developing *for* a good user experience, by designing features one thinks will give the user the desired user experience. The more complex the product design is, the more difficult it is to

identify how to create a successful user experience (Garret, 2011). In this thesis it was desired that the participants would enjoy engaging in the peer support conversations guided by the chatbot. Satisfying UX was valued as it was intended for participants to have recurring interactions with the peer support chatbot, and for this to be realistic it must contribute to a fulfilling experience. To accomplish a satisfying user experience participants were part of the design process, and the prototype was designed and evaluated through several iterations.

2.2 Attention Deficit/Hyperactivity Disorder (ADHD)

As mentioned in the introduction, Attention Deficit/Hyperactivity Disorder (ADHD) is one of the most common psychiatric conditions in the world. The condition is recognized by its three symptoms; inattentiveness, hyperactivity, and impulsivity. The condition will occur sometime during childhood or early adolescent, and can persist throughout adulthood. The cause of ADHD is not fully known, but researchers agree that it is both environmental and genetic factors involved (Hannås, 2019). The three main symptoms of ADHD can be of varying degree and combinations leading to three different subtypes: ADHD hyperactive/impulsive type, ADHD inattentive type, and ADHD combined type. There can be considerable individual variations in the different symptoms regarding frequency, intensity and effects on everyday life (ADHD Norge, 2016a). Figure 2.1 illustrates the subtypes of ADHD and their core symptoms.

	<i>ADHD, Primarily Inattentive</i>	<i>ADHD, Hyperactive-Impulsive</i>	<i>ADHD, Combined Type</i>
<i>Inattentive/ Poor Attention Span</i>	X		X
<i>Impulsive and/or Hyperactive</i>		X	X

Figure 2.1: ADHD subtypes and their core symptoms (Williams, n.d).

2.2.1 Diagnosis

Symptoms of ADHD can be similar to other conditions such as anxiety and depression, making it challenging to diagnose. A medical evaluation that spans over a period of time must be conducted, and clinical interviews performed by a psychologist or a psychiatrist will be administered (ADHD Norge, 2016b). In many cases there will be requested additional information from close family, work colleagues, former teachers, or similar. This is done to

ensure that the symptoms have been prominent since childhood or early adolescent, and not recently occurred due to unknown circumstances (Helsedirektoratet, 2014). Another important prerequisite is that the symptoms must cause impaired function in several areas of daily life, for instance at work/school, relations/family life, and spare time/hobbies (Øie, 2015; Hannås, 2019).

2.2.2 Treatment

The main goal when treating ADHD is to improve daily functioning and reduce symptoms experienced. This is usually achieved by combining medication, facilitated education and/or work, and behavioral therapy (Norsk Helseinformatikk, 2017). When planning which type of treatment the patient will receive, there are several factors that need to be taken into account: Which type of ADHD, the severity, the impact on functioning, and other difficulties (Helsedirektoratet, 2014). The most common treatment method is stimulants. It is estimated that stimulants have a positive effect for 75% of children and 50% of adults with ADHD (Helsedirektoratet, 2016b). Other types of treatments are information about the diagnosis, assistive support for remembering and organizing, help to establish better routines, and facilitations at work/school. If the patient has an additional psychological difficulty therapy is often offered (Øie, 2015). There is expressed a desire for more non-pharmaceutical treatment options (NCCMH, 2009; Intromat, 2016b).

2.2.3 Adult ADHD and Mental Health

ADHD affects life in several ways. Initially, the person with the condition is affected as a consequence of the symptoms leading to impairments in various aspects of their lives. Thereafter, ADHD can affect the person as a result of the stigma associated with this diagnosis. Stigma is designated as “A mark of disgrace associated with a particular circumstance, quality, or person” (English Oxford Dictionary, 2019). Pescosolido et al. (2008) found that kids with ADHD often were stereotyped as “... lazy, bad or aggressive, or considered to have a behavioral or special needs problem rather than a mental health disorder that requires treatment”. These prejudices are incorrect and accentuate an untrue understanding of the condition, which can increase stigma.

Comorbidity, which refers to the addition of at least one additional condition, is common for people with ADHD. Approximately 50% of the kids and 75% of the adults have additional difficulties. For children this is often learning disabilities, behavioral disorders, tics, sleep problems, and mobility problems. For adults this is most commonly affective disorders, such as anxiety, depression, bipolar disease, personality disorders, or substance abuse (Helsedirektoratet, 2014; Øie, 2015; ADHD Norge, 2016a; Hannås, 2019).

2.2.4 Resource-Demand Imbalance in Adults With ADHD

Inattentiveness, hyperactivity and impulsivity persist as the person with ADHD grows older, but the overt signs of hyperactivity and/or impulsivity decline with age. Inattentiveness persists mainly unvaried (Turgay *et al*, 2012). As one grows older the demands one faces increase, at the same time as the support resources one receives decrease. This is illustrated in Figure 2.2. Environmental demands consist of “... academic, occupational, financial and social activities and functioning” (Turgay *et al*, 2012, p. 2). Resources are defined as what one uses to meet these demands. Resources can consist of internal processes (working memory, planning and organizational skills, social skills and similar), and external processes (personal support network consisting of family and other people, and, other means such as planners, medication and alarm clocks) (Turgay *et al*, 2012).

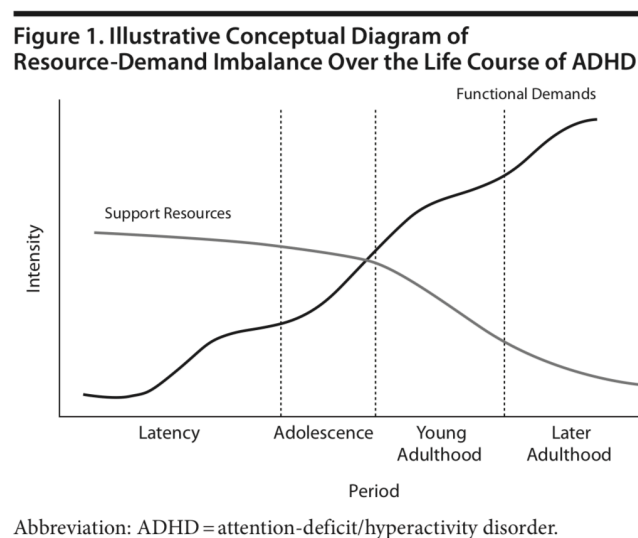


Figure 2.2: A diagram for the conceptual Resource-Demand Imbalance predicted for people with ADHD. It represents how resources and demands will emerge over a life course.
(Turgay *et al.*, 2012)

When getting diagnosed with ADHD as a child, a larger help system is available with customized teaching, medication, adaptive technology, informing and training of parents, and educational monitoring in school. As an adult this help system is limited, and usually merely consists of medication, and possible guidance/therapy if the patient requests it. As an adult, the demands one faces are greater than as a child, and these can be difficult to handle when the support resources have decreased. In work life the person is on his/her own, and there are no concrete measures available to help adults with ADHD (Turgay *et al.*, 2012).

When impairment occurs between support resources available and expected demands in an adult with ADHD this can lead to serious consequences. “Serious functional impairment in adults with ADHD includes educational and occupational underachievement relative to intelligence, increased psychiatric comorbidity, substance use disorder, unemployment, arrest and divorce versus normal controls” (Turgay *et al.*, 2012, p. 3). In other words, adults with ADHD are at risk for underachieving relative to their potential in several aspects of their lives. This can lead to a higher chance of unemployment, criminal behavior, issues with social relationships, and mental health issues. Consequently, many adults with ADHD may experience troubles in relation to their mental health.

In consideration of the examples above it is evident that being diagnosed with ADHD has a clear significance on a person’s life, both physically and mentally. Especially for adults with ADHD the help system is not sufficient: Their resources, both internal and external, are often not adequate with the demands they face in everyday life. Stimulants, which is the most used treatment form, is only efficient for 50% of the adult patients, and it is expressed a desire for non-pharmaceutical treatment options. There has been conducted some studies on the effects of peer support groups and online interventions in adults with ADHD, which has promising results (Pettersson *et al.*, 2017; Sehlin *et al.*, 2018).

2.3 Peer Support

As described in the introduction, peer support involves people in similar situations supporting each other. The support can be expressed by sharing personal experiences, knowledge, and emotional support (Mead, Hilton and Curtis, 2001; Solomon, 2004). Peer support is about both giving and receiving support. Peer support is not an new approach, and several well-known and well-established organizations, Alcoholic Anonymous for example (Alcoholic

Anonymous World Service Inc., 2018), are built on this approach. In AA people who have a problem with alcohol support each other through regular meetings where they encourage one another to become and/or stay sober. These meetings are run by and for alcoholics, hence the term peer support (Alcoholic Anonymous World Service Inc., 2018).

Peer support is used as an ingredient in different forms of treatment, such as in Alcoholic Anonymous (AA). Peer support meetings can be conducted in several different ways, but they have some main characteristics: The groups consist of people in similar situations and the main focus is to support and help each other. These groups often have a person with a counselor-type role. In some cases a professional guides the sessions, in other cases a trained peer has this responsibility, and in other situations there are no one with specific training (Faulkner and Basset, 2012).

Research on peer support shows several positive outcomes for the participants. By being a part of a peer support group one can get a shared identity with the group and one can develop new skills. These skills can be used to cope with disabilities (Faulkner and Basset, 2012). Further, peer support can increase confidence, improve mental health and wellbeing, as well as deal with stigma and discrimination (Faulkner and Basset, 2012). Peer support is proved to reduce symptoms and re-hospitalizations for patients with mental illnesses (Davidson *et al.*, 1999, 2006; Sells *et al.*, 2006). Mancini (2007) points out the role of peer support in enhancing peers' sense of managing their own illness. The mentioned benefits can help the person struggling in their everyday life and in their recovery. There is evidence from several studies indicating that providing support to peers can benefit the one providing the support as well: It can lead to increased social skills, not being dependent on others, and increased occupational functioning (Bracke, Christiaens and Verhaeghe, 2008; Salzer *et al.*, 2013).

One of the most significant advantages of peer support is the feeling of identifying with a social group. It is understood that by identifying with a social group one has the potential of increasing self-esteem and self-efficacy, and become more confident in oneself (McKenna and Bargh, 1998). To be connecting with others in similar situations that share similar concerns, illness symptoms and difficulties, can be comforting and create a greater sense of belonging to a group (Harvey *et al.*, 2007). This may also lead to more efficient recovery, better personal wellbeing, and social integration (Davidson *et al.*, 1999). Peer support offers several benefits, and is expected to serve an important role in future mental health treatment.

2.3.1 Online Peer Support

In recent years as the Internet has become a significant part of a lot of people everyday life, it has also become a promising part of peer support. Online peer support is based on the same principles as traditional peer support, but a key difference is that the participants use Internet technology to communicate, and therefore do not need to meet face-to-face. Such Internet technologies can be forums, chat rooms, private groups on social media (e.g. Facebook groups), and peer support websites. This type of technology offers both opportunities and challenges for peer support.

Mental health issues are a common matter in online peer support. The reasons for this may be that a lot of people experiences it easier and safer to admit their problems when anonymous on the Internet (O’Leary *et al.*, 2018). The fear of being labeled as mentally ill may be a reason why people do not want to admit mental health problems offline. On the Internet people have the opportunity to be anonymous, and thereby not being personally labeled. Mental health problems are a growing concern, and many people experience barriers when it comes to consulting professional help. These barriers can be the person being reluctant to seek help due to symptoms of their illness (e.g. social anxiety, or the feeling of hopelessness), or they may perceive the help system as inadequate. Other reasons why people do not seek professional help can be lack of trained professionals to meet the demand, the cost of receiving professional help, and stigma concerning mental health problems (Kazdin and Blase, 2011). Research on online peer support groups show great results and opens for new possibilities when it comes to mental health treatment. O’Leary *et al.* (2018) and Sehlin *et al.* (2018) argue that people struggling with mental health problems can have positive effect by being a part of an online peer support group, but that this should be a part of treatment, and not a replacement.

2.3.2 Benefits of Online Peer Support

Researchers have found several benefits of online peer support used in relation to mental health issues. People with metal issues can often experience symptoms that interfere with meeting new people (Dickerson *et al.*, 2001), which may cause them to avoid seeking help and support from peers or professionals. The stigma associated with their mental health issues may keep them from admitting their problems, which affects the help they possibly can receive. The Internet gives this group the opportunity to speak to peers without meeting them

in person, by appearing anonymously and only sharing at their own decision. Online peer support has the possibility to promote social connections among people with mental health issues by overcoming obstacles such as stigma and mental health symptoms (Highton-Williamson, Priebe and Giacco, 2015). People may experience greater freedom to share their experiences and feelings online, since they may feel they are not being labeled personally (Stephen *et al.*, 2014; O’Leary *et al.*, 2018).

In online text-based communication there is no need to understand non-verbal language (e.g. body language) and social cues, which for some may be a challenging part of social interactions. The lack of non-verbal language and social cues on the Internet can make people more confident (Highton-Williamson, Priebe and Giacco, 2015). Bargh and other researches acknowledge the importance of self-expressions, specifying that it can have positive effects on the challenging of stigma (Bargh, McKenna and Fitzsimons, 2002; Whitley and Campbell, 2014). On the Internet the participants possess more control than in real life. They can more freely choose their level of engagement and acquire more time to give their response. For some people with mental illnesses this lack of pressure can be a positive and relaxing experience (Stephen *et al.*, 2014). Alvarez-Jimenez *et al.* (2014) argue that online peer support can promote treatment engagement and prevent treatment drop out.

Naslund *et al.* (2016) emphasizes the possibilities for people with serious mental illness to connect with others with similar experience in online peer-to-peer communication. They stress the opportunity to openly disclose their own diagnosis, and to share positive experiences, stories and facts with peers. This could help challenge stigma, myths and misperceptions about mental illnesses (Naslund *et al.*, 2016). Another possible benefit with online peer support is the range of people one reaches – there are no demographic boundaries. Naslund *et al.* (2016) points out that online peer support has a wider demographic target group than traditional treatment since it may reach individuals who are reluctant to seek professional help. Since it is on the Internet, the only constraints are that the person needs to have Internet access and a computer or smartphone. Despite reduction in demographic boundaries, online peer support may cause a greater digital divide. Digital divide is a social issue represented by the gulf between those who have access to computers and Internet, and those who do not (Rouse, 2014). As mental health care becomes more Internet-based this divide may become greater. As described in this section, online peer support has great potential, but there are also some risks one needs to acknowledge.

2.3.3 Risks of Online Peer Support

Being able to be anonymous on the Internet can be both a benefit and a risk. By not knowing who you are talking to it can be uncertainty in the credibility of their stories and advices (Entwistle *et al.*, 2011). Some peers may promote unhealthy attitudes such as self-harm or other destructive behaviors (Ziebland and Wyke, 2012). Others may be hostile on purpose leaving mean comments and perform harassment, even though there is limited evidence of this (Highton-Williamson, Priebe and Giacco, 2015). People with mental health issues are a sensitive group, and this type of behavior can have critical impact.

There may be differences in experiences and skills between participants in peer support, which may create unrealistic expectations. This can further lead to more confusion and anxiety (Ziebland and Wyke, 2012). Not being able to conduct what others in the same situation have done may be perceived as demotivating and embarrassing. Some participants in online peer support systems may incur dependencies on online relationships (Chung, 2013). As explained in section 2.3.2, some people with mental health issues have symptoms making it difficult to meet and interact people. Being able to fulfill their need of social interaction on the Internet, might make them even more socially isolated and withdrawn from the offline world (Lawlor and Kirakowski, 2014). Naslund *et al.* (2016) debates that these individuals may already be severely isolated, and that online peer support may get them to have some social interaction and possibly give them a feeling of group belonging, which subsequently can motivate to recovery (Naslund *et al.*, 2016).

Naslund *et al.* (2016, p.18) specifies that overall, "... benefits of online peer-to-peer support appear to outweigh potential concerns, though sources of risk must be explored further to inform future research efforts". They further emphasize the elevated risks people with mental health issue experience in real life, for example by discrimination, abuse and violent crimes. They allege that "... the risk of connecting with similar others through social media should always be considered in light of existing susceptibility to risks posed by offline encounters" (Naslund *et al.*, 2016, p. 118). By this, it appears that with what we know today, the benefits and potentials of peer support exceed the possible risks and concerns it might offer. The risks of engaging in online activities should also always be compared with offline activities.

2.3.4 Guiding in Online Peer Support

As stated in Section 2.3, guiding in peer support conversations can be conducted in different ways, which also applies for online peer support. Baumeister et al. (2014) conducted a literature review analyzing (1) guiding vs. no guiding in online peer support conversations, (2) the level of guiding, (3) qualities of e-coaches and (4) asynchronous (not occurring at the same time, e.g. e-mail) vs. synchronous communication (occurring at the same time, e.g. instant messaging). The results showed that guidance is a beneficial feature of Internet-based interventions, but that features of e-coaches had no significant differences. Only one article examined (2) the level of guiding, and (3) asynchronous vs. synchronous communication, without any significant effects in either of the outcomes. Baumeister et al. (2014) stressed that guidance is an important feature, and should be used in online interventions if possible. They further noted that the most likely reason for implementing unguided intervention is due to the cost of it. O’Leary et al. (2018) conducted an experiment on peer support chats and anxiety. The participants were divided into two groups where one group was guided by prompts, and the other group was unguided. Results showed that both groups had reduced symptoms of anxiety, but the groups were perceived different. The guided group chats were perceived as “...deeply valuable for gaining solutions and insights, but provoked unwanted focus on troubles in some cases.” (O’Leary *et al.*, 2018, p. 10). The unguided group chats were seen as smooth and easy-going, but could be sensed as a distraction from their problems, rather than emotional support (O’Leary *et al.*, 2018).

Guiding an online peer support conversation has a significant value, and should be included if possible. If used correctly, guiding might have the potential to prevent some of the mentioned risks. In the HCI research field it could be interesting to examine different possibilities for guiding online peer support conversations.

2.4 Conversational Interfaces: Chatbots

Conversational interfaces (CI) have been around since the late 1960s when Weizenbaum (1966) developed ELIZA, the world’s first CI. Even though CIs have been around for decades, it isn’t until recent years they became popular. This is partly due to advancement in artificial intelligence (AI) and the boost in use of messaging platforms (Brandtzaeg and Følstad, 2018). The term *chatbot* is more widely used both in the academic and commercial

field. “Chatbots are conversational agents that provide users' with access to data and services through natural language dialogue” (Følstad and Brandtzaeg, 2017, p. 2). Other expressions such as conversational interface, conversational agent, and intelligent personal assistant, are referring to the same technology (Skjuve and Brandzaeg, 2018). This thesis will from now on refer to the technology as *chatbots*. Chatbots are often considered a text-based technology, but they can also be speech-based such as Amazon Alexa, Apple’s Siri and Google Assistant. Chatbots can have several different areas of use, where the most typical ones are customer services, personal assistant, content curation, and coaching (Følstad, Skjuve and Brandtzaeg, 2018).

2.4.1 Potential Areas of Use

Chatbots have become popular for several reasons: They are cost-effective, easy to use, effective and available for a large part of the population. Chatbots are an innovation when it comes to presenting and retrieving information. They give the user a new way to communicate with machines and businesses, which may result in positive outcomes if used correctly (Skjuve and Brandzaeg, 2018). Chatbots are an emerging trend, and there is still need for more research in this field to explore the full potential of chatbots.

In their article *Chatbots for Social Good*, Følstad et al. (2018) discusses the potential for chatbots to have a beneficial impact on society. Følstad et al. addresses three areas where they believe chatbots have the opportunity to improve social interactions. Firstly, chatbots can have the potential to strengthen people’s autonomy. That is, reducing digital divides and give better access to empowering services. According to Følstad et al. (2018), chatbots can possibly lower the threshold to interact with technology (as chatbots have an conversational nature and are widely accessible) and offer citizens services related to health, engagement and welfare. Secondly, chatbots can improve competence; the knowledge and skill one have to act in one’s own or others best interest. In this scenario chatbots can support education and training, for example by learning a language or new skills. They note that chatbots should preferably be part of a larger educational or training program. Thirdly, chatbots may contribute to social relatedness, meaning “... closeness and connectedness to significant others” (Følstad *et al.*, 2018, p. 4). This case might help oppose against social isolation. Bringing people together, for example in an educational or training program, can be an example of this (Følstad *et al.*, 2018).

2.4.2 Woebot

An example of a well-established chatbot that can be considered to have a beneficial impact on society is *Woebot*. Woebot was developed by researchers at Stanford University, and has contributed with important research on chatbots for mental health purposes. The development team consisted of psychologist, designers, data engineers, storytellers, and data scientists. Their aim was to develop a chatbot that could enter the role as a coach and teach users different techniques and theories they could apply in daily life to reduce symptoms of depression. These techniques and theories were built on cognitive behavioral therapy (CBT), a type of therapy focusing on improving mental health (Fitzpatrick, Darcy and Vierhile, 2017). A study was conducted to test if chatbots such as Woebot had the potential to deliver guidance based on CBT. 70 participants were recruited and divided into two randomized groups: One that had regular interactions with Woebot, and one that had no interaction with Woebot but given an eBook about mental health. At the beginning of the study period there were no significant differences in depression symptoms between the two groups. The group who used Woebot had a significant decrease in depression symptoms while using Woebot and in the time after. The control group had no significant decrease in depressive symptoms. This result indicates that chatbots can be an efficient way of delivering therapy (Fitzpatrick, Darcy and Vierhile, 2017).

Having a chatbot guiding online peer support conversations might be a possible solution to the potential risks mentioned in Section 2.3.3. The chatbot could supervise the conversation, and oversee how the participants behave towards each other. If participants use harmful language or express thoughts of self-harm, the chatbot could interfere by explaining how the participants should behave towards each other or by giving information about emergency hotlines. The chatbot could ban participants who chose not to follow the rules of the conversation. In section 2.3.4 it was recommended for online peer support conversations to be guided, however, it was not clarified how to perform this guiding. A possible solution could be a chatbot guiding such conversation, by proposing topics and facilitate for discussions that will be suitable and valuable for the participants. The guiding could be based on what professional psychologists or peer guiders would have done, which might lead to more realistic goals and expectations in the participants. There are a lot of possibilities for chatbot and peer support, which results in many research opportunities for the field of HCI and interaction design.

2.5 Design Considerations

When designing a prototype there are several factors one needs to take into account. In this section some considerations designers should consider regarding ADHD, mental health and chatbots are presented.

2.5.1 Designing for Adults with ADHD

McKnight (2011) developed a set of general design principles to be used when designing technology for children without excluding those with ADHD. The design principles consist of fifteen suggestions for consideration the developers should take into account. Most of these suggestions are coherent with ordinary design principles. For example, suggestions such as “Design materials so the layout is neat and uncluttered”, “Provide a ‘calm’ environment, with soothing colors”, “No decorations or distractions”, “Organize items in an orderly way” and “Use brief and clear instructions”. These design principles are common sense for many, as it is the norm to try to develop a product that brings out the essence of the product itself (McKnight, 2011).

Sonne *et al.* (2016) have developed some design strategies for developing assistive technology for people with ADHD:

- (1) Provide structure to facilitate activities** – For people with ADHD, structure can be very beneficial and they will often be more likely to complete a task if it has a predictable pattern.
- (2) Minimize distractions** – external distractions can occur extra disturbing for people with ADHD making it easier to lose attention. Therefore, it will often be beneficial to limit external distractions.
- (3) Encourage praise and rewards** – Praising and rewarding children and teenagers with ADHD is a core element in teaching them desired behaviors.

Both McKnight’s (2011) and Sonne *et al.*’s (2016) design principles includes several of the same strategies: When designing for people with ADHD it seems important to minimize distractions and surprises, making a predictable pattern, provide suitable information about their behavior (rewards or redirection), and be clear with instructions and languages. These design principles focused on children with ADHD, but it is anticipated that they will apply to adults since they experience similar symptoms. Additionally, a lot of the design principles are consistent with general design principles. It is therefore reasonable to assume that the same

principles can be used for adults with ADHD as well as children.

2.5.2 Designing for Mental Health

In their study *Design and evaluation guidelines for mental health technologies*, Doherty, Coyle and Matthews (2010) propose several design recommendations and evaluation principles for developing mental health technologies. Overall, their study focuses on knowing the user you are designing for, both considering who they are in context of their background and current issue, but also in order to engage them in the technology. They stress that it is important to use technologies familiar to the intended user, and make it clear that the data is secure. Another important factor they acknowledge is the ethical consideration, both international and local. The designer needs to consider the ethical consequences their product might face. The technology should build on requirements and traditions of mental healthcare setting, and accepted theories and models (Doherty, Coyle and Matthews, 2010).

When evaluating new technologies for mental healthcare the researchers accentuate the need to acquire ethical approval for all evaluations, and that one should evaluate the product in several distinct stages. Further, they described the need to evaluate with both potential users and therapist to make sure the product is suitable for all instances of its intended role. Additionally, the product should be evaluated in clinical practice to ensure that it will function in everyday practice (Doherty, Coyle and Matthews, 2010).

2.5.3 Designing Chatbots

An important factor to consider when developing a chatbot is what capabilities and limitations the chatbot will have. A problem that has occurred in relation to chatbots is user getting negative experiences due to the chatbot not being integrated well enough. When interacting with a chatbot the user expects the chatbot to be able to interact with them and answer questions related to its purpose. If the chatbot is not able to do this, the intention can seem meaningless. A lot of chatbots are not integrated well and can therefore not help the users with what they require. The chatbots may seem “dumb” and are unable to fulfill its purpose. This can give the user a negative impression of the chatbot, and may affect the users view on chatbots in general. Informing the user of what the purpose of the chatbot is and how the chatbot can offer help, is important. This will contribute to a more realistic expectation of the

chatbots abilities. A way to demonstrate this is for the chatbot to give examples and explanations for its abilities in the beginning of the conversations (Jain *et al.*, 2018).

Følstad, Skjuve and Brandtzaeg (2018) have proposed a typology to support developing chatbots for different purposes. They structure chatbot interaction design into two dimensions: *Locus of Control* and *Duration of Interaction*. *Locus of control* refers to who has the role as the leader of the conversation: Is it a chatbot- or a user-driven dialogue? Chatbot-driven dialogues are often common in scripted chatbots where there are little alternative paths for the user. Example of such chatbot-driven dialogues are Woebot and other chatbots who's purpose it to guide or coach (Følstad, Skjuve and Brandtzaeg, 2018). A reason why guide/coach chatbots have little alternative paths is due to the need for limitations and a clear framework for chatbots whose purpose concerns mental health and other sensitive topics. *Duration of interaction* concerns the human-chatbot relation, and if it is supposed to be long- or short-term relationship. A short-term relation is characterized with a one-time interaction, without storage of any data. In long-term relation chatbot and user will have multiple interactions, and the chatbot will often store data about user profile and previous conversations. "Some long-term chatbots exploit the duration of the relation to gradually present a rich set of content, such as a complex story or a game, or to gradually build skills and capabilities in the user, such as in educational, fitness or therapy chatbots" (Følstad, Skjuve and Brandtzaeg, 2018, p. 5).

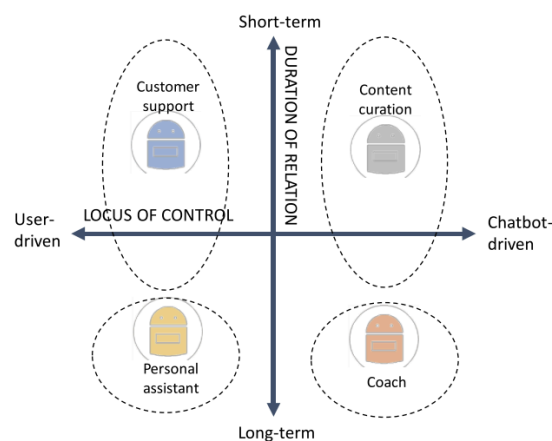


Figure 2.3: "A typology of chatbots with four example chatbot purposes located within the typology dimensions" (Følstad, Skjuve and Brandtzaeg, 2018).

As seen in Figure 2.3, chatbots whose purpose is some type of coaching will usually offer chatbot-driven conversation with the option of long-term relationships. Having a predefined program for the user, often with brief and recurrent sessions, can identify coaching chatbots. These sessions will often build on the previous one. The aim of these sessions is to gradually increase the knowledge and skills in the user. (Følstad, Skjuve and Brandtzaeg, 2018, p. 6)

Ciechanowski et al. (2018) examined the communication between humans and chatbots. *The uncanny valley* is a hypothesis suggesting that if a person has problems distinguishing between a human-like object and its human counterpart it will evoke negative affects. This was examined by studying how people felt when interacting with a standard chatbot (normal, text-based) versus an avatar chatbot (text-based, but including an animated avatar that moved according to the text produced). Their findings showed that the participant preferred the normal, text-based chatbot, and that experience was in general more pleasant (Ciechanowski et al., 2018). Some text-chatbots, such as IKEA's retired *Anna*, can also be perceived as "too human" which makes it difficult for the users to know how to correctly interact with them (Brandtzaeg and Følstad, 2018). The uncanny valley is an important factor to consider when developing chatbots. Related to the uncanny valley is a concept called "uncanny cliff". Uncanny cliffs occur when chatbots have an unexpected and sharp boundary between what they know and what they don't know (Grudin and Jacques, 2019). An uncanny cliff can for example be a pizzeria that has implemented a chatbot to answer questions about their opening hours and menus, but then the chatbot can't answer questions about the pizzerias location. This can lead to another problem, known as a "mission creep". It can be difficult to decide the chatbot's desired capabilities, and if developers always increase the capabilities of the chatbot, it can end up being "a resource black hole" (Grudin and Jacques, 2019).

Personality is expected to improve the user experience between humans and chatbots. As IKEA's *Anna* was considered too human, other chatbots have lacked a personality, which may indicate that there has to be a balance between robot and human factors and traits. It may seem like it is important that the chatbot appear as a robot, but with a distinctive personality. A research by Smestad and Volden (2018) indicated that chatbot personality had an impact on the user experience. They emphasize that the chatbot personality should be dependent on the purpose of the chatbot (it's job), the context it appears in, and its intended user group. Another study amplifies these findings. Thies, Menon and Magapu (2017) examined three different chatbot personalities in a Wizard of Oz trial. Their findings indicate that personality

factors such as the chatbot being reassuring, emphatic, non-judgmental, and funny, were perceived as positive features in a chatbot. It is important to remember that the personality traits must be considered in light of the chatbots context (Smestad and Volden, 2018).

2.6 Chapter Summary

ADHD affects 2,5% of adults in Norway, and stimulants, the most used treatment form, only benefit 50% of adults with the diagnosis. 75% of adults with ADHD have at least one other condition, which in many cases are affective disorders. It is expressed a desire for non-pharmaceutical treatment options. INTROMAT is developing an online self-help program that focuses on teaching adults with ADHD skills and techniques they can use in everyday life to deal with their symptoms. The INTROMAT team expressed a desire to explore the potential of implementing peer support technology in this program. There has been conducted research on online peer support (e.g., Takahashi *et al.*, 2009; Niela-Vilén *et al.*, 2014; Naslund *et al.*, 2016a; O’Leary *et al.*, 2018), which seems promising. Furthermore, research on Woebot (Fitzpatrick, Darcy and Vierhile, 2017), a chatbot delivering guiding, achieved promising results. This motivate for more research on the potential for chatbots to deliver guiding related to mental health.

Chapter 3

Methodology

As presented in Chapter 1, the primary research question of this thesis is:

RQ1: How can we design guided chat-based peer support technology accompanying an online self-help program for adults with ADHD?

This chapter introduces the methodology and methods used in this thesis, which includes the research through design approach, and methods for development and evaluation. The research conducted is mainly qualitative, which will be represented by the selected methods.

3.1 Research Through Design as a Method for Interaction Design Research in HCI

Zimmerman et al. (2007) explain the importance of specifying the role of interaction designers in HCI research, and have therefore created a methodology for interaction design research within the field of HCI together with a set of criteria for evaluating the quality of interaction design research contributions. This methodology is called Research Through Design (RtD) and will be referred to as RtD. Zimmerman et al. (2007) emphasize that designers should focus on making the *right* thing, meaning an artifact that should transform the world from its current state and to the preferred state, and thereby produce knowledge. This focus on the right thing is in contrast to the commercial market where the focus is on making the most successful product (design artifacts with focus on making a commercially viable product). The RtD methodology makes artifacts an important contribution to research.

Zimmerman et al. (2007) defines RtD as a research approach that uses methods and processes from the design practice as a valid method of exploration. In RtD, design activities have a crucial role in the generation of knowledge (Stappers and Giaccardi, 2013). These activities will include reviewing relevant literature, conducting user research, development of prototype, and then evaluating, framing and reframing the problem. This process will usually be done in several iterations (Stappers and Giaccardi, 2013).

3.1.1 Problem Identification

RtD has a strong focus on “wicked problems”; a problem that can never be accurately modeled. The term “wicked problems” was proposed as a result of researchers not being able to apply scientific methods to address certain dilemmas. These problems can be perceived as vague and ambiguous, making it difficult to solve by using traditional research methods.

Design research focus on problems that “... arises from a number of phenomena in combination, rather than the study of a single phenomenon in isolation” (Zimmerman, Forlizzi and Evenson, 2007, p. 496). Zimmerman et al. (2007) notes that interaction designer who engage in research should address wicked problems found in HCI. There is also a desire for the research contributions to demonstrate significant invention. That is, the design artifact should not be a clarification of a product that already exists, but a “... novel integration of theory, technology, user need, and context” (Zimmerman, Forlizzi and Evenson, 2007, p.499).

In this thesis an identified research problem is how to design a chatbot guiding a peer support conversation between adults with ADHD. There has been conducted research on chatbots and mental health (e.g. Woebot) and online peer support (e.g., O’Leary *et al.*, 2018), which are both promising HCI research areas. It will therefore be interesting to examine the potential for designing a chatbot guiding an online group conversation, and the potential to use a chatbot in peer support conversations in relation to an online self-help program for adults with ADHD. It is evident that this thesis researches important problems in HCI. There is a need for more research on designing chatbots, especially how to prototype scripts and chatbot experiences. This thesis will therefore explore the use of Wizard of Oz trials as a design and evaluation method in scripted chatbot development.

3.1.2 Evaluation

Zimmerman et al. (2007) have proposed a set of four criteria for evaluating an interaction design research contribution:

1. Process: When evaluating the contribution one should examine the rigor applied to the methods, and why specific methods were chosen. Rigor refers to the “the quality of being extremely thorough, exhaustive, or accurate” (Hobson, 2004). By this, the interaction designer should employ enough information about the research so the process can be reproduced, and it should also be provided a rationale for the selection of methods.

2. Inventions: The importance of an extensive literature review is stressed, to demonstrate that the result “...advances the current state of the art in the research community” (Zimmerman, Forlizzi and Evenson, 2007, p. 499).

3. Relevance: When discussing design and creation of artifacts there are no expectations about having different researchers producing the same results. Instead the focus should be on expressing the preferred state of the artifact, and explain why this should be considered as the preferred state.

4. Extensibility: The research should be described and documented in a way that will make it available for the community to leverage the knowledge produced by the work (Zimmerman, Forlizzi and Evenson, 2007). The artifact developed can become available for the public, research can be published in scientific journals, and findings can be presented at conferences or in other manners for the public to gain knowledge.

3.1.3 Why Use RtD?

RtD is a methodology tailored for HCI research as it acknowledges artifacts as viable contribution to knowledge and research. RtD facilitates for how interaction designers can contribute with important research inputs. In this thesis RtD was chosen as the methodology due to its focus on developing the *right thing* in form of new artifacts that produces knowledge, compared to other methodologies that lean toward commercially successful products.

3.2 Interaction Design - Beyond the Artifact

Zimmerman et al. (2007) described how interaction designers and the development of artifacts could contribute to HCI research. In its early years interaction design was often described as “designing interactive products” (Kaptelinin and Bannon, 2012). In their book, *Interaction Design – Beyond Human-Computer Interaction*, Preece, Rogers and Sharp (2015, p. 8) expand the definition by describing interaction design as “designing interactive products to support the way people communicate and interact in their everyday and working lives”. Kaptelinin and Bannon (2012) argue that the earlier definition of interaction design should also include designing of “... spaces for human communication and interaction”, expanding the definition even more by including that the final product can be a physical artifact, spaces, or processes. This master thesis has contributed with several outcomes: The concept of a

chatbot guiding a peer support chat, the chatbot scripts and requirements, and design considerations regarding chatbot for guiding group chats.

3.2.1 From User-Centered to Participatory Design

Interaction design has a clear focus on the user. User-centered design (UCD) is often described as the underlying fundament of interaction design. UCD focuses on the importance of analyzing user needs, understanding the context of actual and potential use of the artifact, and having an active involvement from users throughout the design process. The users are kept in focus throughout the development process, and may be included at several stages for testing, or similar. Kaptelinin and Bannon (2012) believe that interaction design must expand beyond the UCD, as UCD usually consists of designing and evaluating a specific artifact. Kaptelinin and Bannon (2013) stress that UCD has been, and will continue to be an important part of interaction design, but that one needs to develop new concepts and methods to support design and evaluation not only for artifacts, but also environments and habitats (Kaptelinin and Bannon, 2012).

Participatory design (PD) is an approach that is aiming to overcome some limitations of UCD, according to Kaptelinin and Bannon (2013). The differences between PD and UCD have become vague over the years, but there are some important differences in the level and type of user involvement. In PD users are an important part of the design and development process: The users either contribute to the design and content development process, or the users are in charge of the whole operation (Kensing and Blomberg, 1998). User involvement may vary from users playing a part in design decisions, to users creating the entire product from scratch. Users may disclose opportunities and challenges designers and developers would not have discovered otherwise. This differs from UCD where the designers and developers focus on keeping the user in focus by defining what they need, and similar measures.

Both the INTROMAT Adult ADHD case and this thesis follow a participatory design approach. The INTROMAT ADHD team includes users throughout their development processes by having an expert group consisting of adults with ADHD who are invited to design workshops and other suitable events. Additionally, several members of the INTROMAT team can be considered as potential future users as they work with patients with ADHD in their everyday life at work. Furthermore, several members in the team are domain

experts (neuropsychologist, clinical psychologist specialized in Internet interventions, clinical psychologist specialized in Goal Management Training) and have relevant expertise that is considered important in the development. In this thesis domain experts were present throughout the design iterations and were participating in evaluation of concepts and establishing requirements. The expert group was present at several design workshops and was participating in the evaluation.

3.2.2 Personas

Since this thesis follows principles from participatory design, RtD empathizes to conduct user research, and the expert group was not able to participate in all arranged design workshops, it was considered valuable to develop personas. Personas are detailed descriptions of real or hypothetical users that are applied in the design process to keep the user in focus (Blanton *et al.*, 2009). Creating personas made it easier to keep the user in focus throughout all stages of the design process. There are several ways of making personas, and various types of information designers can chose to include. In this thesis the description of the personas included general information (name, age, occupation, education, and similar), description of interest and hobbies, typical behavior, and frustrations and goals in connection with their ADHD diagnosis. The type of information that was included based on factors expected to potentially affect the users interaction with the prototype. It was developed three personas in this thesis to present a wider selection of potential users. The personas are presented in Section 4.3.1.

3.3 Ethical Approval

The INTROMAT ADHD case has its own ethical approval for workshops with volunteer adults with ADHD, which also applied to this thesis. In the design workshops it was only taken notes in a notebook and no personal information was recorded. Prior to the collection of data in the evaluation phases (WoOz trials) the Norwegian Center of Research Data (NSD) approved the project. See Appendix D for form of approval. The application to NSD included collecting data from two WoOz trials, including one consisting of volunteers (adults with ADHD) in the INTROMAT ADHD case. The data collected could be gathered through online chats, questionnaires, and focus groups.

3.4 Prototyping

Prototyping concerns a process of developing a type of user interface and designing interactive products. Prototyping does not necessarily mean that one has to develop software, since there are different types and degrees of prototyping techniques. Prototyping is often divided into low-fidelity or high fidelity prototyping, depending on methods and characteristics. Low-fidelity prototypes are usually limited in its functions, and offers restricted interaction opportunities for the user. Examples of low-fidelity prototypes are sketches, wireframes and Wizard of Oz trials (Rudd, Stern and Isensee, 1996). Low-fidelity prototypes are often cheap and easy to make, which makes them an effective way of expressing and exploring possible design alternatives. Low-fidelity prototypes are especially useful in the beginning of project, as it is a productive way of gathering requirements and analysis, and may function as a proof of concept. Low-fidelity prototypes also offer limitations: Their limited interaction opportunities may provide little error checking, and thereby important design decisions can be overlooked. Additionally, the interfaces are often far from the final product, and it may therefore be problematic to get a thorough and useful evaluation. High fidelity prototypes are more time-consuming and expensive to make, but are fully interactive and offers complete functionality, which facilitates better user feedback. Limitations include the prototype being more complex and expensive to change, so it will be ineffective as a proof of concept or to gather requirements (Rudd, Stern and Isensee, 1996).

MacKenzie (2013, p.128) stressed that designers and developers should only use as much “...time, effort, and investment as are needed to generate useful feedback and evolve an idea”. He further stressed that the goal of prototyping is to learn about strength and weaknesses of the concept, and identify opportunities and directions for the prototype. This thesis is inspired by MacKenzie’s (2013) view on prototyping, and chose to focus on low-fidelity prototypes as proof of concept and to gather requirements. In this thesis it was first developed design illustrations of what a peer support technology could be (sketches). Later it was conducted a Wizard of Oz trial (low-fidelity prototype) on adults without an ADHD diagnosis to test the concept and the scripts, as well as establish further requirements. Based on results from the first Wizard of Oz trial changes were made in the script and more requirements were settled. The renewed script and requirements were evaluated in a new, *alternative* Wizard of Oz trial (Q. Yang *et al.*, 2019) with adults with ADHD. It was considered to be useful to develop low-fidelity prototypes in this thesis, as the main focus was

to examine the potential for designing chatbots to be used in a group conversation, and to establish essential requirements for that purpose.

3.4.1 Design Workshops

Several workshops were conducted throughout the design process. Design workshops can be an effective way of making decisions that are carefully discussed and considered. Design workshops often consist of different types of people from different professions that are collaborating on the same project. Potential users of the product should also be a part of design workshops, as a way of making sure the process has a participatory design approach. The design workshop consisted of the author, a neuropsychologist, clinical psychologists with different expertise (Internet interventions and Goal Management Training), two psychology students, and a HCI researcher. An expert group consisting of several adults with ADHD was invited to the design workshops. One expert was able to participate in the first meeting (Section 4.2.1) and the first design workshop (Section 4.2.2).

A focus of the design workshops was to brainstorm about different ideas and possibilities. The brainstorming sessions lead to important discussions, which further promoted design decisions and requirements. The design workshops also worked as an arena where concepts and design suggestions could be presented and evaluated by both domain experts and the expert group. Design workshops were carried out with regular intervals throughout the project, as a means of showing progress and discuss the road ahead.

3.4.2 Conceptual Design

Conceptual design concerns articulating design ideas that meet the agreed upon requirements. Conceptual design are often made in the beginning of a project, where one describes an abstraction of what the purpose of the product will be, how people can interact with it, and what concepts are needed to understand how to interact with it. A concrete design is much more specific and will include information about details such as colors, images, icons, menus, and so on. In concrete design, alternatives are considered at every point (Preece, Rogers and Sharp, 2015). In this thesis it was developed two conceptual designs; first for a web-based application, and later a peer support chatbot. It can be discussed if the chatbot requirements and scripts apply as a conceptual or a concrete design, but due to its complexity in requirements it will be considered as a complex conceptual design in this thesis. A chatbot

personality, script and requirements do not have details in the same way as other types of interface, as many of the design choices will be predefined by the software it will be used in.

3.5 Evaluation

Since this project was an iterative process, several types of evaluations were conducted. The first evaluations were feedback from the expert group and domain experts on conceptual models in design workshops. Later in the project it was conducted a Wizard of Oz (WoOz) trial followed by online questionnaires on adults without ADHD diagnosis. Results were used to improve scripts and requirements, which was tested in a WoOz trial with adults with ADHD. The second WoOz trial was followed by online questionnaires and a focus group.

3.5.1 Wizard of Oz

Wizard of Oz (WoOz) is a low-fidelity prototyping method, which can be used to test requirements and evaluate concepts. In this method the user interacts with a piece of software as though interacting with the product in developing. The variation with this technique is that it in fact is a human operator that simulates the software's response to the user. This operator is referred to as the wizard (Dahlbäck, Jönsson and Ahrenberg, 1993). Several researchers have used this method in relation to natural language interfaces. Wizard of Oz experiments can be an efficient way of testing, for example, the potential for a new chatbot. With this method one can easily test the user preferences, scripts, personality attributes, response times, and other important factors the chatbot possesses (Dahlbäck, Jönsson and Ahrenberg, 1993; Q. Yang *et al.*, 2019). Gurdin and Jacques (2019) have written a paper that concerns a phenomena they refer to as *human bot* or *humbot*: people that, for various reasons, masquerades as a bot. One reasons for why people do this is to test chatbot traits in a WoOz trial, they state. They further state that WoOz trials depend on deception, as the participants should not know they were talking to a human before the trial is done.

An Alternative Wizard of Oz Method

Q. Yang *et al.* (2019) have proposed a way of sketching natural language processing (NLP) experiences. NLP experiences are technologies that use natural language as part of their interface, for example chatbots. There have been disagreements and uncertainties in the field of HCI on how to design and prototype for these experiences. To achieve an efficient way of

sketching NLP experiences, Q. Yang *et al.* propose an *alternative* WoOz method. The reason behind this alternative WoOz method is that there currently are limitations in the technology used to develop NLP experiences (e.g. machine learning and artificial intelligence technologies). Q. Yang *et al.* (2019, p. 8) therefore state that it is important to demonstrate how “...different intelligent features produce different kinds of errors, each of which can have different UX consequences”. Q. Yang *et al.* (2019) stress the importance of capturing these errors, and propose the use of *possible errors* (errors one can expect to happen in the real product due to limitations in current technology) when conducting WoOz trials for NLP experiences.

In this thesis the Wizard of Oz method was used as a proof of concept, to establish further requirements, to evaluate the script for the chatbot, and to explore the opportunities of using WoOz methods when designing a scripted chatbot. The first WoOz trial followed the standard layout, while the second WoOz trial followed the suggestion of an alternative WoOz trial by facilitating the trial for likely errors (Q. Yang *et al.*, 2019).

3.5.2 Sampling of Participants

In this thesis it was sampled participants for two different WoOz trials. In the first WoOz trial participants were recruited based on convenience sampling. That is, the participants were sampled due to their availability. Convenience sampling is a form of non-probability sampling. Non-probability sampling refers to participants not being selected using a random sampling method, which causes the probability of some units of the population to be more likely selected than others (Bryman, 2016). It was recruited nine participants to the first WoOz trial. The requirements concerned the participants being above the age of 18, and having access to Internet and a computer or smart phone. The participants were not required to have a diagnosis of ADHD as it was anticipated that adults with or without ADHD would have somewhat the same behavior and experience with the concept.

In the second WoOz trial it was desired to test the concept and the scripts on potential users, which were adults with ADHD. Participants were sampled based on purposive sampling. That is, non-probability sampling methods where participants are sampled due to their relevance to the research question, not on random basis (Bryman, 2016). The adults with ADHD who were asked to participate were already volunteers in the INTRMAT project. A domain expert in

the INTROMAT ADHD case sent an email to a list of voluntaries explaining the project and asking them to participate. Participants in the second WoOz trial received a gift card of 500NOK. A total of 5 adults with ADHD enlisted to participate in the trial, but one withdrew. Results from neither convenience sampling nor purposive sampling can be generalized to the general population (Bryman, 2016).

3.5.3 Online Questionnaire: Measuring User Experience

In both WoOz trial participants were sent an online questionnaire after each chat they had participated in. The intention was to register data regarding the user experience and their opinion of the concept, shortly after having conducted the chat session. The online questionnaires after ‘Chat 1’ and ‘Chat 2’ were almost identical, except that the questionnaire belonging to ‘Chat 1’ contained a question about previous chatbot experiences. The online questionnaires were a combination of statements the participants should rank on a likert scale from 1 (negative) to 7 (positive), and questions where they could elaborate. The questionnaires were used in both WoOz trials making it easy to look at similarities and differences between the two trials. Online questionnaires give a lot of possibilities in form of appearances and choices of layout. On the other hand it can be challenging to get participants to answer online questionnaires. In this thesis the researcher has to remind several participants to answer the online questionnaires. The questionnaires can be viewed in Appendix A.

3.5.4 Focus Group

After the adults with ADHD had finished the WoOz trials they were asked to participate in a focus group concerning their experiences. Only two participants qualified to participate, due to the others not attending the whole trial. According to Bryman (2016) the typical focus group size is between six to ten members. When using focus group to collect qualitative data it is common to have several groups to collect data from, since one group will often not meet the researcher’s needs. In this thesis it was considered to be sufficient with one focus group, since this was only one part of the data collected, and it was only necessary to interview adults with ADHD who actually participated in both ‘Chat 1’ and ‘Chat 2’.

Bryman (2016, p. 507) states that “Clearly, the moderator [of the focus group] has to straddle two positions: allowing the discussion to flow freely and intervening to bring out especially salient issues, particularly when group participants do not do so”. The role of the moderator is

an important, but difficult, part of conducting a good focus group. The moderator in the focus group had defined a small number of generic questions beforehand, which were used to guide the conversation to desired topics. Topics included the participants' perception of the chatbot, their general perception of the concept, and design considerations. There are some limitations with focus groups. Firstly, focus groups can be difficult to organize since one has to make the participants agree to participate, as well as getting them to turn up at a particular time and place. Further, the recording one collect are time-consuming to transcribe, and the data can be difficult to analyze. Participants can speak at the same time, or interrupt each other.

Participants can also influence each other: Some participants may be dominating the conversation, while other may appear reluctant to answer due to personal differences. The moderator can of course try to moderate certain problems, but it may be difficult at times (Bryman, 2016).

By conducting a focus group after the WoOz trials the researcher got a more thorough understanding of the participants experience with the concept and the simulated chatbot. Participants were able to express and discuss different aspect and opinions of the chatbot personality and role, the script, and the concept in general, compared to only getting to answer on the online questioner. The focus group is consistent with the participatory design approach, as the participants and future users got to influence future design considerations.

3.6 Chapter Summary

This chapter introduced the research framework chosen to answer the research questions of this thesis. Further, selected methods for development and evaluation was described and elaborated why they were chosen, including their limitations. Table 3.1 illustrates the different phases, purposes, methods, sample sizes and data collected in this thesis.

Phase	Purpose	Method	Sample	Participants	Data
Preface	Discuss potential peer support technologies	Meeting	N=8	The INTROMAT team + expert group	Notes
Conceptual Design	Discuss possible solutions and specify requirements	Design Workshops	N=7-8	The INTROMAT team + expert group	Sketches and notes
Conceptual Design	Develop the chatbot script together with domain experts	Design Workshop	N=7	The INTROMAT team	Script and notes
Wizard of Oz Trial 1	Evaluate user experience and script in 'Chat 1' and 'Chat 2'	Log and questionnaire from 'Chat 1' and 'Chat 2'	N=9 N=9	Adults, convince sampling	Chat log and questionnaire answers
Wizard of Oz Trial 2	Evaluate user experience and script in 'Chat 1' and 'Chat 2'	Log and questionnaire from 'Chat 1' and 'Chat 2'	N=4 N=2	Adults with ADHD	Chat log and questionnaire answers
Wizard of Oz Trial 2	Get a more in-depth understanding of the user experience	Focus Group	N=2	Adults with ADHD	Voice recording, which has been transcribed

Table 3.1: An overview of the different phases, purposes, methods, samples and data collected throughout this thesis.

Chapter 4

Conceptual Design and Establishing Requirements

This chapter addresses the design of a conceptual peer support chatbot from iteration 1-3

First Iteration: The project started with a meeting, followed by a literature review and sketching of different alternatives. The alternatives were presented at a design workshop, which lead to agreement on several aspects of requirements. A conceptual design was expressed and evaluated by the INTROMAT team and expert group.

Second Iteration: Development of personas. Designing of a low-fidelity prototype: A web-based chat system designed in Adobe XD. The prototype was presented and discussed with the INTROMAT team and the expert group.

Third Iteration: Based on literature review and feedback on last iteration: A conceptual design of a peer support chatbot was made and discussed with the INTROMAT team. A new design workshop that focused on the chatbot conversation was held, leading to a scripted prototype.

There are some roles that are repetitive throughout the design process that might be beneficial to define. Further, the author of this thesis has three roles throughout this project: author, researcher and the wizard in the Wizard of Oz trials.

INTROMAT: INTROMAT, INtroducing personalized TReatment Of Mental health problems using Adaptive Technology, is a project that intends to improve mental health by the use of innovative technology.

INTROMAT ADHD Case: A sub project of INTROMAT that is developing an online self-help program for adults with ADHD.

The INTROMAT Team: A term used to referring to the group of people working on INTROMAT's adult ADHD case.

Domain Experts: Specific members in the INTROMAT team with different types of expertise in psychology.

Expert Group: A group of volunteering adults with ADHD in the INTROMAT project.

Expert: A member of the expert group.

4.1 Tools

4.1.1 Adobe XD

Adobe eXperience Design (XD) is a prototyping software produced by Adobe Systems. By using Adobe XD the designer can make realistic looking prototypes of apps or web-based services (Adobe, 2019). Adobe XD was chosen as a prototyping tool in this thesis due to its well-made interface, and its fitting functionalities. Adobe XD was used in the first and second design iterations to express design alternatives and a conceptual design.

4.1.2 Slack

Slack (Slack, no date) is a web-based chat service. Slack was the software used to conduct the chats in the Wizard of Oz trials. In Slack one makes an online *workspace*, that is, a group in Slack where all desired people are included. Inside this workspace, it can be created distinct *channels*, where only certain people are included and have access. In relation to this thesis this was considered efficient, since one could have all participants in the Wizard of Oz trial in the same online workspace, but in separate channels. Slack facilitates for workspace awareness in their chat. When a user is typing it is visualized to the other participants in that chat channel. This was considered an important feature, as it would be useful for both the participants and the wizard to know when others were typing.

4.2 First Iteration

The aim of the first iteration was to develop a conceptual design for a peer support system and produce essential requirements. To accomplish this it was emphasized to achieve a thorough understanding of the project and its established requirements, as well as attain an accurate comprehension about related research themes. The project started with a meeting with the INTROMAT team and an expert. The INTROMAT team included researchers from several different fields: Neuropsychology, clinical psychology with expertise in Internet interventions, clinical psychology with expertise in Goal Management Training, and a HCI researcher. The first iteration ended with a design workshop.

4.2.1 First Meeting: Discussion of Possibilities

One of the main intentions with the first meeting was to examine the opportunities for implementing peer support technology in the program, and which requirements one had to take into account. There were no specific ideas envisioned, so the assignment was particularly free and undetermined, but both the expert and the INTROMAT team expressed a positive opinion towards implementing peer support technology. The INTROMAT team mentioned that the peer support technology should be a part of the online self-help program, and it should concern the needs of adults with ADHD. After the first meeting a literature review concerning ADHD, technology, peer support, and online interventions, was conducted. Based on the literature review a PowerPoint presentation with potential requirements presented against each other was made. Some possibilities were sketched up in Adobe XD to create visualizations. This presentation was then demonstrated at the first design workshop.

4.2.2 Design Workshop: Specifying Requirements

The aim of the first design workshop was to discuss possible solutions and specify requirements for peer support technology. This workshop was important to create a basis for the upcoming design iterations. An expert and the INTROMAT team participated in the workshop. The PowerPoint presentation included several aspects of peer support technologies that needed to be determined. These aspects were expressed by describing possible solutions and illustrated with sketches and models (e.g. Figure 4.1 and Figure 4.3). The presentation entailed brainstorming and discussion amongst the participants, leading to several requirements being established:

The purpose of peer support

It was important to define why and how peer support could benefit adults with ADHD who participated in the online self-help program. There were two fundamental possibilities: It could support them on a general, emotional level, or it could be closely linked to the self-help program. Based on a discussion it was decided to be most beneficial for peer support to concern the exercises and skills, since that is the essence of the self-help program. The way this could be done was by having regular conversations between peers that were related to topics and exercises in the program. This is supported by literature, which indicates that online peer support can have a positive effect on learning, as well as coherence to self-help

programs (Section 2.3.2). Figure 4.1 illustrates the different purposes for the peer support technology.

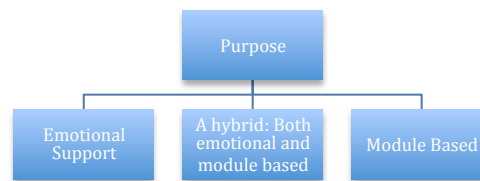


Figure 4.1: A visualization of requirement choice regarding purpose of peer support.

Chat Composition

There are several ways online peer support could be managed. The three practices discussed in this case were: one-to-one chats, group chats, and forums, as visualized in Figure 4.2. The two types of chats could vary from consistent to random interlocutors. One-to-one chats were excluded based on the fact that it was desired a peer system where several participants would support each other. There is always a risk of participants dropping out of the program, or forgetting their appointments. Since chats require more involvement and participation than a forum, the idea was that group chats would be most valuable in relation to the purpose of the program. It was further considered to be beneficial to have regular groups since one will develop familiarity, the same basis for the conversations, and solidarity with the other participants. Domain experts from the INTROMAT team expressed that with offline self-help programs (e.g. Goal Management Training) it is common to have the same group of participants following each other throughout the program. Findings from the literature review also support the decision of group chats with regular participants. In a recent study about different online channels for social-support groups, D. Yang et al. (2019) found that compared to public forums, private channels (similar to chats) lead to stronger emotional and informational support, as well as a stronger reciprocity effect between the participants.

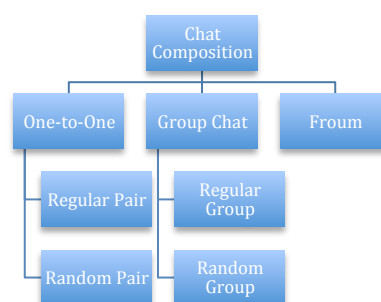


Figure 4.2: A visualization of the requirement choice regarding group composition.

Structure of the Chat

There can be different types of structure in a peer support chat, as described in Section 2.3.4. It could be a very structured and fully guided chat, meaning that the participants were guided by a person/technology that outlined the structure of the whole conversation. Or it could be a “freely” chat without any specific structure or guiding, where all the participants have responsibility for the conversation. It is also possible to have a hybrid of these two. Figure 4.3 demonstrates possible differences between a structured and a semi-structured chat. Since it earlier was decided that the conversations would have a specific purpose and deal with topics related to the program, it was concluded to be more natural to have some structure and guiding in the chats. The chats should have themes to be discussed, but it was not specified directly how. Since it was decided to have a group chat with some kind of guiding it indicates that the chat will be synchronous or semi-synchronous.

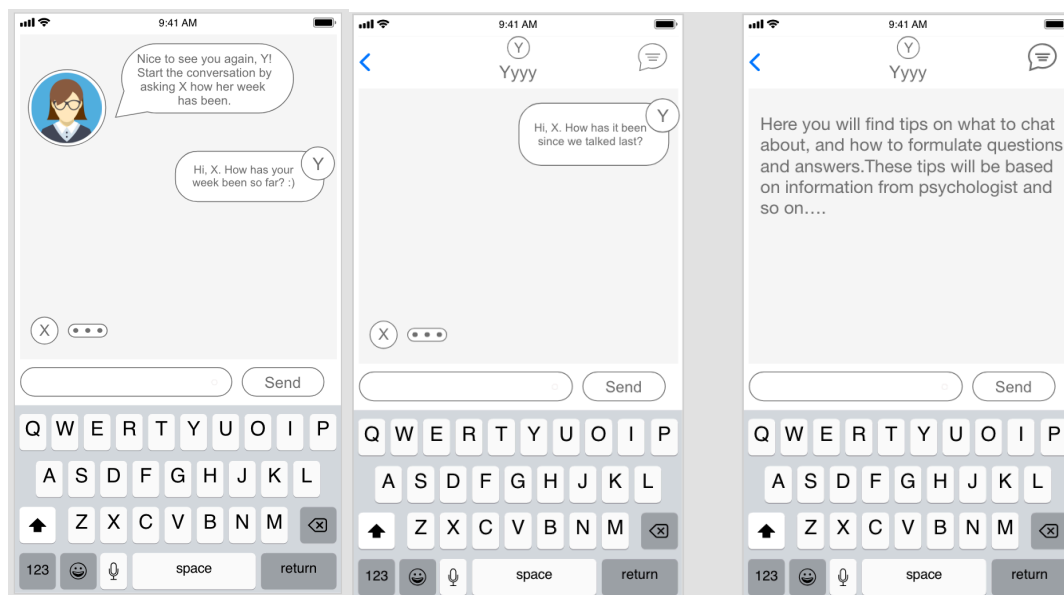


Figure 4.3: The left chat illustrates a fully guided chat, while the two right ones demonstrate a semi-guided chat

Chat Regulations

There was also a need to determine how chat sessions should be organized. Should the chats be open 24/7, or only at certain times? It was stated as important for the INTROMAT team to know what was talked about in the chat, especially in regards to negative influence or harmful intentions. Because these sessions would need to be supervised it was decided that the chats should be held at scheduled times. This was also stressed to be a possible benefit in the way of it being a routine. It was also decided that the technology should be web-based, and not an

app. The reason behind this was to simplify the process for the participants by not having to download an app. The online self-help program will also have important security features (such as login with Bank-ID) that naturally apply for the peer support system to have as well.

Personal Information

The question concerning the level of anonymity in the chats was raised. In chat programs participants can be anything from completely anonymous to entirely public to each other. Three mockups presenting three different types of user profiles were presented to create a better understanding of how possible information descriptions could be. This is illustrated in Figure 4.4. Some domain experts thought that one should not be entirely exposed to each other due to personal risks. The expert stressed the importance of not being embarrassed of the diagnosis in order to do well in daily life, and thereby meant that the profiles should not be fully anonymous. The HCI researcher noted that group therapy in general is not anonymous, and that this chat could be compared with group therapy on certain aspects. After a thorough discussion the end conclusion was that the user can to some degree chose which information they want to share, and it is recommended to share for example first name and some general personal information (gender, age group, hobbies) to help contribute to a better group dynamic. By using first names it was believed that the conversation would seem more natural. But as stated above, these are only recommendations, and the participants themselves decide the level of anonymity.

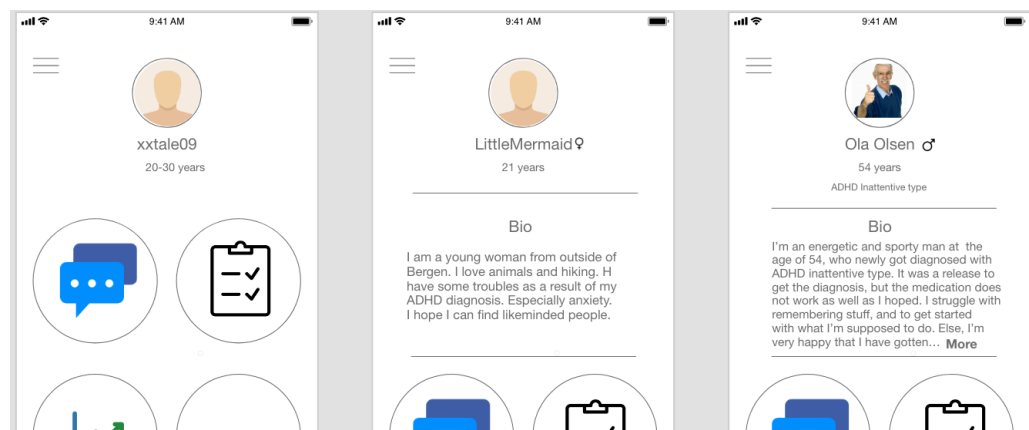


Figure 4.4: A visualization of three different user profiles with various degree of personal information. The far left represent a fully anonymous profile, the middle a quasi-anonymous profile, and the far left a fully open profile.

4.2.3 Summary

In the first design workshop it was decided that the peer support system should be a group chat with a consistent group composition, and with scheduled sessions. The topic should primarily concern the self-help program, and the chat conversations should be semi-guided or guided to ensure this. The participants should be recommended to share some personal information to contribute to a better group dynamic, but it is not required. In the end, the user is personally responsible for which type of information they want to share. The domain experts stress the necessity of peer support in their online self-help program, as it has the opportunity to give the participant a sense of group belonging, even though it is online. This correlates with research in Section 2.3.2, which expressed the usefulness of peer support in interventions.

4.3 Second Iteration

Before starting the design in the second iteration, it was created three user personas to help keep the user group in focus under further development processes. In the second iteration the aim was to develop a low-fidelity prototype, which in this case was a web-based chat system designed in Adobe XD. The prototype was presented and discussed with the INTROMAT team and the expert.

4.3.1 Personas

Based on a thorough literature review on the topic, as well as an analysis on interviews with adults with ADHD (Østheim, 2011), three personas were made. The reasons for making three personas were to include different subtypes and problems common for adults with ADHD. The outline for the personas was based on literature in Section 3.2.2. They were printed out and kept as inspirations through all design iterations. All personas can be seen in Appendix B. Figure 4.5 shows one of the personas.

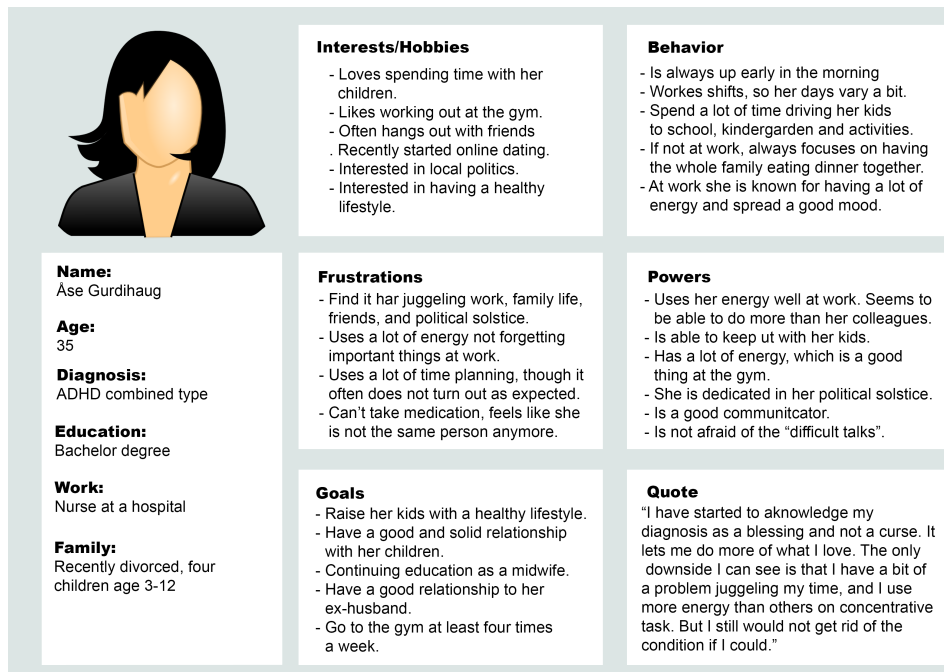


Figure 4.5: Persona 1 – Woman, 35 years old with ADHD Combined Type.

4.3.2 Advancement of the Idea

Based on the requirements and discussions from the first design workshop, the idea of a chatbot guiding the peer conversation emerged. Research from the literature review (Fitzpatrick, Darcy and Vierhile, 2017; Følstad *et al.*, 2018) supported this idea, and indicated that this is an area in the HCI field that needs more research. In the previous design workshop it was decided that the chat should be somewhat guided, without specifying *how*. To demonstrate the new concept to the INTROMAT team a sketched prototype of a chatbot that guided a web-based peer support conversation was made in Adobe XD.

4.3.3 Design Workshop: A Peer Support Chatbot

When developing the first prototype for a peer support chatbot, the main goal was to demonstrate the idea and possible requirements to the INTROMAT team. This prototype contributed to a discussion in the workshop, which lead to a settlement about specific features the chatbot could have and how it could operate. Compared to the earlier idea it would be important to develop a script for the conversations, and give the chatbot a personality and features to promote desired achievement. As stated in Section 2.5.3, chatbots which role is to guide usually has a chatbot-driven and scripted conversation style. The same requirements as

decided in the first design workshop still counted as requirements for this prototype, but there would be a need to address more requirements with the advancement of the idea.

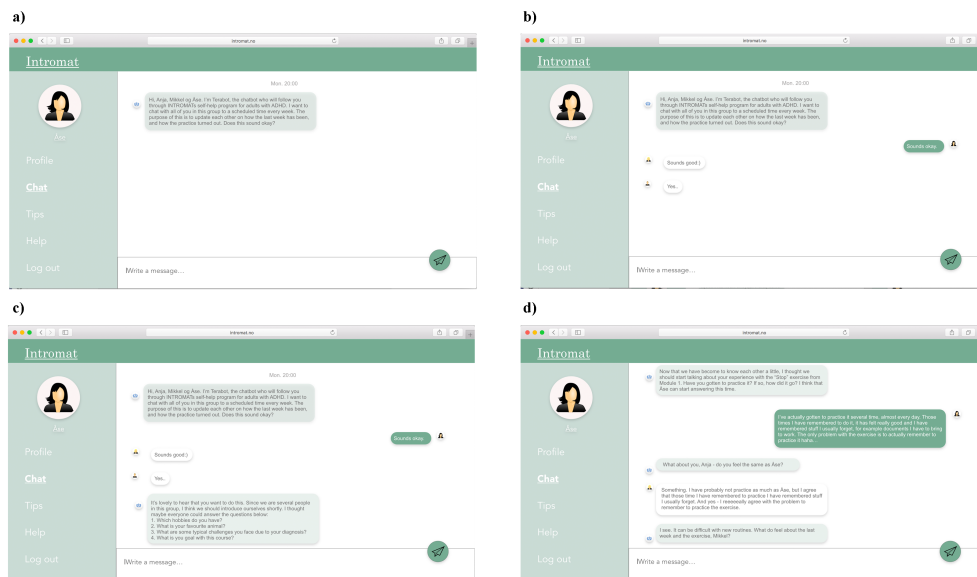


Figure 4.6: Segments of the prototypes designed in Adobe XD and presented to the INTROMAT team.

A prototype that illustrates the concept was developed in Adobe XD, and was demonstrated for the INTROMAT team in the second design workshop. No expert (adult with ADHD) was able to join this workshop. Figure 4.6 shows segments of the prototype that was presented. The focus was mainly on the concept itself and the type of conversation the chatbot should promote. The INTROMAT team agreed that the conversation should continue to revolve around the associated exercise in the program, together with several other requirements:

- The chatbot should, as in Figure 4.6, guide the participants through the conversation.
- The chatbot should behave professionally, but should also be empathic, non-judging, and have a sense of humor.
- The users need some information on the purpose of the chat, and how they are expected to behave.
- The chatbot should make sure that all participants take part in the conversation and are included. To manage this, the chatbot can for example ask specific participants specific questions, and encourage participants to give each other feedback.

- The chatbot should also have the responsibility to introduce new themes to the conversation and make sure that there is a certain flow in the chat. That is, that all participants are answering and involved.
- The chatbot being made was just a prototype, so it was decided that it should be concentrating about one specific exercise from the self-help program. The selected exercise was “The Stop Exercise”. “The Stop Exercise” is a cognitive exercise where the goal is to stop for a short time to gather ones thoughts. Being focused on only one exercise was considered to be an efficient way to test proof of concept.

At the end of this design workshop some time was used to discuss how typical group therapy was conducted. This information contributed to the layout and development of the script. Several of the domain experts had experience from managing group therapy for adults who participated in Goal Management Training (GMT) (Levine *et al.*, 2011) from an intervention program that treated ADHD patients with reduced inhibitory control and executive functions. The chatbot was not supposed to conduct group therapy, but it was considered useful to adopt some components from it. It was considered especially interesting how facilitators got participants to be involved in the conversations and how the guiding took place. The domain experts with expertise from GMT expressed how they conducted an introduction round in the beginning of the program to create a sense of group belonging. How participants were asked to present themselves varied, but participants would often repeat similar information as those presenting before them. The facilitator would further contribute to the conversation by including participants, introduce topics and help keeping a flow in the conversation.

Based on the approval of the concept and the new requirements, a new design process started. It was considered important to have a solid script for the chat conversations, which would be the main focus for the next iteration. Scripts are common in chatbots that guide and/or have a therapeutic role, as it is important to know what type of information and advice the chatbot gives. Well-designed scripts were considered valuable both due to user experience, but also since adults with ADHD can be considered a vulnerable group and it is crucial to have evaluated and considered all possible outcomes of the conversation.

4.3.4 Summary

In the second iteration it was specified that the guiding in the peer support chat should be lead by a chatbot. The chatbot should be responsible for including all participants in the conversation, introduce new themes, and keep a flow in the conversation. The conversations should consider specific exercises from the online self-help program. This thesis took base in “The Stop Exercise”. The chatbot has to follow a script, as the conversations need to be carefully planned. Some information from GMT meetings was used as inspiration in the layout of the script.

4.4 Third Iteration

In the third iteration the intent was to develop the chatbot script and conduct a trial based on this. Information from the previous design workshop was used to draft a script for the chatbot conversation. This draft would work as a base in the new and upcoming design workshop. In this design workshop the chatbot script was further changed and improved, and it was promoted information and tips that would be needed to take into consideration. The script was further revised after the design workshop, and was later sent to the domain experts for a read-through and evaluation. The evaluated script was then used to perform a Wizard of Oz trial.

4.4.1 Design of the Chatbot

As decided in the previous design workshop, and supported by literature from Section 2.5.3 the chatbot would have a chatbot driven dialogue, as the chatbot would be guiding the conversation. The interaction between participants and chatbot would be long-term, since the participants were supposed to have several interactions with the chatbot throughout the online self-help program. The chatbots’ personality was another important consideration. As studies show it would be important to make the users understand that it was a chatbot and not a real person they chatted with, as well as what the chatbots limitations were. The chatbot was given a name and a profile picture to demonstrated this. Figure 4.7 shows the chatbots’ profile picture, which was used throughout the trials in this thesis. The chatbot was given the name *Terabot*, indicating to the user that this was a chatbot and not a real person. *Tera* refers to *therapist* (*terapeut* in Norwegian) and *bot* refers to *chatbot*, hence *Terabot*.



Figure 4.7: The profile picture used for the chatbot, Terabot.

The script was created to concern *The Stop Exercise*. The chatbot should appear professional, but also have a distinct personality. It was decided that the conversation should begin with the chatbot explaining its role in the program, its limitation and how it expected the participants to behave. This was followed by the chatbot explaining the purpose of that specific chat, before moving on to the introduction round. This resulted in a considerable amount of uninterrupted text, but it was deliberated to be information that was important to be presented early in the conversation. The chatbot started to introduce itself before asking the participants to answer a couple of questions about themselves (name, age, what they struggle with in everyday life, and what they hope to learn in this program). In the rest of the script the conversation went on by having the chatbot asking specific participants questions about the exercise, and asked other participants to give feedback and to give their experience. The questions regarded understanding of the exercise, which situations to use it in, and so on. The example below uses the personas to illustrate this.

...

Chatbot: How do you understand the purpose of this exercise, Åse?

Åse: **Describes her understanding of the exercise**

Chatbot: I see. Have you understood it in the same way, Anja?

Anja: **Describes her understanding of the exercise**

Chatbot: *(Tries to interpret if they have the same understanding or not, e.g. Anja saying "Yes...")* I can see that Åse and Anja have somewhat of the same understanding of the exercise. How did you understand the exercise, Mikkel?

Mikkel: **Describes his understanding of the exercise**

Chatbot: "Good thinking, all of you! The way I understand the exercise it is about calming down and gather one's thoughts. Often when we are doing something it is very easy to go into autopilot, and then one can forget what one is doing or supposed to do. That happens to me, at least!"

...

The example above illustrates the first solution to include all participants, as well as how the chatbot could make it possible to guide a group chat. In the introduction round the chatbot would specify that participants will be asked specific questions they are supposed to answer, and that they should not talk/write without being asked. The reason behind this is due to the chatbots' limitations. If participants wrote without the chatbot being prepared and without context it could make the conversation chaotic and lead to errors. In the example above the chatbot asks the participants one by one to answer a question and give feedback, so it will be easier for the chatbot to keep control of the conversation. By asking the participants if they have understood it similar to the others, all participants get to answer every question, as well as participants getting feedback on their answers. The idea was that this could lead to a feeling of group belonging. In the example above the chatbot does not need to be particularly smart. There is no need for the chatbot to understand what's being said, but it would be advantageous if it understood when the answers were affirmative or denying. Considering the flow of the conversation, the assumption was that the chatbot would ask a specific participant a question, get an answer, and then continue on to the next question/answer. This example has several faults that could create problems; for example by participants writing several answers to the same question, or if they answered with a question. That is potential problems that need to be taken into consideration. There is also the possibility of participants not answering at all. The first drafted script entailed several flaws and challenges, but that was something that continued to be refined in the design workshop.

4.4.2 Design Workshop: Improving Chatbot Script

The design workshop started with a review of the drafted script, followed by a discussion. The domain expert in Goal Management Training (GMT) argued that it would be better to divide the chat into two: One chat before the participants started practicing the exercise, and one after they had practiced. The chatbot could then contribute to the participants having a thorough understanding of the exercise beforehand, and then letting the participants explain their experiences after. By having peer support in such scenario the thought was that the participants would be more likely to get the true understanding of an exercise, as well as being able to learn from others. It could also be comforting to see others struggle with the same as you, or motivation to see others understand the exercise in a different way. It could also be motivating to practice an exercise when one knows that one will be discussing it with

peers later. The domain experts stressed that this is the way they would have done it in GMT: You build a foundation and expectation towards the exercise beforehand, then practice the exercise, and then talk about the execution and experience after.

Regarding the participants introducing themselves, it was discussed how this could best be done. Some questions can be difficult to answer and could possibly lead to more challenges than benefits for the participants. The participants should be allowed to share the type of information they themselves wanted to. It was therefore decided that the chatbot would start by introducing itself, before asking the participants to do the same. The chatbot would not ask about specific information, but when presenting itself it might influence the participants to give the same type of information, if they wanted to. The chatbot introduces itself with name, occupation, hobbies and favorite animal.

It was also expressed opinions that it seemed a bit much when the chatbot asked the same question to all participants every time. Based on a discussion it was decided that the chatbot should ask one question that related to all the participants and where all the participants were expected to answer. When all participants had answered the chatbot would move on to the next question. An example is demonstrated below. The design workshop was further used to specify what type of information one desired in 'Chat 1' and 'Chat 2'.

Based on the new knowledge from the design workshop the two scripts were further processed before they were sent to the domain experts for a read-through. When the scripts were reviewed and approved they were ready to be used in a Wizard of Oz trial. The scripts (Norwegian) can be seen in Appendix C.

...

Chatbot: In which situations in your everyday life can you imagine using this exercise?

Åse: **Describes when/where she can imagine using this exercise**

Anja: **Describes when/where she can imagine using this exercise**

Mikkel: **Describes when/where he can imagine using this exercise**

Chatbot: It is exciting to read about what situations you are thinking of using the exercise. I've learned some new ones I hadn't thought of before.

...

4.4.3 Summary

Based on information from the design workshop the chatbot script related to “*The Stop Exercise*” was divided into two distinct scripts: One that worked as an introduction to the exercise, and one that focused on the execution and experience with the exercise. The chatbot would ask one question that all participants were expected to answer. It was further decided what type of information the participants needed in the chat, how they should be asked to introduce themselves, and what type of question the chatbot should ask about the exercise in ‘Chat 1’ and ‘Chat 2’.

4.5 Chapter Summary

This chapter presented design iteration 1-3. At this point in the thesis it was collected several types of data based on literature review, meetings and design workshops. The concept had developed from designing a web-based peer support chat, to a web-based peer support chat guided by a chatbot. This thesis focused on the design of the chatbots’ requirements, personality traits, and developing scripts. The basis for the scripts was “The Stop Exercise”. Based on expertise from Goal Management Training, it was determined that it would be efficient to conduct one chat before the participants started to practice the exercise, and one chat after they had practiced for some time. It was developed two scripts that would be tested in a Wizard of Oz trial in the next phase of this thesis.

Chapter 5

Wizard of Oz Trials

This chapter addresses two Wizard of Oz (WoOz) trials. As stated in Section 3.5.1, WoOz is a low-fidelity prototyping method, which can be used to test requirements and evaluate a concept. In this thesis WoOz trials were chosen for two reasons: Firstly as a way of evaluating the concept, the scripts and to specify requirements. Secondly, it was considered valuable to examine the opportunities of mentioned method when designing scripted chatbots.

First evaluation: The first Wizard of Oz Trial, accompanied by online questionnaires, was conducted following the design choices made in iteration 1-3. The WoOz trial worked as a proof of concept and an evaluation of the chatbot personality, scripts and requirements.

Second evaluation: Some adjustments to the scripts and requirements were made. The second Wizard of Oz trial was inspired by the study of Q. Yang et al. (2019), whom proposed an alternative WoOz trial by entailing likely errors. The participants were adults with ADHD. The trial was followed by questionnaires and a focus group.

5.1 First Wizard of Oz Trial: Evaluation of Scripts and Proof of Concept

The Purpose

The first Wizard of Oz trial was conducted to evaluate the scripts and requirements, as a proof of concept that using a chatbot to guide peer support conversations had potential, and to explore the benefits and limitations of the WoOz method in this scenario. The participants were adults without any prerequisite of being diagnosed with ADHD, as it was not considered to be necessary at this point in the project. The participants were divided into three groups consisting of three participants, which were the groups used in the chat sessions. The participants were separately introduced to “The Stop Exercise” in form of a document describing the exercise in detail, in addition to the participants being allowed to ask the researcher for clarification. In a real life scenario the exercise would be a part of the self-help program where the users would have gotten a more thorough introduction to the exercise in form of text, video and assignments.

The Wizard

The author of this thesis had the role as *the wizard*. The author took the role as the chatbot by following the scripts made, but also decided that some improvising could generate useful information for future development. Nevertheless, the author tried to stick to the script to the extent possible.

The Trial Structure

The participants were sent an online folder containing information about the Stop Exercise, the trial layout and login information. The participants had a few days to familiarize themselves with the exercise before the first chat. They were not supposed to start practicing the exercise yet, only try to understand the essence of it. The time for the first chat session was settled based on when the participants were free to participate. The time was decided some days in advance, and the participants were reminded about it earlier the same day. Figure 5.1 visualizes the chat environment from the wizards view. This differs from the participants view, as they would only see one chat group on the left side of the screen, as well as no participants on the list below.

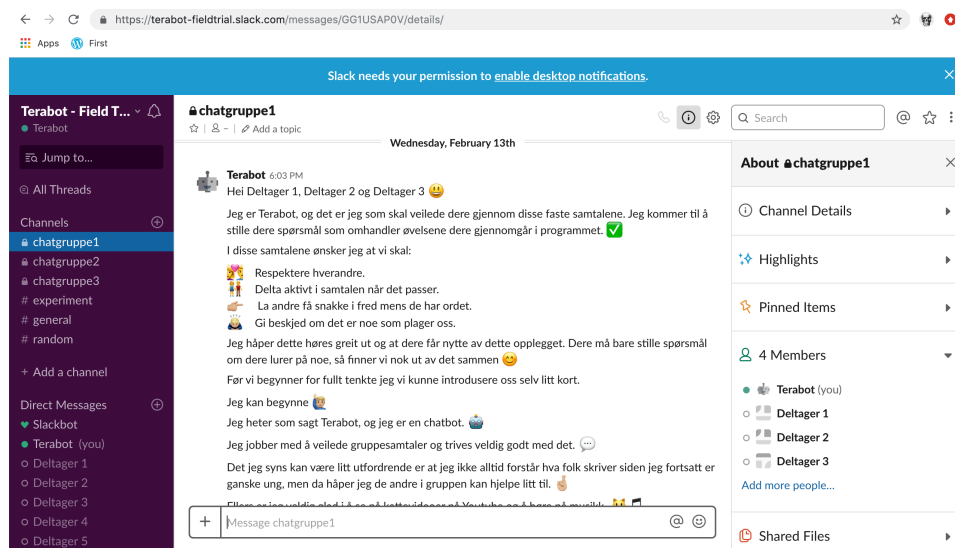


Figure 5.1: A screenshot of the chat environment from the wizards' perspective. On the left one can see the channels for the three different chat groups, and below one can see a list of some of the participants.

All participants received usernames and passwords from the researcher so they did not have to use any private accounts. The participants were anonymous from each other, and were told

that they could choose what information they wanted to share with the other members, but were not recommended to share identifiable information. The reason for them being anonymous was because they were recruited based on convenience sampling and some participants might have knowledge about each other. The author decided that one would not want the participants to be reluctant to answer questions due to other chat members knowing them.

The Chat Sessions

The first chat concerned the understanding of the exercise. The chat started by the chatbot explaining the purpose of the chat, the chatbots' role and limitations, as well as how it wanted the participants to behave in the chat. Then, the chatbot introduced itself and asked the participants to do the same, before the conversation went on as described in Section 4.4. The participants discussed their understanding of the exercise, and were urged to practice it until the next chat session. When the first chat was finished the participants were asked to fill out a questionnaire regarding their experience, which they received by e-mail. The author scheduled the next chat with the participants as soon as the first was completed, and reminded them about it the same day as it was supposed to take place.

The second chat took place approximately one week later. The participants were supposed to have practiced the exercise in the meantime. The second chat considered how the practicing of the exercise had been, and if the exercise felt useful for the participants. The layout was relatively similar to the first chat. When the second chat was finished, the participants filled out the second questionnaire. The researcher then revealed for the participants separately that it wasn't a chatbot that guided the conversations but a researcher and why this was the circumstance. None of the participants expressed a negative perception of this, and some even stated that they felt that the chatbot was too good to be true, since it didn't make any big mistakes or errors.

5.1.1 Results: The Wizard's Notes and Chat Logs

A part of the evaluation was the wizard taking notes if something notable or unexpected happened in the chats, for example user behavior that wasn't anticipated. Additionally, the researcher analyzed the content of the chat logs. The researcher made some notions

considering group compositions, group dynamics, requirements one should consider in the future as well as benefits and limitations with the use of the WoOz method.

Composition of Groups

Group 1 (participants 1-3) and 2 (participant 4-6) had similar group compositions, both consisting of two women and one man in their mid twenties. Group 3 (participant 7-9) had a more diverse group composition, consisting of two women and one man, with ages that ranged from mid twenties to mid sixties.

Group Dynamics

Based on the analysis of the chat logs, the researcher experienced the three groups to have different group dynamics. In Group 1 the conversation had a nice flow, and the participants answered complementary. The tone was interpreted as positive, and it seemed like the participants connected with each other. Sometimes participants referred to each other (username) in their answer, as seen in the example below:

...

Terabot: "Did you experience anything positive with this exercise?"

Participant 1: "I managed to collect my thoughts during a stressful situations, which usually is very difficult for me!"

Participant 3: "I agree with Participant 1 ☺"

...

In Group 2 the conversation appeared more direct, without any unnecessary conversations. The answers were often shorter than in Group 1, but seemed to answer the questions to the same extent. Participants in Group 2 also referred to each other by several occasions when answering questions, especially if they agreed with another participants answers. Neither of the participants in Group 3 referred to each other during the two chat sessions. In Group 3 the wizard had to improvise more, as participants sometimes did not directly answer the question, did not answer at all, answered very vaguely, or asked Terabot a question.

Improvising Answers

There were several occasions where the script did not match the outcome of the participants' answers. In Group 3 the participants answered very vaguely when asked to introduce

themselves and only stated their profession, etc “I’m a student”. The next question in the script was “Did you have anything in common with the other participants?” The participants’ vague answers resulted in the wizard improvising a follow-up question regarding their hobbies. Participants sometimes wrote several answers to acknowledge what another participant had answered, but then the wizard gave them some extra time before moving on to the next question. In some occasions a participant did not answer the question at all. The wizard waited for a while before specifically referring to that participant and asked them to answer. The example below illustrate such scenario:

...

Terabot: “How do you understand the purpose of the exercise?”

Participant 9: *Answers*

Participant 8: *Answers*

30 seconds without activity

Terabot: “What do you think, @Participant7?”

Participant 7: *Answers*

...

Benefits and Limitation with the Wizard of Oz Method

As described in Section 3.5.1, the WoOz method can be valuable for testing a new chatbot potential, evaluate scripts, define requirements and/or learn about what type of user input and behavior one can expect. Based on this WoOz trial the concept of a chatbot guiding a group conversation seemed feasible, and worth exploring more. The scripts, chatbot personality traits and requirements appeared well designed, but with potential for some improvements.

Low-fidelity prototyping methods, such as the WoOz method, may provide little error checking, and thereby important design decisions can be overlooked (Section 3.4). In this WoOz trial there was not made any measures to regulate this limitation. In the first WoOz trial it was considered valuable to be able to improvise, but in the next trial one could make room for demonstrating chatbot errors too. To make a more realistic chatbot experience the wizard could have strictly followed a script that included all possible answers. This includes the wizard having to follow the script if participants did not understand the question, asks a question, or similar. It could also be useful to determine some time constraints the wizard should follow before interfering in the conversation.

5.1.2 Results: Data from the Chats

Based on the quantitative data from the chats, the first factor to address was reliability for participants to be ready when the chats started. All participants were explained that it was important to be logged into the chat and be ready by that scheduled time, as the chatbot wouldn't start before all participants were logged in. Only one group was ready in time for both chats. The two other groups started between 1 to 4 minutes after they were scheduled to. It was usually one participant in each group who wasn't ready in time. This was a factor that needed to be considered in future design, as it can cause problems if not handled correctly. There could be measures to ensure that all participants remember the session, or a way for the chat to continue without them.

There were also some differences in duration of chats, as seen in Table 5.1. Group 2 stand out as they had shorter conversation in both 'Chat 1' and 'Chat 2' than the other groups. These results agree with the researchers' opinion of Group 2 being efficient in their conversation.

Group Number	Chat 1	Chat 2
1 (N=3)	24 minutes	29 minutes
2 (N=3)	17 minutes	20 minutes
3 (N=3)	26 minutes	23 minutes

Table 5.1: The durations of 'Chat 1' and 'Chat 2' measure in minutes for each group.

5.1.3 Results: Online Questionnaires

After each chat session participants answered an online questionnaire consisting of 12 questions ('Chat 1') and 11 questions ('Chat 2'), as explained in Section 3.5.3.

The General Experience of the Group Conversation

Overall the participants expressed a positive experience with both chats. Participants were asked to describe the experience of the group conversation they recently participated in. All answers to this question were interpreted positive. Answers in the first chat included the conversation being nice, the participant having a good experience, the conversation being surprisingly helpful, easy to understand, and positive feelings towards the concept. In the

second chat statements expressed the experience as educational, informative, exciting, nice, positive and affirmative, and generally a good experience.

Level of Connectedness

Participants were asked to what extent they felt connected to the other participants in the chat. The value 1 indicates a very low level of contact, while 7 indicates a very large level of contact. As visualized in Figure 5.2, the level of connectedness increased from 'Chat 1' to 'Chat 2'. In 'Chat 1' the average was 4,78, which increased to 5,89 in 'Chat 2'. There can be several reasons for this: It can be a natural increase due to having talked to each other over two chats, or it could have something to do with the differences in the scripts.

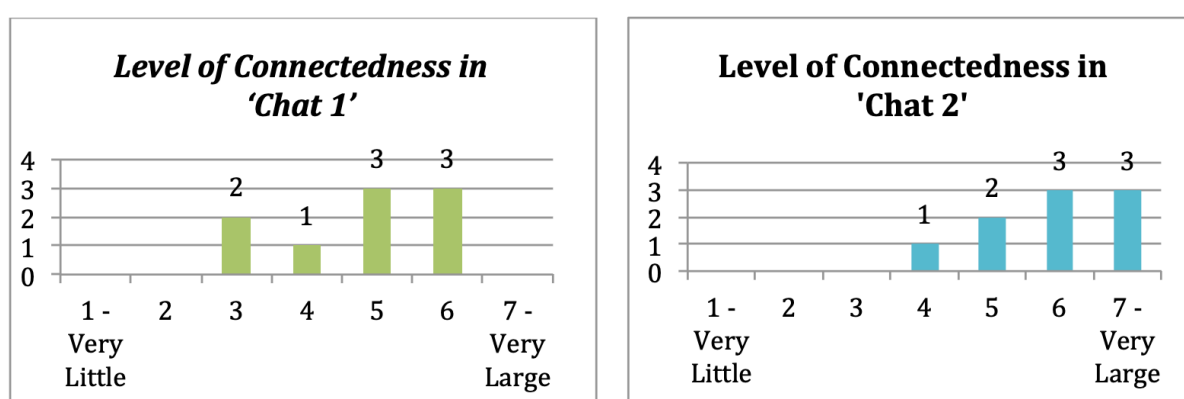


Figure 5.2: A diagram illustrating the level of connectedness in 'Chat 1' and 'Chat 2'

The Value of Discussing the Exercise with Peers

There was an agreement between participants in both 'Chat 1' and 'Chat 2' that it was useful to discuss the exercise with peers. In both chats, the average answer was 6, where 7 was the top score indicting *very useful*. The results show that it was considered valuable to discuss the exercise before starting to practice it, as well as how the practicing went in a later chat.

The Impression of Terabot, the Chatbot

Several questions were related to the impression of Terabot, the chatbot, and its role in the conversations. Regarding the overall impression of Terabot, all participants expressed it was *very positive* (6-7). The average score based on both chats was 6,5 of 7 possible. One question regarded the usefulness of the chatbot in the conversation. In both chats the chatbot was perceived as useful. It was a small decrease from the first to the second chat (6,67 to 6,56).

Participants were asked to assess how disturbing they experienced the chatbot in the conversation. In the first questionnaire one participant assessed 2, which indicate the chatbot to be *very disturbing*. The rest of the participants answered *not disturbing at all* (6-7). There is the possibility that the participant who assessed the chatbot to be *very disturbing* in questionnaire one misread the question, as all participants in the second questionnaire thought the chatbot to be not disturbing at all.

Characteristics of Chats

The participants were further asked to answer some general questions regarding the chats' characteristics. The distribution of who spoke/chatted was perceived as evenly distributed with most participants assessing 7, *very even*, in both questionnaires. A related question regarded if the participants felt that they got to say what they wanted to. As seen in Table 5.2 below, participants in general felt satisfied regarding this.

Statement	Scores	Chat 1 (N=9)	Chat 2 (N=9)	Mean (N=18)
To what extent did you experience contact with the other participants in the chat?	1 = very little, 7 = very large	4,78	5,89	5,335
To what degree did you find it useful to discuss the 'Stop Exercise' with people in the same situation?	1 = very useless, 7 = very useful	6	6	6
What was your impression of Terabot (the chatbot)?	1 = very negative, 7 = very positive	6,44	6,56	6,5
How disturbing did you experience Terabot, the chatbot, in the conversation?	1 = very disturbing, 7 = not disturbing at all	4,89	6,67	5,78
How useful did you find Terabot, the chatbot, for the conversation?	1 = very useless, 7 = very useful	6,67	6,56	6,615
How did you experience the distribution between who spoke/chatted?	1 = very little, 7 = very large	6,56	6,56	6,56
To which degree did you experience to say what you wanted to?	1 = very little, 7 = very large	6,67	5,89	6,28

Table 5.2: Average Scores from the online questionnaire the participants answered after both chats in WoOz 1.

Specifying of what they liked and potential for improvements

In the end of the questionnaires participants were asked to express what they liked with the chat, if anything, and to suggest areas of improvement. Participants expressed many of the same opinions as in the first question when it came to explaining what they liked. Participants further expressed that they felt the questions were easy to understand. Several participants mentioned that Terabot had an important role, and did well in managing the group and pay attention to all participants. One participant wrote, “Everyone got to say what they wanted to, and Terabot asked clear questions and gave explanations along the way, and had a very positive tone. It was also useful to see who wrote at the bottom of the chat”. This statement seemed to summarize several of the opinions.

In consideration to improvements, there were some suggestions, but none were mentioned by several participants. One participant felt it was unnecessary with the introduction round. Another suggestion was to ask specific participant distinct questions instead of asking every question in plenum. One participant noted that he/she liked the use of emojis in ‘Chat 2’ better, as he/she perceived it as a bit much in ‘Chat 1’. The last suggestion for improvement was to review some of the questions (‘Chat 2’), as it could be difficult to answer certain of them. All these suggestions were taken into consideration with reference to the other data and experiences that was gathered through the first Wizard of Oz Trial.

5.1.4 Summary

Nine adults participated in the first Wizard of Oz trial, which included them having to engage in two group chat sessions and practice a cognitive exercise. The participants were asked to assess two online questionnaires regarding the experience. Overall, the participants evaluated the chatbot and the sessions as positive. The researcher registered some user behavior and chatbot features that should be taken into consideration in future design. The aims of the trial were to evaluate the scripts and test the potential for a chatbot to guide an online peer conversation, as well as make some opinions of the WoOz method. The results from the chats and the questionnaires indicated that it is a great potential for chatbots to guide these types of conversation, and that the concept should be explored further.

5.2 Adjustments and Specifications

Based on the first Wizard of Oz trial there were some areas of adjustments that needed to be considered before conducting the second WoOz trial. In the beginning of ‘Chat 1’ the chatbot had stated that participants should “Notify if something bothers you”. This statement was removed, as the chatbot would not be able to answer unanticipated statements or questions, and one could not rely on other participants to answer instead. It could have been developed key words participants could use in these settings, but this was not prioritized in this phase of the project. There were made some small adjustments in how the chatbot formulated some questions and answers and some emojis were removed. These changes were considered to have minor impact on the user experience. Since some questions (e.g. “Have you practiced the exercise since the last chat?”) could cause three different outcomes (all, some or none participants have practiced) it was decided to develop three different chatbot answers based on what participants answered. As stated in Section 5.1.3, it could be useful for the wizard to follow a script fully. It was further developed a script with prompts the chatbot could use in different scenarios. Which scenarios the prompts could be used were specified in the document.

It was settled that the wizard would ask a question, give participants some time to answer, and when the first participant had answered the wizard would wait 10 seconds. If no participants had any activity (writing/answering) in 10 seconds the chatbot would ask the other participants to answer. If there were activity from the participant but then it disappeared, the wizard would start to wait for 10 new seconds. The number 10 was chosen as a base to explore what would be a realistic time constraint. When all participants had answered the question, the chatbot would move on to the next question. Figure 5.3 is a simplified model for the chatbot conversation, including the time constraint.

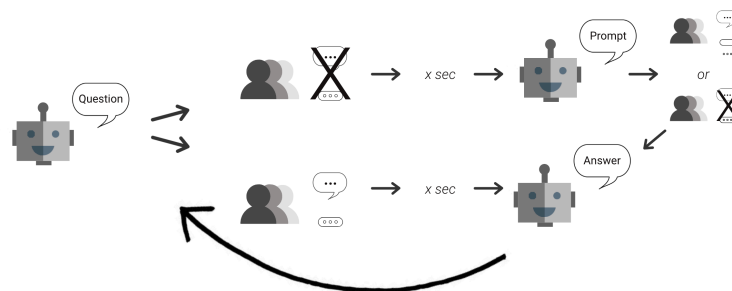


Figure 5.3: An illustration of a possible framework for the chatbot conversation: The chatbot asks a question, oversees if participants write/answer. If participants' answers the chatbot

waits for x seconds (in case participants having a discussion) before asking the next question. If participants do not write/answer in x seconds the chatbot follow up with a prompt and wait a new round of x seconds. If there still isn't any activity from the participants the chatbot moves on to the next question.

5.3 The Second Wizard of Oz Trial: Evaluation with Adults with ADHD

In the second WoOz trial the wizard strictly followed the scripts made, and the only interference the wizard did was based on the separate prompt script. It was specified in which situations the prompts could be used, and it was considered important that the wizard followed this rule. The reason for the strict rules was to make the experience as realistic as possible, including possible errors. By following the script, the prompt document, and the rules made for the wizard, the experience should be more similar to a real chatbot. The choices made for the second WoOz trial agrees with the suggestion of Q. Yang *et al.* (2019) whom proposed to include the possibility for the chatbot to make errors.

The second Wizard of Oz trial was conducted with adults with ADHD as participants. The purpose of this trial was to examine if adults with ADHD would have the same user behavior and experience as the participants in the first Wizard of Oz trial. It was also valuable to examine the benefits and limitations with the *alternative* WoOz trial suggested by Q. Yang *et al.* (2019). The researcher asked the participants if they were comfortable with using their first name in the chats, which all participants answered yes to. The chats were recorded so they later could be evaluated for typical user behavior.

The structure of the alternative WoOz trial was identical to the first WoOz trial:

- Participants got an introduction to the Stop Exercise and some days to familiarize with it
- Participated in the first chat and answered an online questionnaire
- Practiced the exercise for a couple of days
- Participated in the second chat and answered an online questionnaire
- Was asked to participate in a focus group

The Trial

The second WoOz trial was facilitated for two different chat groups consisting of three participants each, but due to different circumstances both groups decreased to two

participants. In the first chat session in both groups one participant did not show up. After waiting for ten minutes while trying to contact the missing participant without luck, the wizard decided to start the chat session with only two participants. It was further decided that participants must have been a part of ‘Chat 1’ to participate in ‘Chat 2’; to make sure participants had the same basis for the conversation. Both chat groups completed the first chat without further difficulties.

‘Chat 2’ was only successful for Group 1. The reason why ‘Chat 2’ was unsuccessful for Group 2 was due to a participant not answering. Both participants in Group 2 were online, but one of them seemed to have forgotten the chat. The wizard tried to contact the participant who was inactive, but without luck. After 10 minutes the wizard cancelled the chat and asked the online participant if she was able to reschedule, but she was not. It was therefore not possible to complete ‘Chat 2’ with Group 2. Participants in Group 1, which finished both chats and questionnaires, were asked to participate in a focus group. Both accepted.

All participants who did not participate in the chats they were scheduled for contacted the wizard later. There were different reasons for why they had not attended the chat: Sickness, oversleeping and getting distracted with something else. These are scenarios that can happen in real life, and therefore important to take into consideration. In the end of this trial, the results ended with two groups of two participants that completed the first chat and questionnaire, and one group of two participants completed the second chat and questionnaire. The participants of the group that completed both chats participated in a focus group.

5.3.1 Results: The Wizard’s Notes and Chat Logs

Composition of Groups

Group 1 consisted of a man and a woman, who according to themselves were both “elderly”. Group 2 consisted of two women with different ages (34 and 66), according to their presentation in the chat.

Group Dynamics

In the analysis of the chat logs, the researcher interpreted the participants in Group 1 to have a good dynamic; as they often referred to each other by using each others name and gave feedback unsolicited. Both seemed talkative and answered complementary. The researcher

interpreted the chats as the participants had a positive experience. The conversation between participants in Group 2 was interpreted as more strained, compared to Group 1. One of the participants seemed to be doing other things simultaneously, as she often logged in and out of the chat, which lead to pauses in the conversation. Group 2 only participated in one conversation, so it was limited data to analyze.

Use of Prompts

The wizard had to use prompts several times throughout the conversations. One of the participants in Group 2 had logged out in the middle of the chat, and did therefore not answer Terabots question. When the online participant had answered, the wizard waited 10 seconds before prompting the offline participant to answer. When the offline participant did not answer, Terabot moved on to its next answer, which was a “Goodbye” statement. The participant who did not answer came online again, and answered the previous question, which resulted in the participants continuing the conversation after the chatbot had said goodbye. One of the participants then asked a question, and the chatbot sent a prompt that it didn’t understand and that if it was important, the participant could ask the person in charge of the program. That ended the conversation. The wizard did not use any prompts with Group 1.

Future Requirements

One dilemma the wizard experienced in several occasions was that when the wizard thought the participants were done discussing (both had answered and the wizard waited for 10 seconds after the last answer) the wizard started to write the upcoming answer/question. At the same time one of the participants started to write as well. This caused a dilemma for the wizard: Would the chatbot have continued to write its response, or would it stop and let the participants finish? In this trial the wizard chose to continue with its response, as this was a dilemma there were no requirements for and this was an alternative WoOz trial, and it could possibly contribute to useful feedback.

Benefits and Limitation with the alternative WoOz Method

The second WoOz trial tried to regulate some limitations of low-fidelity prototyping by having strict rules the wizard had to follow. The intent was that this could create a more realistic chatbot experience for the participants, as the wizard could not improvise when the script was inadequate. These rules generated some chatbot errors throughout the trials, as the

wizard had to follow the script or prompt document. This is considered positive, as one can learn from the possible errors, and as a functioning chatbot would most likely generate errors.

Even though the second WoOz trial allowed for likely errors, there are still limitations to this approach. The wizard followed the scripts and the prompt document, but a chatbot will have even more limitations and potentially cause other types of errors as well. Nevertheless, this early in the design process one cannot consider all possible errors, and the main goal was to explore the possibility of designing a peer support chatbot, which is achieved by the use of WoOz.

5.3.2 Results: Data from the Chats

Both of the first chats started with a delay as it was supposed to be three participants in each chat, but one did not log in. The other two participants in both groups were ready at the scheduled time. The same was halfway true for ‘Chat 2’, but one of the participants who were online and ready a while before the chat was supposed to start, got distracted and started to do something else instead of participating in the chat session. This may signal that if participants are logged into the chat too early they can get distracted while waiting for the chat session to begin. This is a dilemma that needs to be considered in future design.

In ‘Chat 1’ the time used to complete the chat was very similar between the two groups, as illustrated in Table 5.3. Group 1 used 28 minutes on ‘Chat 2’. An interesting finding is that the time used to conduct both chats are comparable to the groups who used the longest time completing the chats in the first trial. In the first WoOz trial the longest time used to complete ‘Chat 1’ was 26 minutes, and ‘Chat 2’ was 29 minutes. These results are especially interesting since it is fewer participants in each group in the second WoOz trial, which is why one could have expected shorter chat sessions. Both WoOz trials were of small size, so differences in the duration of the chats do not signify anything, but it could be an aspect worth exploring further.

Group Number	Chat 1	Chat 2
1 (N=2)	25 minutes	28 minutes
2 (N=2)	26 minutes	-

Table 5.3: The durations of ‘Chat 1’ and ‘Chat 2’ measure in minutes for each group.

5.3.3 Results: Online Questionnaires

Four participants answered the first online questionnaire, and two participants answered the second one, which is represented in the data in Figure 5.9. No participant expressed having experience with chatbots beforehand. The participants described the first chat as: “Pleasant, with a topic that interests me a lot”, “It went well. I didn’t know what I expected, but it felt unexpectedly good”, “Okay”, and “It’s okay, but it’s characterized by different life situations and age”. When asked if they liked anything in particular, participants answered: “Good guiding of the conversation from Terabot. Nice co-participant who followed the theme constructively. Mutual”, “It was a nice scene, we got good guidance from Terabot”, “One got time to reflect”, and “We were different, which allows for reflection”. Only one participant answered on wanting to change something, stating: “Conversations within the group are useful, but this did not Terabot facilitate for”.

After ‘Chat 2’ only one participant answered on the overall experience. The participant described it as “Very good”. When they were asked to describe if they liked anything they stated “Relevant” and “Yes, very pleased. Chats in front and after training provided an extra dimension and encouragement to the exercise, and good reflection afterwards”. After ‘Chat 2’ one participant stated that they felt the chat being a bit short. Another participant noted that the awareness of a robot being present increased slightly, especially when the chatbot asked questions that were already answered. The participant concluded by saying that “But the robot is not so stupid;-)”.

As one can see in Table 5.4 the overall impression of the concept is positive, as the lowest score, how disturbing they experienced the chatbot in the conversation, is 4,75. Both level of contact with the other participants and distribution between who spoke got the highest score, with 6,5 of 7 possible. This reinforces the researchers’ impression about a well functioning group dynamic. Further, it is interesting to see that there was a decrease in the participants overall impression of the chatbot and how disturbing they found the chatbot to be for the conversations. The results are visualized in Table 5.4.

Statement	Scores	Chat 1 (N=4)	Chat 2 (N=2)	Mean (N=6)
To what extent did you experience contact with the other participants in the chat?	1 = very little, 7 = very large	5	7	6
To what degree did you find it useful to discuss the 'Stop Exercise' with people in the same situation?	1 = very useless, 7 = very useful	4,5	6	5,25
What was your impression of Terabot (the chatbot)?	1 = very negative, 7 = very positive	5	4,5	4,75
How disturbing did you experience Terabot, the chatbot, in the conversation?	1 = very disturbing, 7 = not disturbing at all	5,75	4	4,88
How useful did you find Terabot, the chatbot, for the conversation?	1 = very useless, 7 = very useful	5,33	5,5	5,42
How did you experience the distribution between who spoke/chatted?	1 = very little, 7 = very large	5,75	6,5	6,13
To which degree did you experience to say what you wanted to?	1 = very little, 7 = very large	6	6	6

Table 5.4: Average scores from the online questionnaire the participants answered after both chats in the alternative WoOz trial.

5.3.4 Results: Focus Group

Overall, the participants in the focus group appeared satisfied and impressed by the concept and of Terabot. When asked what they thought about the concept and the experience one participant answered:

“There are two aspects I’m thinking about: One is that it was very useful to be forced to formulate in advance what one is supposed to do. That helps me a lot. [...] The questions that were raised were very relevant. Thinking in advance “when am I supposed to practice it [the exercise]?” was very relevant. I also thought it was really nice to have peer support. I often experience that people without ADHD can sometimes make me feel like they are there to educate. I can often experience that other people are “proper” and that they want to educate us. That’s a bit condescending, and you feel like you have to compensate in a way for how you behave. Having

conversations with others who are in the same situation is very rewarding, and it gives me very much. I can breathe out differently, and I feel freer and I can get a lot in return, I think. So I think it was very nice with those two aspects: That one got a preparation and that one also had someone to talk to” (Authors translation).

The statement above represented much of what was expressed in the focus group. The participants mentioned several times the value of preparing themselves for the exercise, as well as the importance of having someone to relate to. Both participants agreed that the peer support chats (guided by the simulated chatbot following the scripts) lead to a feeling of positive pressure towards practicing the exercise and participating in the chat sessions. Participants further described Terabot as “nice”, “helpful”, and that it “had an inviting tone”. The questions Terabot asked were interpreted as relevant. Participants also noted that they liked when Terabot praised them when they had e.g. remembered to practice the exercise. One participant stated that:

“Sometimes it [the chatbot] gave some praise, such as “Nice that everyone has dealt with it”, or something similar. [...] Yes, that [the praise] felt nice” (Authors translation).

Considering the preferable size of the chat groups, the participants seemed a bit unsure, as there were benefits with both smaller and larger groups. After some discussion they agreed that three to four participant would be suitable. They stressed that they had very intense chat conversations, which they both preferred, as one could not do other stuff at the same time. The composition of participants could be based on similarities in age so it would be easier to connect with each other.

Having a set time for when the chat should start could be problematic, but so could a more indefinite time period, according to the participants. Having a very strict time for the chat could cause a lot of stress. One participant stated that:

“I use a lot of energy on being on time, so maybe it could be possible to have a casual conversation for a while [in the chat], and then start the actual chat a bit later”.

The other participant agreed with this idea. The participants expressed a solution where the chat session would start at e.g. 18:00, but the ten first minutes of the chats would consider more casual topics. When the time was 18:10 the “real” chat regarding the program and the exercise would start. This idea facilitate for the possibility of participants being a bit late, at the same time as participant who are ready on time do not have to wait for the others. One participant suggested that this idea could also improve the following discussion about the exercise, as participants could have better familiarity to the other participants.

Several times in the chat sessions participants experienced that they had more to say on the current topic when Terabot moved on to the next one. The participants recognized that it was difficult to determine how long one should stay on one topic and how the chatbot could know when the conversation had changed to an unrelated theme. The participants agreed that it was useful and important for Terabot to set some constraints in the conversations so the conversation would be relevant towards the program. Nevertheless, the participants had some suggestions for how Terabot could recognize valuable conversation between participants. If Terabot was moving on to the next question (starting to write the question) and a participant started to write, it was considered positive if Terabot stopped writing and let the participants discuss some more. Terabot could also ask participants if they wanted more time to discuss, if they seemed to have a lot to say (several messages and activity). One participant suggested that Terabot should pay attention to if participants used each other’s names:

“Maybe it is an idea to give some extra time [for discussion] if a participant uses another participant’s name. It can be a signal of them having a conversation, which one should possible give some space for” (Authors translation).

A solution for expressing more peer support was articulated. Often the participants would recognize what the other participant had stated. It could be time consuming to answer everything one recognized, and it could be difficult to do that at the same time as one followed Terabot’s conversation. One possible solution that was expressed would be to integrate feedback symbols in the chat, similar to ‘likes’ on Facebook. Then participants could give a symbol feedback on messages where they recognized the problem, or similar. These symbols should be of positive context, and promote peer support.

Overall both participants were positive to the concept and what was done to this point in the project. One participant stressed this:

“I think it is very positive that someone has tried to do this [develop peer support technology for the online self-help program for adults with ADHD]. I think you have succeeded in finding a good format [chatbot] for it. It can easily happen that someone talks too much, or interrupts, or yes... I was very impressed on how well it functioned and the benefits of it. [...] Sometimes it feels like it is developed technology in relation to what it [technology] can do, not what is needed or what is effective. This [the peer support chatbot] felt like it was developed based on “How can we best facilitate for a good situation for those who will have this conversation?” [...]”.

(Authors translation)

5.3.5 Summary

In the second Wizard of Oz trial four participants completed ‘Chat 1’ and its questionnaire, and two participants completed ‘Chat 2’, its questionnaire, and participated in a focus group. Based on data from the chat logs and the results from the questionnaires and focus group, the concept of a peer support chatbot guiding a group conversation between adults with ADHD seems feasible. The results from the questionnaires were positive, and both participants in the focus group were satisfied and impressed with the concept and Terabot. The participants suggested some possible design solutions, such as group compositions, how the chatbot could foster more peer support between participants, how one could customize the time frame of the chat considering participants being late.

5.4 Chapter Summary

This chapter has presented the first Wizard of Oz trial and its results, adjustments made to the scripts, and the second Wizard of Oz trial and its results, including a focus group. The concept of a peer support chatbot guiding a group conversation between adults with ADHD seems feasible, and the participants in the focus group contributed with suggestions for how to make the concept even better.

Chapter 6

Discussion

The overall research question of this thesis concerned how to design a guided chat-based peer support technology accompanying an online self-help program for adults with ADHD. Based on findings from the literature review presented in Chapter 2 and design workshops presented in Chapter 4, the research aim got specified to include conversational interfaces. From this, two new research questions arose concerning 1) how one could design for a conversational interface which role is to guide a group conversation between adults with ADHD, and 2) what opportunities and limitations the Wizard of Oz method offers when used in the design of conversational interfaces. To answer the research questions the scripted chatbot, Terabot, was designed as described in Chapter 4. Terabot was evaluated through two trials of Wizard of Oz. This was all accomplished by using methods described in Chapter 3.

This chapter will discuss the process and results, compare the design of Terabot to the design considerations mentioned in Section 2.5, answer this thesis's research questions, and explain limitations with this study.

6.1 The Potential of Using a Chatbot for Peer Support

The concept of having a chatbot guiding a peer support conversation was perceived positive both within the INTROMAT team and amongst the participants who engaged in the Wizard of Oz trials. The results from the WoOz trials were satisfactory and indicated that guided group chats based on predefined scripts have potential in relation to online self-help programs. Due to this prototype being low-fidelity and limitations with the WoOz method a real chatbot must be developed and evaluated to validate the concept of a chatbot guiding the peer support conversation. Even though the second WoOz trial facilitated for likely errors, it is possible that a real chatbot would have caused other types of errors and problems not predicted in this thesis. The only way of validating the potential for a real chatbot guiding a peer support conversation is by further developments and tests.

Further, this thesis is a proof of concept that the scripts and format of Terabot was perceived as useful and satisfactory when discussing a cognitive exercise, both amongst adults with and without ADHD. This might indicate that the format of Terabot is not limited to an online self-help program for adults with ADHD, but could also be used in similar interventions with other purposes and participants. Based on the results from the online questionnaires and the focus group, the questions and layout of the chats were experienced as relevant and beneficial for the exercise. Many of the choices and requirements used in the design of Terabot might be of inspiration to future researchers who wants to explore the potential for chatbots in group contexts or peer support conversations.

High-fidelity chatbot prototypes are costly and time consuming to develop, and may therefore not be suitable for the early stages of a project. Low-fidelity chatbot prototypes, such as Terabot, are less time-consuming and low-cost. For these reasons low-fidelity chatbot prototypes can be very useful in the beginning of a project as it is a productive way of gathering requirements and analysis, and may function as a proof of concept.

6.1.1 Design Choices – The Emergence of Terabot

The design of Terabot happened through several iterations of design and evaluation. In the beginning of the project it was expressed a desire for peer support technology in relation to the online self-help program INTROMAT is developing. Based on this desire the idea of a peer support chat appeared, before the idea further advanced to the concept of chatbot guiding the peer support conversation. As mentioned in Section 2.3, online peer support has great potential, but also accompanying risks. Since the peer support technology would be used in relation to mental health, it was considered especially important to deal with these risks. Based on a literature review and a design workshop it was speculated that a chatbot might have the potential to deal with these risks, as well as benefit the peer support conversation and be less expensive than a human operator. From that time and on, this thesis considered how a chatbot could guide peer support conversations between adults with ADHD who participated in an online self-help program.

Throughout the development of Terabot, several design choices were made, as described in Chapter 4 and 5. Several requirements were supplemented throughout the iterations as new knowledge appeared, and adjustments in already determined requirements were conducted.

As the design choices have already been explained in previous chapters, only some specific choices and their constraints and possibilities will be discussed in this section.

Scripted Chatbot

When designing and developing requirements for chatbots with guiding purposes one must carefully consider and evaluate, especially if the purpose is for mental health intentions. This thesis agrees with the theory presented in Section 2.5.3 stating that scripted chatbots are appropriate for guiding, as there are limited alternatives for the participants. Chatbots that are designed to guide will often have a chatbot driven dialogue, indicating that the chatbot is in control of the conversation. This agrees with the use of scripted chatbots, as the designers can develop the whole layout and script, and thereby have better control of the possible outcomes. Since the conceptual chatbot in this thesis would be scripted it was important to design and plan its purpose, what questions to ask, personality traits and evaluate the different requirements. This was done in design workshops with help of expertise from domain experts. The domain experts provided useful information and suggestions on how to guide, which questions to ask, and other important factors to consider.

Guiding

According to Baumeister et al. (2014) online interventions should always be guided if possible. Section 2.3.4 presents literature regarding guiding in online peer settings. The study of O’Leary et al. (2018) show that participants who engaged in peer support chats that were guided perceived that conversation as “...deeply valuable for gaining solutions and insights...” (O’Leary *et al.*, 2018, p. 10). The unguided peer chats from the same study were perceived as smooth and easy-going, but was considered more of a distraction from problems than a solution. Evaluations of Woebot (Fitzpatrick, Darcy and Vierhile, 2017) show that participants who receive therapy from a chatbot based on Cognitive Behavioral Therapy, experience decrease in depressive symptoms. Terabot was designed to mainly consider an exercise from the self-help program, and it therefore seems preferable for the peer conversation being a way of gaining insights and solutions. The guiding of Terabot was based on expertise from the domain experts in the INTROMAT team. Several of the domain experts have experience and knowledge from Internet interventions and Goal Management Training interventions.

Scheduled Chats

One requirement that set some constraints for the design alternatives was the need for the chat to be monitored, as expressed by the INTROMAT team. This requirement entailed the chats to have scheduled sessions. The reason for the need of human monitoring was due to the possibility of negative behavior in the conversations (e.g. harassment or self-destructive behavior) and that the project has a responsibility for both the products they offer and the participants. The requirement of having scheduled sessions caused design restrictions, as other possibilities were not considered. Based on both WoOz trials, it might seem like having scheduled times for the chats could cause some difficulties. In the first WoOz trial two out of three groups experienced participants being late, which resulted in the other participants having to wait. The waiting varied between one and four minutes. In the second WoOz trial both groups experienced a participant not showing up at all in ‘Chat 1’. This led to participants in both groups having to wait for 10 minutes before the chat started with only two participants.

The participants in the focus group expressed a possible solution that would be valuable to consider in future design. Both participants appreciated that the chats were held at scheduled times so they knew when they should participate and that it would be an efficient conversation. They suggested that instead of the chat beginning to discuss the exercise, it could be some casual small talk first guided by the chatbot. For 10 minutes the chatbot could ask participants questions about their everyday life and other topics not specifically related to the self-help program, but with some other value (e.g. emotional support). After 10 minutes the conversation would change to the exercise and program. This way participants who were ready on time would not have to wait for the others, as well as participants who logged in too late could still participate.

Peer Support Features

The reason for having a peer support conversation in relation to the online self-help program is due to the usefulness of speaking to others who understands you, which has several benefits expressed in Section 2.3. Based on results from the online questionnaires from both WoOz trials the level of experienced contact with other participants was acceptable but could still be improved. One possibility is that it would naturally improve as the participants interacted with each other over time, which the results presented in Section 5.1.3 and 5.3.4 imply. However, other design choices could possibly also contribute to an increased degree of contact. In ‘Chat

2' it was focus in the script on achieving better peer connections by asking participants to express some positive remarks to the other participants. In future work the scripts could be adjusted to include more conversation where participants refer to each other or are asked to give each other feedback.

Participants in the focus group had suggestions to help facilitate for peer contact. If participants used each other's names in the conversation, the chatbot could take that as an indication that the participants had a valuable conversation and therefore give them some extra time to chat before asking the next question. Another solution could be to add the possibility of expressing symbols on other participants' answers, similar to a Facebook 'like'. Participants could express a symbol on answers they recognized themselves in, or similar. This could possible result in more interaction in the conversation and increase the level of peer connection. This could be beneficial, as it would not interrupt the chatbots conversation, and may give participants an increased feeling of being understood.

Group Composition

In the WoOz trials participants were randomly assembled into chat groups. Levine et al. (2011) had groups based on similarities in their Goal Management Training intervention program. Participants were selected to groups based on age and difficulties. The participants in the focus group agreed that it would be preferable if the other group members were similar to them, at least in age. The participants expressed that it would be easier to relate to the other participants then, which is an important part of peer support conversations.

6.1.2 The Design of Terabot vs. Design Considerations

As described in Section 2.5, there are several design considerations one could take into account when developing a new product. In this thesis these suggestions regarded designing for ADHD, mental health and chatbots. The design considerations from Section 2.5 will be used to compare with the design choices made in this thesis.

Designing for ADHD

When designing for ADHD it was described as important to keep structure. Terabot had a chatbot-driven conversation style, which indicates that the chatbot decided the structure of the conversation and the topics. The conversations were scripted, and domain experts with

experience from group therapy assisted in the development. The chat sessions were held at scheduled times with consistent groups, and there were two chats per exercise/module. These circumstances promote structure.

Praise and re-direction was explained to be important when designing technologies with regards to ADHD. Terabot had embedded praise and re-direction in its scripts. For example, when participants were asked if they had practiced the exercise, Terabot would use their answers as input to its own answer: If all participants had practiced Terabot would praise all of them. Had none of them practiced Terabot would recommend them to do it after that chat session. Had some of them practiced and others not, Terabot would give a divided answer as to who gave the different answers. The chatbot also re-directed participants who didn't answer at all by referring to them specifically for an answer.

The last consideration related to designing technology for ADHD was to minimize distractions. Participants were supposed to actively participate in the conversation when the chat was scheduled. One participant from the focus group stated that she liked the intensity of the chat, since she didn't have time to do other stuff simultaneously. However, it seemed like one participant in the second WoOz trial did other things while participating in the chat, as she logged in and out of the chat throughout the whole conversation. There is also the possibility that she had a bad connection. Nevertheless, Terabot was designed to have an effective conversation style where participants were expected to actively participate throughout the session, but some participants might get distracted anyhow.

Designing for Mental Health

When designing for mental health technologies it is important to know whom you are designing for. In this thesis that was achieved by having participatory design, doing a literature review, and reading interviews with adults with ADHD. Ethical consideration was taken into account, and was in fact one of the reasons why the interface of the peer support technology turned out to be a chatbot. It was recommended to use technology familiar to the intended users, but even though only few participants in the Wizard of Oz (WoOz) trials had experience with chatbots, the results were positive.

In Section 2.5.2 it was described that technology for mental health should build on requirements and traditions of mental healthcare setting and accepted theories and models.

Terabot builds on the concept of peer support and Goal Management training, and mental health domain experts were part of the design team. Section 2.5.2 further described the need to evaluate the technology with both potential users and therapists to make sure the product is suitable for all instances of its intended role (Doherty, Coyle and Matthews, 2010). Terabot was discussed through several design workshops with the domain experts, and was evaluated in a WoOz trial with potential users. One design consideration that is not yet fulfilled, is the evaluation in connection with clinical practice to ensure that it will function in everyday practice (Doherty, Coyle and Matthews, 2010).

Designing Chatbots

Important considerations when designing chatbots will be the capabilities and limitations. Terabot was designed to guide a peer support conversation between adults with ADHD concerning an exercise/module from an online self-help program. In the beginning of the conversation Terabot state that it has limitations and do not understand unanticipated questions. If participants were to ask questions Terabot could not answer Terabot has replies in it's "Prompt" list stating that it does not understand the question, but if it is important the participant should contact the administrators of the program. During the Wizard of Oz (WoOz) trial this did not become a problem. The participants in the focus group expressed that they were aware of the limitations and claimed that people had to behave like adults, and not try to trick the chatbot.

In Section 2.5.3 it was stated that a reason why guide/coach chatbots have little alternative paths is due to the need for limitations and a clear framework for chatbots whose purpose concerns mental health and other sensitive topics. As previously explained Terabot fulfill this. Research presented in Section 2.5.3 express the importance of chatbot personality. When designing the scripts, focus was on demonstrating certain personality traits in Terabot, such as trustworthiness, niceness and it being non-judgmental. During both WoOz trials only positive words were used to describe Terabot, and Terabot was valued as important for the conversation and as likable in the questionnaires.

Three related concepts that also are important to consider are: The uncanny valley, the uncanny cliff, and the mission creep. The uncanny valley was carefully considered when designing Terabot, hence Terabot was designed to display clearly that it was a chatbot, by use of profile picture, things it said in it scripts and its name. Uncanny cliffs occur when chatbots

have an unexpected and sharp boundary between what they know and what they don't know. During the WoOz trials there were some instances where Terabot was asked reasonable questions it couldn't answer. One cannot prepare a chatbot for all related topics and questions, especially early in the design process. In future work these questions should be considered. When designing chatbots it can be difficult to decide which capabilities one want for the chatbot. If the designers always increase the capabilities of the chatbot, it can end up being "a resource black hole", also known as mission creep. It will therefore be important to carefully discuss and decide what type of information a chatbot will have knowledge of in the future.

6.2 Research Questions

This thesis had three research questions that will be discussed in this section.

6.2.1 Designing Guided Chat-based Peer Support

The primary research question of this thesis concerned:

RQ1: How can we design guided chat-based peer support technology accompanying an online self-help program for adults with ADHD?

An important factor to consider when designing peer support technology for adults with ADHD is to know the potential users. There are several ways of doing this. By having a participatory design approach potential users can have a strong impact on the product, and contribute with valuable information the designer would not have known otherwise. Chapter 2 and many of the design alternatives presented in Chapter 4 demonstrate different solutions for designing a guided chat-based peer support technology that could be accompanying an online self-help program for adults with ADHD. Some possibilities are, but not limited to, group chats, forums and one-to-one chats. The level of guiding could vary, but as stated in the Introduction, guiding should always be included if possible. The purpose of having peer support technology in connection with an online self-help program for adults with ADHD could also vary, but the two considered in this thesis was either emotional or exercise based support. There are endless possibilities for how each of these could be designed.

6.2.2 Designing Guiding Conversational Interfaces

RQ2: How can one design a conversational interface which role is to guide a group conversation between adults with ADHD?

As stated by Følstad and Brandtzaeg (2017), HCI researchers and interaction designer must consider conversation as objects of design, due to the change in focus from graphical user interfaces to conversational user interfaces. This makes the designing of scripts an important part of the development of chatbots. Designing of scripts do not only include defining questions and answers for the chatbot, it should also consider how desired personality traits could be visible through text. Further the designer must predict user behavior and input, which can be a complex task. Knowing the purpose of the chatbot and the intended user group is essential in the design process, as the chatbot should be customized according to these factors.

This thesis focuses on the design of a chatbot guiding peer support conversation between adults with ADHD who participated in an online self-help program. The designer ought to be familiar with the user group and the purpose of the self-help program. Both potential users and domain experts participated in design workshops regarding peer support technology for the self-help program. The scripts that were designed were reviewed by domain experts with expertise in Internet intervention and Goal Management Training, before the scripts were further evaluated in Wizard of Oz trials. Results from the reviews and trials were used as input for adjustments and specification of design choices. The concept of the chatbot was evaluated, refined, and evaluated again before the final prototype was completed.

6.2.3 The Wizard of Oz Method Used in the Design of Conversational Interfaces

RQ3: What are the opportunities and limitations of using the Wizard of Oz method when designing conversational interfaces?

Terabot can be viewed as low-fidelity chatbot prototype. There are, as stated in the introduction, some challenges with evaluating chatbot prototypes. Interaction designers and HCI researchers have previously been focused on designing and evaluating graphical user interfaces, but must now also address conversational interfaces and conversation prototypes.

The Wizard of Oz (WoOz) is a method used for testing chatbot prototypes, and can serve different purposes including getting relevant information about the user experiences, the scripts made, chatbot personality features, some technical conditions, and user behavior and input. In this thesis two WoOz trials were conducted as a way of evaluating Terabot, and to get input for further design choices. The results from the trials lead to valuable information and improvement of the prototype.

Nevertheless, WoOz trials are a low-fidelity prototype method and offers limitations. The results from both WoOz trials were good, but it is a test of a simulated bot, and we cannot assume that the results would transfer to a fully implemented and automated chatbot. There are many factors, both technical and user behaviors, that are difficult for designers and researcher to anticipate based only on WoOz trials. WoOz trials might be most valuable in the beginning of a project to test if the concept has potential, to gather information about specific design choices, and to analyze user behavior. However, it should not be used as a validation for such chatbot technology, but rather as an encouragement that the concept has potential to be further tested in a real chatbot. WoOz also offers an efficient and low-cost way of evaluating scripts, chatbot personality traits, requirements, and to gathering information about user behavior and input. If designers do not conduct a WoOz trial or similar evaluations early in the design process, there will probably be a higher chance of the designer missing important details, which again can lead to a less effective end results.

The alternative WoOz trial suggested by Q. Yang et al. (Q. Yang *et al.*, 2019) recommend designers to include the possibility for likely errors. In this thesis the second WoOz trial facilitated for likely errors. An interesting result is that the second WoOz trial in this thesis got lower average scores on the user experience compared to the first trial. The sample of participants in these trials were limited, so the results could have been a coincidence, but it can also be an indication that the alternative WoOz approach may result in lower user experience scores as it present more likely outcomes from the wizard. Anyhow, the author of this trial appraises both variations of the WoOz trial as valuable, but for different purposes. The ordinary trial was perceived as valuable in the very beginning as a proof of concept and to establish some expectations about user behavior and evaluation of scripts. The alternative WoOz trial was perceived as more realistic, making the proof of concept in this trial seem more genuine and valuable.

6.3 Research Limitations

As described, this thesis has not designed and evaluated a functioning chatbot, so the concept of a chatbot guiding a peer support conversation between adults with ADHD is not validated, but this thesis encourage to explore the concept more based on results form two WoOz trials. Additionally, there were a limited number of participants in the WoOz trials, especially adults with ADHD. Further, this thesis aimed to follow a participatory design approach. In this respect, the number of adults with ADHD who participated in the design workshops was limited and could preferably be larger.

Another limitation concerns the trial layout. In a real case scenario the participants would be candidates in an online self-help program for adults with ADHD, and the peer support chatbot would only be a minor part of that program. Based on the presented limitations for this project, the results from this thesis are very specific and can only be emphasized in light of this specific project and case. Still, the gathered data and trials have given valuable insights into the process and are considered sufficient to provide answers to the research questions.

6.4 Chapter Summary

This chapter has discussed the design and evaluation of the low-fidelity prototype, Terabot. Several design choices were discussed more thoroughly, and the final prototype was compared to related design considerations. This chapter ended with answering the research questions of this thesis.

Chapter 7

Conclusion

The research presented in this thesis was done as a part of the INTROMAT project, which is a project aiming at improving mental health by the use of innovative technology. One of the cases in INTROMAT concerns the development of an online self-help program for adults with ADHD. By using the research through design methodology and participatory design approach, this thesis has explored the potential of peer support in connection with the online self-help program. The design and evaluation of a low-fidelity chatbot prototype was part of this exploration. In the evaluation the Wizard of Oz (WoOz) method was used, which lead to the author making some experiences regarding opportunities and limitations.

Through three design iterations consisting of literature review and design workshops the first prototype of Terabot was created. Terabot is a low-fidelity chatbot prototype consisting of a concept, two scripts, personality traits and requirements. Terabot can thus be considered what Følstad and Brandtzaeg (2017) refer to as a conversation prototype. The aim of Terabot was to guide participants in the self-help program through regular chat sessions that concerned an exercise from the program, more specifically *The Stop Exercise*. The Stop Exercise would be accompanied by two chat sessions: One that would focus on building a foundation and expectations towards the exercise, while the other would concern the execution and experience with the exercise. Having two chats for one exercise was considered valuable for both the feeling of peer support and for motivating participants to practice the exercise.

Two WoOz trials were conducted based on the low-fidelity chatbot prototype. In the first WoOz trial the wizard pretended to be Terabot by following the script, but was allowed to do some improvising if needed. The second WoOz trial was inspired by Q. Yang's (2019) suggestion to facilitate such prototypes for likely errors. This suggestion resulted in the wizard having to strictly follow the scripts, a prompt document, and certain time constraints. The prototype received satisfactory results in both WoOz trials, but the scores were slightly better in the first trial. As this thesis had a small number of participants one cannot conclude why there is a difference, and there are several possible reasons: It could be the likelihood of errors in the alternative WoOz trial, it could be small changes in the scripts, differences

between the participants, or completely random. Nevertheless, in the future it would be interesting to explore differences in results by using the ordinary vs. the alternative WoOz trials, and compare these results to an evaluation of a fully developed chatbot.

It is important to emphasize that the findings in this thesis are related to a low-fidelity chatbot prototype, and can therefore not be stated to apply for a functioning chatbot. Due to this prototype being low-fidelity and limitations with the WoOz method a real chatbot must be developed and evaluated to validate the concept of a chatbot guiding the peer support conversation. Nevertheless, this thesis shows that the concept of a chatbot guiding a peer support conversation between adults with ADHD who participate in an online self-help program is perceived as positive from both domain experts and potential users. This thesis encourages development of a functioning chatbot based on the prototype and findings from this thesis.

7.1 Future Work

There are many concepts presented in this thesis that would be interesting to explore further. In relation to the WoOz method it would be valuable to explore the differences between the ordinary and alternative method to a wider extent. In regards to Terabot, it could be useful to explore the design suggestions from the focus group (Section 5.3.5). Further, it could be beneficial to evaluate the low-fidelity prototype of Terabot with more participants with ADHD, as the selection in these WoOz trials was limited.

Additionally, the prototype from this thesis only examined conversations concerning one exercise, *The Stop Exercise*. The results demonstrated that the participants thought it was useful to discuss the exercise based on the guiding. How the participants would react to discuss other types of exercises or several exercises, is not known. It could therefore be beneficial to explore results from an upscaled trial concerning several exercises and chat sessions. Last but not least, it would be useful to develop a high-fidelity prototype based on the design of Terabot, to examine the real potential of a chatbot guiding a peer support conversation between adults with ADHD who participate in an online self-help program.

Bibliography

- ADHD Norge (2016a) *Hva er ADHD?*, ADHD Norge. Available at: <http://adhdnorge.no/voksen/hva-er-adhd/> (Accessed: 20 August 2018).
- ADHD Norge (2016b) *Utredning*, ADHD Norge. Available at: <http://adhdnorge.no/voksen/utredning-2/> (Accessed: 20 August 2018).
- ADHD Norge (2016c) *Voksenliv*. Available at: <http://adhdnorge.no/avdeling/voksen/> (Accessed: 20 August 2018).
- Adobe (2019) *Adobe XD: Design already ahead of its time*. Available at: <https://www.adobe.com/no/products/xd.html> (Accessed: 24 April 2019).
- Alcoholic Anonymous World Service Inc. (2018) *Alcoholics Anonymous*. Available at: <https://www.aa.org> (Accessed: 1 September 2018).
- van Allen, P. (2018) 'Prototyping ways of prototyping AI', *Interactions*, 25(6), pp. 46–51. doi: 10.1145/3274566.
- Alvarez-Jimenez, M. *et al.* (2014) 'Online, social media and mobile technologies for psychosis treatment: A systematic review on novel user-led interventions', *Schizophrenia Research*. Elsevier, 156(1), pp. 96–106. doi: 10.1016/J.SCHRES.2014.03.021.
- Atkinson, M. and Hollis, C. (2010) 'NICE guideline: Attention deficit hyperactivity disorder', *Archives of Disease in Childhood: Education and Practice Edition*, 95(1), pp. 24–27. doi: 10.1136/adc.2009.175943.
- Bargh, J., McKenna, K. and Fitzsimons, G. (2002) 'Can You See the Real Me? Activation and Expression of the "True Self" on the Internet John', *Journal of Social Issues*, 58(1), pp. 33–48. doi: 10.1111/1540-4560.00247.
- Baumeister, H. *et al.* (2014) 'The impact of guidance on Internet-based mental health interventions - A systematic review', *Internet Interventions*. Elsevier B.V., 1(4), pp. 205–215. doi: 10.1016/j.invent.2014.08.003.
- Blanton, M. *et al.* (2009) 'Human-Computer Interaction', in *Encyclopedia of Database Systems*. Boston, MA: Springer US, pp. 1327–1331. doi: 10.1007/978-0-387-39940-9_192.
- Bracke, P., Christiaens, W. and Verhaeghe, M. (2008) 'Self-Esteem, Self-Efficacy, and the Balance of Peer Support Among Persons With Chronic Mental Health Problems', *Journal of Applied Social Psychology*, 38(2), pp. 436–459.
- Brandtzaeg, P. B. and Følstad, A. (2018) 'Chatbots: Changing User Needs and Motivations', *Interact*, XXV.5, pp. 38–43.

- Bryman, A. (2016) *Social Research Methods*. 5th edn. Oxford: Oxford University Press.
- Chung, J. E. (2013) 'Social interaction in online support groups: Preference for online social interaction over offline social interaction', *Computers in Human Behavior*. Pergamon, 29(4), pp. 1408–1414. doi: 10.1016/J.CHB.2013.01.019.
- Ciechanowski, L. *et al.* (2018) 'In the shades of the uncanny valley: An experimental study of human-chatbot interaction', *Future Generation Computer Systems*. Elsevier B.V., 92, pp. 539–548. doi: 10.1016/j.future.2018.01.055.
- Dahlbäck, N., Jönsson, A. and Ahrenberg, L. (1993) 'Wizard of Oz studies - why and how', *Joint Agent Workshops and Symposium*, 6(4), pp. 258–266.
- Davidson, L. *et al.* (1999) 'Peer Support Among Individuals With Severe Mental Illness : A Review of the Evidence', pp. 165–187.
- Davidson, L. *et al.* (2006) 'Peer Support Among Adults With Serious Mental Illness : A Report From the Field', 32(3), pp. 443–450. doi: 10.1093/schbul/sbj043.
- Dickerson, F. B. *et al.* (2001) 'Outpatients with schizophrenia and bipolar I disorder: Do they differ in their cognitive and social functioning?', *Psychiatry Research*, 102(1), pp. 21–27. doi: 10.1016/S0165-1781(01)00247-5.
- Doherty, G., Coyle, D. and Matthews, M. (2010) 'Design and evaluation guidelines for mental health technologies', *Interacting with Computers*. Elsevier B.V., 22(4), pp. 243–252. doi: 10.1016/j.intcom.2010.02.006.
- English Oxford Dictionary (2019) *Stigma*. Available at: <https://en.oxforddictionaries.com/definition/stigma> (Accessed: 15 May 2019).
- Entwistle, V. A. *et al.* (2011) 'How information about other people's personal experiences can help with healthcare decision-making: A qualitative study', *Patient Education and Counseling*. Elsevier, 85(3), pp. e291–e298. doi: 10.1016/J.PEC.2011.05.014.
- Faulkner, A. and Basset, T. (2012) 'A helping hand: Taking peer support into the 21st century', *Mental Health and Social Inclusion*, 16(1), pp. 41–47. doi: 10.1108/20428301211205892.
- Fitzpatrick, K. K., Darcy, A. and Vierhile, M. (2017) 'Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial', *JMIR Mental Health*, 4(2), p. e19. doi: 10.2196/mental.7785.
- Følstad, A. *et al.* (2018) 'Chatbots for Social Good', *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–4. doi: 10.1145/3170427.3185372.

- Følstad, A. and Brandtzaeg, P. B. (2017) 'Chatbots and the new world of HCI', *Interactions*, 24(4), pp. 38–42. doi: 10.1145/3085558.
- Følstad, A., Skjuve, M. and Brandtzaeg, P. B. (2018) 'Different Chatbots for Different Purposes : Towards a Typology of Chatbots to Understand Interaction Design'.
- Garret, J. J. (2011a) *The Element of User Experience*. 2nd edn. Berkley, CA: New Riders.
- Garret, J. J. (2011b) *The Elements of User Experience: User-Centered Design For the Web and Beyond*. Second. Berkley, CA: New Riders.
- Grudin, J. and Jacques, R. (2019) 'Chatbots, Humbots and the Quest for Artificial General Intelligence', *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, p. 11. doi: <https://doi.org/10.1145/3290605.3300439>.
- Hannås, B. M. (2019) *ADHD, Store Medisinske Leksikon*. Available at: <https://sml.snl.no/ADHD> (Accessed: 6 January 2019).
- Harvey, K. J. et al. (2007) "'Am I normal?" Teenagers, sexual health and the internet', *Social Science & Medicine*. Pergamon, 65(4), pp. 771–781. doi: 10.1016/J.SOCSCIMED.2007.04.005.
- Helsedirektoratet (2014) *ADHD/Hyperkinetisk forstyrrelse – Nasjonal faglig retningslinje for utredning, behandling og oppfølging*. doi: 10.2307/2298845.
- Helsedirektoratet (2016a) *ADHD, Helsenorge*. Available at: <https://helsenorge.no/sykdom/psykiske-lidelser/utviklingsforstyrrelser/adhd> (Accessed: 24 October 2018).
- Helsedirektoratet (2016b) *Legemiddelbehandling av ADHD/ Hyperkinetisk forstyrrelse*. Available at: <https://helsedirektoratet.no/retningslinjer/adhd/seksjon?Tittel=legemiddelbehandling-av-adhd-hyperkinetisk-9281#ved-langtids-bruk-av-legemidler-for-adhd/hyperkinetisk-forstyrrelse-bør-effekter-og-bivirkninger-vurderes-regelmessig,-minst-årlig.anbefaling> (Accessed: 25 February 2019).
- Hewett, T. T. et al. (1992) *Curricula for Human-Computer Interaction ACM Special Interest Group on Computer-Human Interaction Curriculum Development Group*. Available at: <http://delivery.acm.org/10.1145/2600000/2594128/a1-hewett.pdf?ip=145.136.150.146&id=2594128&acc=NO> RULES&key=0C390721DC3021FF.4AD871FF6AD78CEE.4D4702B0C3E38B35.4D4702B0C3E38B35&__acm__=1529680811_ac1165f18b27bf82a73f8603e7557848.
- Highton-Williamson, E., Priebe, S. and Giacco, D. (2015) 'Online social networking in people with psychosis: A systematic review', *International Journal of Social Psychiatry*,

61(1), pp. 92–101. doi: 10.1177/0020764014556392.

Hobson, A. (ed.) (2004) *The Oxford Dictionary of Difficult Words*. New York, New York, USA: Oxford University Press Inc. Available at:

https://books.google.no/books?id=Vm_mNJiflwGC&pg=PA378&lpg=PA378&dq=rigor+the+quality+of+being+extremely+thorough,+exhaustive,+or+accurate&source=bl&ots=d2V_nRe9JO&sig=ACfU3U38jRFuRLEPBS5t66qDPLJHseAYig&hl=en&sa=X&ved=2ahUKEwjB4cmrzbBiAhUHwqYKHdKeDqkQ6AEwAHoEAcQAQ#v=onepage&q=rigor the quality of being extremely thorough%2C exhaustive%2C or accurate&f=false.

Hurvitz, N. (1970) 'Peer self-help psychotherapy groups and their implications for psychotherapy', *Psychotherapy: Theory, Research & Practice*, 7(1), pp. 41–49. doi: 10.1037/h0086549.

Intromat (2016) *Cognitive training in ADHD*. Available at: <http://intromat.no/cases/cognitive-training-in-adhd/> (Accessed: 8 September 2018).

Jain, M. *et al.* (2018) 'Evaluating and Informing the Design of Chatbots', *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*, pp. 895–906. doi: 10.1145/3196709.3196735.

Kaptelinin, V. and Bannon, L. J. (2012) 'Interaction design beyond the product: Creating technology-enhanced activity spaces', *Human-Computer Interaction*, 27(3), pp. 277–309. doi: 10.1080/07370024.2011.646930.

Kazdin, A. E. and Blase, S. L. (2011) 'Rebooting psychotherapy research and practice to reduce the burden of mental illness', *Perspectives on Psychological Science*, 6(1), pp. 21–37. doi: 10.1177/1745691610393527.

Kensing, F. and Blomberg, J. (1998) 'Participatory Design: Issues and Concerns', *Computer Supported Cooperative Work*, 7(3–4), pp. 167–185. doi: 10.1023/A:1008689307411.

Kooij, S. J. (2010) 'European consensus statement on diagnosis and treatment of adult ADHD: The European Network Adult ADHD', (1), pp. 1–24. doi: 10.1186/1471-244X-10-67.

Laudan, L. (1977) *Progress and Its Problems: Towards a Theory of Scientific Growth*. 1st edn. University of California Press.

Lawlor, A. and Kirakowski, J. (2014) 'When the lie is the truth: Grounded theory analysis of an online support group for factitious disorder', *Psychiatry Research*. Elsevier, 218(1–2), pp. 209–218. doi: 10.1016/J.PSYCHRES.2014.03.034.

Levine, B. *et al.* (2011) 'Rehabilitation of Executive Functioning in Patients with Frontal Lobe Brain Damage with Goal Management Training', *Frontiers Human Neuroscience*, 5(February), pp. 1–9. doi: 10.3389/fnhum.2011.00009.

- McKenna, K. and Bargh, J. (1998) 'Coming Out in the Age of the Internet: Identity "Demarginalization" Through Virtual Group Participation Katelyn', *Journal of Personality and Social Psychology*, 75(3), pp. 681–694.
- McKnight, L. (2011) 'Designing for ADHD: in search of guidelines', *Digital Technologies and Marginalized Youth*, 44(0). doi: 10.1080/21564574.2006.9635545.
- Mead, S., Hilton, D. and Curtis, L. (2001) 'Peer Support : A Theoretical Perspective', *Psychiatric Rehabilitation Journal*, 25(2), pp. 1–29.
- Naslund, J. A. *et al.* (2016) 'The future of mental health care: Peer-To-peer support and social media', *Epidemiology and Psychiatric Sciences*, 25(2), pp. 113–122. doi: 10.1017/S2045796015001067.
- Niela-Vilén, H. *et al.* (2014) 'Internet-based peer support for parents: A systematic integrative review', *International Journal of Nursing Studies*, 51(11), pp. 1524–1537. doi: 10.1016/j.ijnurstu.2014.06.009.
- Norsk Helseinformatikk (2017) *Behandling av ADHD*. Available at: <https://nhi.no/sykdommer/psykisk-helse/adhd/adhd-behandling/> (Accessed: 24 January 2019).
- O'Leary, K. *et al.* (2018) 'Suddenly, we got to become therapists for each other', *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–14. doi: 10.1145/3173574.3173905.
- Øie, M. G. (2015) *Hva er ADHD?* Available at: <https://www.psykologforeningen.no/publikum/videoer/videoer-om-psykiske-lidelser/hva-er-adhd> (Accessed: 20 March 2019).
- Østheim, A. (2011) *I Arbeid med ADHD: En gruppe voksne med ADHD forteller om sine arbeidserfaringer*. 1st edn. Møklegaard Print Shop.
- Oulasvirta, A. and Hornbæk, K. (2016) 'HCI Research as Problem-Solving', in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, New York, USA: ACM Press, pp. 4956–4967. doi: 10.1145/2858036.2858283.
- Pescosolido, B. A. *et al.* (2008) 'Public knowledge and assessment of child mental health problems: Findings from the national stigma study-children', *Journal of the American Academy of Child and Adolescent Psychiatry*. The American Academy of Child and Adolescent Psychiatry, 47(3), pp. 339–349. doi: 10.1097/CHI.0b013e318160e3a0.
- Pettersson, R. *et al.* (2017) 'Internet-Based Cognitive Behavioral Therapy for Adults With ADHD in Outpatient Psychiatric Care: A Randomized Trial', *Journal of Attention Disorders*, 21(6), pp. 508–521. doi: 10.1177/1087054714539998.
- Pliszka, S. (2007) 'Practice Parameter for the Assessment and Treatment of Children and

- Adolescents With Attention-Deficit/Hyperactivity Disorder', *Journal of the American Academy of Child & Adolescent Psychiatry*. Elsevier, 46(7), pp. 894–921. doi: 10.1097/CHI.0B013E318054E724.
- Preece, J., Rogers, Y. and Sharp, H. (2015) *Interaction Design - beyond human-computer interaction*. 4th edn. Chichester, West Sussex: John Wiley & Sons Ltd.
- Rouse, M. (2014) *Digital Divide*. Available at: <https://whatis.techtarget.com/definition/digital-divide> (Accessed: 24 May 2019).
- Rudd, J., Stern, K. and Isensee, S. (1996) 'Low Vs High'.
- Salzer, M. S. *et al.* (2013) 'Benefits of Working as a Certified Peer Specialist : Results From a Statewide Survey', 36(3), pp. 219–221. doi: 10.1037/prj0000016.
- Sehlin, H. *et al.* (2018) 'Experiences of an internet-based support and coaching model for adolescents and young adults with ADHD and autism spectrum disorder –a qualitative study', *BMC Psychiatry*. BMC Psychiatry, 18(1), p. 15. doi: 10.1186/s12888-018-1599-9.
- Sells, D. *et al.* (2006) 'The Treatment Relationship in Peer-Based and Regular Case Management for Clients With Severe Mental Illness', *Psychiatric Services*, 57(8), pp. 1179–1184. doi: 10.1176/ps.2006.57.8.1179.
- Skjuve, M. and Brandzaeg, P. B. (2018) 'Measuring user experience in chatbots : An approach to interpersonal communication competence'.
- Slack (no date) *Slack*. Available at: <https://slack.com/> (Accessed: 24 April 2019).
- Smestad, T. L. and Volden, F. (2018) 'Chatbot Personalities Matters Improving the user experience of chatbot interfaces'.
- Solomon, P. (2004) 'Peer Support/Peer Provided Services Underlying Processes, Benefits, and Critical Ingredients', *Psychiatric Rehabilitation Journal*, 27(4), pp. 392–401.
- Sonne, T. *et al.* (2016) 'An assistive technology design framework for ADHD', *Proceedings of the 28th Australian Conference on Computer-Human Interaction - OzCHI '16*, pp. 60–70. doi: 10.1145/3010915.3010925.
- Stappers, P. and Giaccardi, E. (2013) *The Encyclopedia of Human-Computer Interaction*. 2nd edn. The Interaction Design Foundation. Available at: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design>.
- Stephen, J. *et al.* (2014) 'Talking with text: Communication in therapist-led, live chat cancer support groups', *Social Science and Medicine*. Elsevier Ltd, 104, pp. 178–186. doi: 10.1016/j.socscimed.2013.12.001.
- Takahashi, Y. *et al.* (2009) 'Potential benefits and harms of a peer support social network

- service on the internet for people with depressive tendencies: Qualitative content analysis and social network analysis', *Journal of Medical Internet Research*, 11(3). doi: 10.2196/jmir.1142.
- Thies, I. M., Menon, N. and Magapu, S. (2017) 'How Do You Want Your Chatbot ? An Exploratory Wizard-of-Oz Study with Young , Urban Indians', 1, pp. 441–459. doi: 10.1007/978-3-319-67744-6.
- Turgay, A. *et al.* (2012) 'Lifespan persistence of ADHD: The life transition model and its application', *Journal of Clinical Psychiatry*, 73(2), pp. 192–201. doi: 10.4088/JCP.10m06628.
- Weizenbaum, J. (1966) 'ELIZA---a computer program for the study of natural language communication between man and machine', *Communications of the ACM*, 9(1), pp. 36–45. doi: 10.1145/365153.365168.
- Whitley, R. and Campbell, R. (2014) 'Stigma, agency and recovery amongst people with severe mental illness', *Social Science & Medicine*. Pergamon, 107, pp. 1–8. doi: 10.1016/j.socscimed.2014.02.010.
- Yang, D. *et al.* (2019) 'The Channel Matters: Self-disclosure, Reciprocity and Social Support in Online Cancer Support Groups', p. 12. doi: 10.1145/3290605.3300261.
- Yang, Q. *et al.* (2019) 'Sketching NLP : A Case Study of Exploring the Right Things To Design with Language Intelligence', *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, pp. 1–12.
- Ziebland, S. and Wyke, S. (2012) 'Health and Illness in a Connected World: How Might Sharing Experiences on the Internet Affect People's Health?', *THE MILBANK QUARTERLY A MULTIDISCIPLINARY JOURNAL OF POPULATION HEALTH AND HEALTH POLICY*, 90(2), pp. 219–249.
- Zimmerman, J., Forlizzi, J. and Evenson, S. (2007) 'Research through design as a method for interaction design research in HCI', *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, p. 493. doi: 10.1145/1240624.1240704.

Appendix A – Online Questionnaire

Opplevelse av 'Chat 1'

Dette skjemaet skal fylles ut etter at man har deltatt i 'Chat 1'.

1. 1. Hvordan opplevde du gruppesamtalen du nettopp deltok i?

2. 2. I hvor stor grad opplevde du kontakt med de andre deltagerne i chatten?

Mark only one oval.

	1	2	3	4	5	6	7	
Svært liten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Svært stor

3. 3. I hvilken grad opplevde du det som nyttig å diskutere 'Stopp-øvelsen' med andre i samme situasjon?

Mark only one oval.

	1	2	3	4	5	6	7	
Svært unyttig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Svært nyttig

4. 4. Har du erfaring med chatbot fra før? Hvis ja, vennligst utdyp i hvilken sammenheng og grad av erfaring.

5. 5. Hva var ditt inntrykk av Terabot (chatboten)?

Mark only one oval.

	1	2	3	4	5	6	7	
Svært negativt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Svært positivt

6. 6. Hvor forstyrrende opplevde du Terabot (chatboten) i samtalen?

Mark only one oval.

	1	2	3	4	5	6	7	
Svært forstyrrende	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Absolutt ikke forstyrrende

7. 7. Hvor nyttig opplevde du at Terabot (chatboten) var for samtalen?

Mark only one oval.

	1	2	3	4	5	6	7	
Svært unyttig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Svært nyttig

8. 8. Hvordan opplevde du at fordelingen mellom hvem som snakket/chattet var?

Mark only one oval.

	1	2	3	4	5	6	7	
Svært ujevn	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Svært jevn

9. 9. I hvilken grad opplevde du at du fikk sagt det du ønsket?


Mark only one oval.

	1	2	3	4	5	6	7	
Svært liten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Svært stor

10. 10. Var det noe du synes var bra med gruppechatten?

11. 11. Var det noe du ville ha forandret med gruppechatten?

Appendix B - Personas



Name:
Anja Petterson

Age:
22

Diagnosis:
ADHD inattentive type

Education:
Currently studying (3. year) of becoming a primary school teacher

Work:
Part.time job in a clothing store

Family:
Single. Lives by herself in a dorm. Has a good relationship to her parents. Single child.

Interests/Hobbies

- Interested in hanging with friends, but to a certain degree (takes energy)
- Take pictures and post them in social medias
- Likes to go shopping for clothes
- Often listen to audiobooks
- Loves walking her dog, especially in the mountains
- Watch TV shows

Behavior

- Have a hard time getting up in the morning, usually not before 11 am.
- Her mood is affected by lack of sleep.
- Often use a lot of time trying to study, but in fact just wast it on the computer.
- This gives her anxiety.
- Often have several social plans during the day.
- Sometimes forget dinner.

Frustrations

- Often have a hard time studying. Find it hard to start and prioritize.
- Is often disrupted by other ideas and thoughts.
- Has a lot of compulsive thoughts that affects her everyday life.
- Her sleep pattern is not good. Difficult getting to bed at night.

Powers

- Sometimes when interested in a subject she hyperfocuses.
- Good socially, and has a lot of friends.
- Is intelligent - has good grades even though study routine isn't good.
- Is doing well in social medias
- Good with kids and animals.
- Playful and funny.

Goals

- Establish a good study routine
- Make use of the day, not waste it on unimportant nonsense
- Get control of her compulsive thoughts, especially those considering sockets.
- Get a better sleep routine
- Say more 'no' when people ask to hang out.

Quote

"There is so much i feel i have to do everyday, but my schedule does not seem to find the time. This makes me very stressed. I also feel a lot of stress due to my studies and my compulsive thoughts."



Name:
Mikkel Andreasen

Age:
46

Diagnosis:
ADHD combined type

Education:
Vocational education as a car mechanic

Work:
Used to be a car mechanic, but is now under disability pension.

Family:
Married with two children (13 and 17 years old). Been married for 23 years.

Interests/Hobbies

- Interested in cars and all types of vehicles
- Likes to make art using old parts from vehicles
- Being creative in general
- Enjoys making food
- Playing football with the kids
- Test new technology

Behavior

- Wakes up 7am every morning
- He is always doing something
- Has a hard time relaxing
- Usually spend the day helping friends and neighbours with cars and other things
- Makes dinner everyday
- Spends the evening in the garage making some kind of creative project

Frustrations

- Has pain due to wear of the lumbar spine
- Can't manage to do everything he used to due to pain
- Can't seem to finish his projects
- Has inner turmoil because of not being able to do all that he used to
- Remembering details

Powers

- Has a lot of energy and is constantly doing something
- Very creative
- Very practically skilled
- Good socially
- Close to his family
- Funny

Goals

- Find some hobbies that are not very active
- Be able to find peace without being active
- Be better at remembering appointments and messages
- Get a more positive view on life

Quote

"I want to be able to feel inner peace more. I'm constantly doing something, and I have a hard time relaxing even though my body needs it."

Appendix C – Chatbot Script

*Note: The use of smileys is not represented in these scripts, as Microsoft Word is not customized for it.

Chat 1

Chatbot:

Hei [deltagernavn]

Jeg er Terabot, og det er jeg som skal veilede dere gjennom disse faste samtalene. Jeg kommer til å stille dere spørsmål som omhandler øvelsene dere gjennomgår i programmet.

Chatbot:

I disse samtalene ønsker jeg at vi skal:

Respektere hverandre.

Delta aktivt i samtalen når det passer.

La andre få snakke i fred mens de har ordet.

Gi beskjed om det er noe som plager oss.

Chatbot:

Jeg håper dette høres greit ut og at dere får nytte av dette opplegget. Dere må bare stille spørsmål om dere lurer på noe, så finner vi nok ut av det sammen

Chatbot:

Før vi begynner for fullt tenkte jeg vi kunne introdusere oss selv litt kort.

Jeg kan begynne!

Chatbot:

Jeg heter som sagt Terabot, og jeg er en chatbot.

Jeg jobber med å veilede gruppesamtaler og trives veldig godt med det.

Det jeg syns kan være litt utfordrende er at jeg ikke alltid forstår hva folk skriver siden jeg fortsatt er ganske ung, men da håper jeg de andre i gruppen kan hjelpe litt til. Ellers er jeg veldig glad i å se på kattevideoer på Youtube og å høre på musikk.

Insert ‘kattebilde’

Chatbot:

Hva med dere?

Deltagere svarer

Chatbot:

Så kjekt å bli bedre kjent med dere! Har dere noe til felles det de andre har nevnt?

Deltagere svarer

Chatbot:

Nå som vi alle har introdusert oss, er det på tide å begynne å snakke om øvelsen.

Øvelsen vi skal ha frem til neste chat er “Stopp-øvelsen”. Vi kan alle fra tid til annen ha behov for å gi oss selv tid til å tenke gjennom hva vi holder på med, hva vi ønsker å oppnå, og hvordan vi vil gå frem for å få til dette. Dette kan “Stop-øvelsen” hjelpe oss med.

Chatbot:

Hvordan forstår dere hensikten med denne øvelsen?

Deltagere svarer

Chatbot:

Godt tenkt! Slik jeg har forstått øvelsen så handler den om å roe ned og samle tankene sine. Ofte når vi holder på med noe er det lett for å gå på autopilot, og så glemmer man hva man egentlig holder på med eller skal. Det skjer i alle fall med meg.

Chatbot:

Personlig liker jeg å bruke denne øvelsen om morgningen når jeg har det travelt. Da bruker jeg den slik at jeg ikke skal glemme viktige ting mens jeg stresser rundt.

Chatbot:

I hvilke situasjoner i hverdagen deres kunne dere brukt denne øvelsen?

Deltagere svarer

Chatbot:

Det er spennende å lese om hvilke situasjoner dere tenker å bruke øvelsen. Jeg har lært noen nye jeg ikke hadde tenkt på før som jeg ønsker å teste.

Chatbot:

Av egen erfaring vet jeg at det kan være vanskelig å huske på å øve på slike øvelser som dette når hverdagen blir hektisk. Hvordan dere skal minne dere selv på øvelsen i dagene som kommer?

Deltagere svarer

Chatbot:

I dag har vi blitt litt bedre kjent med hverandre, forstått hensikten med øvelsen, diskutert hvilke situasjoner vi kan bruke øvelsen i, og hvordan vi skal huske på å utføre øvelsen i en travel hverdag. Jeg tror det er nok for i dag. Takk for i dag, så gleder jeg meg til å høre hvordan det har gått neste uke!

Chatbot:

Ps: Husk å tren på øvelsen frem til neste gang!

Hadebra

Chat 2

Chatbot:

Hei, og takk for sist!

Nå er jeg spent på å høre hvordan det har gått!

Chatbot:

Forrige uke forberedte vi oss på å trene på Stopp-øvelsen.

Hensikten med denne øvelsen var at vi skal stoppe opp og gi oss selv tid til å tenke over hva vi holder på med, hva vi ønsker å oppnå, og hvordan vi vil gå frem for å få til dette.

Chatbot:

I dag skal vi snakke om hvordan treningen har gått.

Først av alt: Har dere fått testet ut stopp-øvelsen siden forrige chat?

Deltagere svarer

Chatbot:

Dette høres allerede lovende ut! Ca hvor ofte har dere fått trent på øvelsen?

Deltagere svarer

Chatbot: Om noen IKKE har fått øvd på øvelsen

Kan du si litt om hvorfor du ikke har fått brukt øvelsen, @deltager_navn?

Chatbot:

Da får du bare følge med på hvordan de andre har opplevd treningen, og kanskje du lærer noe av det.

Chatbot:

Jeg synes dere har vært flinke! Kan dere fortelle litt om i hvilke situasjoner dere har øvd på den?

Deltagere svarer

Chatbot:

Ble det de samme situasjonene dere hadde sett for dere i forrige chat?

Deltagere svarer

Chatbot:

Jeg skjønner.

Var det noe dere opplevde som positivt med denne øvelsen mens dere trente?

Deltagere svarer

Chatbot:

Hvis ja: Det er alltid kjekt å høre at man får noe positivt ut av en øvelse!

Hvis nei: Det er ikke alltid man opplever noe nevneverdig positiv i øvelsene, men jeg håper dere lærte noe uansett!

Evt: Her seg jeg det var litt blandede erfaringer, men slik er det ofte. Noen øvelser faller i smak hos noen, og ikke hos andre - det er helt normalt.

+ dette etterpå:

Jeg lurar også på om dere opplevde noe som utfordrende med øvelsen mens dere trente?

Deltagere svarer

Chatbot:

Hvis ja:

Ja, jeg skjønner... Det kan være mange grunner til at vi opplever en øvelse utfordrende, men jeg tror det viktigste da er å øve i en grad og i et tempo som passer for oss og vårt liv ☺

Hvis nei:

Det er veldig kjekt å høre! Jeg liker når det går knirkefritt!

Her ser jeg det var litt blandede erfaringer, men slik er det ofte. Det kan være mange grunner til at vi opplever en øvelse utfordrende, men jeg tror det viktigste da er å øve i en grad og i et tempo som passer for oss og vårt liv ☺

+ dette etterpå:

Hvor relevant følte dere denne øvelsen var for dere?

Her kan dere gjerne utdype litt om begrunnelsen for svaret deres, for eksempel si litt om hvorfor den er relevant eller ikke i deres hverdag, eller lignende ☺

Deltagere svarer

Chatbot:

Er dette en øvelse dere kommer til å fortsette å bruke i hverdagen deres?

Deltagere svarer

Chatbot:

Føler du at du og dine chat-venner har opplevd denne øvelsen likt? Forklar gjerne hvordan du mener dere har forstått den likt eller ulikt

Deltagere svarer

Chatbot:

Jeg synes at dere alle har vært veldig flinke! Har dere noen avsluttende, positive kommentarer å komme med til dine chat-venner?

Deltagere svarer

Chatbot:

Dette har vært nok en hyggelig chat! Det har vært veldig lærerikt for meg, og jeg håper det har vært nyttig for dere også. Vi snakkes igjen, takk for meg.

Appendix D - NSD



NSD sin vurdering

Prosjekttittel

The design and potential of using peer support technology in an online self help program for adults with ADHD

Referansenummer

297609

Registrert

15.11.2018 av Oda Elise Nordberg - Oda.Nordberg@student.uib.no

Behandlingsansvarlig institusjon

Universitetet i Bergen / Det samfunnsvitenskapelige fakultet / Institutt for informasjons- og medievitenskap

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Frode Guribye , Frode.Guribye@uib.no, tlf: 41237111

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Oda Elise Nordberg, odaelisenordberg@gmail.com, tlf: 97067566

Prosjektperiode

06.08.2018 - 01.06.2019

Status

14.03.2019 - Vurdert

Vurdering (4)

14.03.2019 - Vurdert

Vi viser til endringer registrert i meldeskjema 12.3.2019. Endringene innebærer at det tilføyes et tredje utvalg, bestående av voksne med ADHD. Dette innebærer samtidig at det vil bli registrert særlige kategorier av personopplysninger om helseforhold, og behandlingens lovlige grunnlag må således oppdateres i henhold til dette. Den følgende vurderingen erstatter således våre vurderinger av 7.12.2018, 13.12.2018 samt 22.1.2019.

Det er vår vurdering at behandlingen vil være i samsvar med personvernlovgivningen, så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg 14.3.2019 samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan fortsette.

MELD VESENTLIGE ENDRINGER

Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html

Du må vente på svar fra NSD før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle alminnelige personopplysninger samt særlige kategorier av personopplysninger om helseforhold frem til 1.6.2019.

LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 nr. 11 og art. 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse, som kan dokumenteres, og som den registrerte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være den registrertes uttrykkelige samtykke, jf. personvernforordningen art. 6 nr. 1 a, jf. art. 9 nr. 2 bokstav a, jf. personopplysningsloven § 10, jf. § 9 (2).

PERSONVERNPRINSIPPER

NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet, vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER

NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1 f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og eventuelt rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET

NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Appendix E – NORSRII Abstract

ABSTRACT BOOK – NORSRII 2019 - PRESENTATIONS

Title

The design of a chatbot for guiding peer support chats between adults with ADHD

Presented by

Oda Elise Nordberg

Organization

Department of information science and media studies, University of Bergen

Authors

Oda E. Nordberg, Suresh Kumar Mukhiya, Frode Guribye, Eivind Flobak, Robin Kenter, Daniel A. Jensen, Astri Lundervold, Tine Nordgreen

Abstract

Peer support can be a valuable addition to treatments delivered as Internet-Based Interventions. The interactions between peers can take place in an online quasi-synchronous chat environment, but the chat should be guided to help the participants keep on topic and share relevant experiences.

This study explores how we can design a chatbot to provide guidance for peer support in groups participating in a web-based self-help program for adults with ADHD. The self-help program is inspired by the Goal Management Training program, which is presently under clinical evaluation in groups of adults with ADHD. The design of the chatbot is based on workshops with clinicians/domain experts and adults with ADHD to establish needs and requirements for the peer support and identify how a chatbot can facilitate and guide the conversation in an online quasi-synchronous group chat.

The first prototype has been tested in a field trial (with three groups of three participants) through using the software Slack and a Wizard of Oz technique. The aim was to test the format of the chat and to explore the potential of a bot guiding the group chat. In the trial the participants got an introduction to a stop exercise where they practice to stop and focus on the task at hand, then participated in one chat session to discuss expectations and their relevant preconceptions of the task, then they practiced for one week, and finally, they had a second group chat to discuss their experiences from doing the exercise. The data gathered in these sessions also serve as input to the design of chatbot functionality (such as identifying intents and keywords). The preliminary findings suggest that the participants were positive to the chatbot guiding them through the conversation. They also gave positive feedback on the format and found it useful to have a discussion related to the exercise.

The next iteration of design and evaluation will include a prototype of a chatbot that both guides and monitors the conversation (e.g. making sure everyone answers the questions within a certain time). This full version will be tested by five groups of three people and will include post-intervention interviews with all participants, and analysis of the interaction in the chatroom. The goal is to

ABSTRACT BOOK – NORSR II 2019 - PRESENTATIONS

understand how we can guide the group chat with a chatbot and how we can tie these peer support sessions to the self-help program and have the chatbot help the group to discuss relevant issues.

