Bachelor's thesis

NTNU
Norwegian University of Science and Technology

Candidate number: 10100

Chatbots as a Supportive Tool for Eating Disorders: A Scoping Review

Bachelor's thesis in Psychology
May 2022

NTNU
Norwegian University of
Science and Technology

Candidate number: 10100

Chatbots as a Supportive Tool for Eating Disorders: A Scoping Review

Bachelor's thesis in Psychology
May 2022

Norwegian University of Science and Technology

NTNU
Norwegian University of
Science and Technology

Chatbots as a Supportive Tool for Eating Disorders: A Scoping Review

Candidate number: 10100

Bachelor thesis in psychology: Mental Health Chatbots

PSY2900

May 2023, Trondheim

Norwegian University of Science and Technology

Supervisor: Lucas Bietti

Preface

This thesis marks the end of my bachelor's degree in psychology at the Norwegian University of Science and Technology. This study is emerged from a bigger research project, of scoping reviews of mental health chatbots. The chosen research question arose from my curiosity in the progress of technology that is becoming widely more available and how it may influence the health care system. Eating disorders affects the individual both physically and psychologically, but establishment of mental health chatbots may supplement the existing help resources. The increase of eating disorders amongst students and peers is concerning, and mapping the availability of low-threshold supportive help has been a rewarding experience.

I want to thank my supervisor Lucas Bietti for his guidance and support. His knowledge has helped me with the process of approaching and writing this thesis. I would like to thank my second reviewer, for assisting me in the retrieval of literature and analysis. As her unbiased view increased the reliability of my work. Also, I want to thank my fellow students of this research project, for the advice and conversations. Lastly, I want to thank my friends and family for their support and care.

Writing this thesis has been challenging, but also enriching. I have a greater understanding of how to conduct a review of literature, and a deeper understanding of the thesis theme. I hope the thesis reflects my interest in conveying insights on the theme, and that the reader will find it rewarding to read.

Abstract

Background: Chatbots are data programs that can simulate human conversations. These are becoming more common, also related to the field of health. Eating disorders is a highly prevalent mental illness, where research shows that adolescents and young adults are at high risk. Due to several barriers, individuals do not get the treatment they need. However, chatbots may be offered as a supportive tool.

Objective: This thesis aims to provide an overview of the existing literature related to chatbots targeted at eating disorders. Specifically looking into the features that deals with user experience of adolescents and young adults.

Method: Five databases were used in the search. Additionally, backwards reference list checking of relevant studies was conducted. Selection of sources were carried out by two reviewers. Data was synthesized through thematic analysis, leading to six themes related to this scoping review's purpose.

Results: Of 248 retrieved records, 8 studies were included in this review. These reported on both existing and upcoming chatbots, as well as recommendations prior development. All chatbots were rule-based, mostly accessible via mobile applications and conveyed psychoeducation. Most interventions targeted participants at risk of developing eating disorders. Effectiveness indicated increase of help-seeking post-intervention. Usability was limited in accuracy. However, satisfaction stated promising results and personalization was appreciated by users.

Conclusion: Chatbots targeted at eating disorders is a rather nascent area within healthcare, which needs further development. Per now, it may offer as a supplement in supporting adolescents and young adults in their struggles with eating problems and eating disorders.

Key words: Mental health chatbot, Eating disorder, User experience, Adolescents and young adults, Scoping review

Chatbots as a Supportive Tool for Eating Disorders: A Scoping Review

Eating disorders (EDs) is a highly prevalent mental illness, especially among adolescents (Beilharz et al., 2021; Galmiche et al., 2019). Several studies have shown a large increase of eating disorders among adolescents in recent years (Galmiche et al., 2019; Harrer et al., 2020; Nemesure et al., 2023). A recent systematic showed that the onset of anorexia, bulimia, and binge-eating disorder was most prevalent before the age of 24 years (Galmiche et al., 2019). Eating problems puts a severe strain on the individual's quality of life. EDs also represents a significant economic burden on a societal level, as it puts pressure on the health care system and one may be unable to work (Beilharz et al., 2021). Another meaningful factor is the strain put on the family and friends, as one of their loved ones struggles with an ED (Beilharz et al., 2021; Shah et al., 2022). Thus, the prevention of the disorder should be a public health priority (Galmiche et al., 2019; Harrer et al., 2020). Several risk factors are identified, for example negative body image (Beilharz et al., 2021; Chan et al., 2022; Galmiche et al., 2019; Shah et al., 2022). The identification of risk factors can support the work of early prevention in individuals who experience eating problems. A challenge facing the health care system are long waiting lists to see professionals, and the fact that many individuals don´t seek the help they need, due to reasons such as stigma and lack of knowledge (Beilharz et al., 2021; Rohrbach et al., 2022).

The advancement of today's technology progresses rapidly, and chatbots are used widely in several industries, for example in customer service. There is an emerging belief in the use of chatbots, also known as conversational agents, to complement the health care system (Chan et al., 2022; Fitzsimmons-Craft et al., 2022). Few studies have examined the user experience of digital conversational agents on the issue of mental health (Fitzsimmons-Craft et al., 2022). This thesis consists of a scoping review to map the available literature on chatbots and digital interventions that target individuals with eating disorders and eating

problems. Among the available chatbots that concentrate on mental health, there are only a limited number that concentrates on EDs (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022). With the increasing need for professional health care, it may be relevant to map the user experience of chatbots and analyse this in relation to further development. The research question of this scoping review aims to identify the user experience of young adults and adolescents, when interacting with chatbots and digital interventions targeted at eating disorders.

Objectives

According to the ICD-classification, which is the International Classification of Diseases, EDs are characterized by abnormal eating behaviour and preoccupation with food as well as prominent body weight and shape concerns (11th ed.; ICD-11; WHO, 2019, §06). Further on this may include deliberate weight loss and restricted eating, or in other cases, overeating with lack of control, purging and excessive training to counteract the "fattening" effect (11th ed.; ICD-11; WHO, 2019, §6B81). The symptoms depend on the specific ED however, presentation of each disorder will not be elaborated as this thesis focuses on EDs in general.

Research shows that there has been a significant increase in the prevalence of EDs, since the early 2000s (Galmiche et al., 2019; Nemesure et al., 2023). In a sample of adolescents, a meta-analysis showed an increase from 3,5% in 2000-2006, to 7,8% in 2013-2018 (Nemesure et al., 2023). In addition, many individuals show signs of an ED, but do not reach the criteria for diagnosis. The development from child to young adult is a risk factor, where especially many students report subthreshold eating problems (Harrer et al., 2020). These symptoms are important to detect early, because of the severe consequences when left untreated (Shah et al., 2022). Individuals with an ED do often neglect the severeness of illness and have compromised treatment motivation (Fitzsimmons-Craft et al., 2022; Shah et al., 2022). This can have longstanding negative effects on the individual, such as comorbid

mental disorders and somatic conditions, giving poor quality of life (Beilharz et al., 2021; Chan et al., 2022; Galmiche et al., 2019; Harrer et al., 2020). This shows the need for early intervention, as there is a clear treatment gap once the individual has developed an ED (Fitzsimmons-Craft et al., 2022).

A major risk factor is dissatisfaction with body image (Beilharz et al., 2021; Galmiche et al., 2019). Concerns surrounding body image are so widespread that they are often considered natural among women, and can cause significant distress (Nemesure et al., 2023). This reflects the high prevalence of EDs among women and girls (Beilharz et al., 2021). In a study conducted by Chan et al. (2022) about 23% of college-age women in America had levels of weight and shape concerns that put them in risk for EDs. There is a thin-ideal internalization in the society, meaning that individuals "buys into" ideals of attractiveness that are defined by the society (Fitzsimmons-Craft et al., 2022). This concept is enhanced by image sharing on social media (Beilharz et al., 2021). This risk is therefore related to the age of the population, given that the majority of young adults are active on social media platforms (Beilharz et al., 2021). The risk factors mentioned above are cultural conditions, whereas differences across countries may be due to various genetics and eating behaviours (Galmiche et al., 2019).

Several studies show that many individuals don´t receive the health care they need, due to for example, reluctance for help seeking and stigmatization (Beilharz et al., 2021). This is worrying, since a longer duration of untreated illness seems to indicate poorer outcome (Beilharz et al., 2021; Rohrbach et al., 2022). A challenge when diagnosing adolescents is that psychologists need to be aware that they are under development, going through changes both physically and psychologically. The youth years are demanding for some, and adolescents may resort to various ways to deal with difficult thoughts related to food and the body (Rojewska et al., 2022). The years spent in university is shown to be a

crucial period in young adults' lives, as it coincides with the peak of many mental and behavioural disorders (Harrer et al., 2020). The individuals´ life situation needs to be balanced with the recommended treatment guidelines for EDs, which says that early prevention strategies give the best prognosis (Beilharz et al., 2021; Shah et al., 2022). The strongest therapeutic approach in treating EDs is cognitive behavioural treatment (CBT-E), where the core concept is addressing the overevaluation of weight and shape (Beilharz et al., 2021). In general, there is a lot of misinformation related to "healthy" eating and weight (Beilharz et al., 2021). A typical feature of an individual suffering from an EDs is the belief that one's body size is irrational compared to what is true (Beilharz et al., 2021). Psychoeducation is a method in CBT-E which communicates relevant information about EDs and is associated with small yet significant improvements on body image (Beilharz et al., 2021).

In line with the increase of mental illness, the demand for psychologists also rises (Rojewska et al., 2022). Consequently, the psychologist lacks time and resources to devote for each individual patient, and this is one external barrier for why many with EDs go untreated (Fitzsimmons-Craft et al., 2022; Harrer et al., 2020; Nemesure et al., 2023; Rojewska et al., 2022; Shah et al., 2022).

Other barriers for seeking traditional health services include stigmatization, where individuals may feel uncomfortable in reaching out (Beilharz et al., 2021). Ali et al. (2017) presented that stigma and shame was the most frequent barriers for help-seeking. A reason for this tendency is that many interpret eating disorder as a choice of lifestyle, and not an illness on a par with other mental disorders (Ali et al., 2017). This is an internal-external factor, as it indicates lack of information among the public, which rather leads individuals to neglect the severeness of their illness (Ali et al., 2017; Fitzsimmons-Craft et al., 2022; Shah et al., 2022). An anonymous digital agent to converse with may feel less intimidating, and

more approachable, especially for younger people who grew up as part of a digital generation (Beilharz et al., 2021; Chan et al., 2022). Hence is the development of digital interventions and chatbots highly relevant and important in closing the current treatment gap (Beilharz et al., 2021; Fitzsimmons-Craft et al., 2022).

Practical factors like economical concerns, are a barrier hindering individuals in contacting professional help (Ali et al., 2017; Beilharz et al., 2021). Especially this difficulty is seen among people from rural areas (Ali et al., 2017). The use of digital interventions is a low or no-cost offer for the users (Beilharz et al., 2021). Additional benefits are that the psychoeducation provided by digital services are evidence-based and considered more reliable than information in other social media (Beilharz et al., 2021). In some cases, the use of digital intervention aims to be a steppingstone for individuals to seek other supportive services (Beilharz et al., 2021; Fitzsimmons-Craft et al., 2022; Rohrbach et al., 2022; Shah et al., 2022).

Chatbots are computer programs that simulate human conversations (Shah et al., 2022). Besides interacting by writing, they can engage orally and visually as well. The use of such digital tools are already widely used, for example Siri and Alexa (Chan et al., 2022). They are becoming more common in the world of health and medicine, as chatbots can provide an easy-accessible platform for psychoeducation and coping skills (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022). For example, Beilharz et al. (2021) presents the development of a chatbot named KIT, which is based on CBT-E. KIT is targeted to people with body image problems and eating concerns, including information for those who seek help to helping others. The format of chatbots intent to provide an easily accessible resource, which can supplement the traditional support services (Beilharz et al., 2021).

One approach to development is rule-based chatbots, which are programmed to generate responses according to certain keywords (Beilharz et al., 2021; Chan et al., 2022;

Fitzsimmons-Craft et al., 2022; Shah et al., 2022). In some cases, the chatbot may also be triggered by specific keywords. If the chatbot asks a question, to which the user responds "hurting myself", the user will be provided with a referral to the crisis hotline (Fitzsimmons-Craft et al., 2022). These sort of chatbots are cost-efficient, as they are easy to make. In addition, they are revisable, and their responses are controllable by the developers and professionals (Chan et al., 2022; Shah et al., 2022). On the downside, this may limit the interaction, as the conversation is predefined (Chan et al., 2022). Another approach uses artificial intelligence (AI) to generate responses. The AI learns progressively by algorithms that search through all digitalized information, such as from prior chatbot interactions, and collecting the ratings of former users and experts (Chan et al., 2022). This method enables a more fluid and natural conversation, as the chatbot is able to respond on inputs from the individual. However, the AI-based approach demands more resources (Chan et al., 2022; Shah et al., 2022). Rule-based chatbots are more commonly used in treatment of EDs, as they are more predictable, thus protecting the user (Chan et al., 2022; Rojewska et al., 2022).

A legitimate concern when relying on digital conversational agents is that they may respond inappropriately and potentially reinforce harmful behaviours (Chan et al., 2022). An example of this could be misunderstanding a user's irony when asked to complement his or her own body, saying "I liked it when I was under weight and could see my bones". This can possibly lead the chatbot to further praising and statements like "… let´s keep working on making you feel this good more often" (Chan et al., 2022). This makes chatbots less justifiable in interaction with patients and other individuals, especially AI-based which are not controlled by experts (Shah et al., 2022). Challenges like these are under development, as experts tries to craft responses that "always work", by being more neutral while maintaining a warm tone (Chan et al., 2022). Another feature about chatbots is that they are developed to keep the response length short, as this is in line with the current texting culture (Fitzsimmons-

Craft et al., 2022; Shah et al., 2022). This is pruned towards youths and may make the user more likely to stay committed during the digital intervention (Fitzsimmons-Craft et al., 2022; Shah et al., 2022).

Methods

To fulfil the objective of examining the user experience of young adults and adolescents interacting with chatbots targeted at eating disorders, this thesis carried out a scoping review. This approach is more appropriate than a systematic review, due to the relatively new and broad nature of the research question. A scoping review synthesizes knowledge through mapping evidence on a specific topic, including ongoing research (Tricco et al., 2018). This is done by systematically identifying main concepts, theories, sources, and knowledge gaps in the available research literature (Tricco et al., 2018). This is a suitable method to provide scientific evidence on the effect of digital mental health interventions (Hutton et al., 2016). A scoping review should facilitate transparency, making it possible for other researchers to gather the same results. This is enhanced by following the guidelines recommended by the Preferred Reporting Items for Systematic reviews and Meta-Analysis extension for Scoping reviews (PRISMA-ScR). The aim of PRISMA-ScR is to provide the readers with a greater understanding of relevant terms, core concepts, and key items of the scoping review (Tricco et al., 2018).

Eligibility criteria

This study follows PCC, which stand for population, concept, and context. PCC defines the criteria which studies had to convey, to be included. The chosen population is adolescents and young adults. Subjects was considered within this age group between 13 and 30 years old. The concept are EDs, which includes individuals who self-report being at risk, as well as individuals clinically screened for EDs. The context is the user experience of chatbots and

digital interventions that enable interaction between users and conversational agents, including literature that aims to improve the user experience. The review included empirical studies therefore, it excluded other scoping and systematic reviews. There were no language restrictions in the data search, and no restrictions were set to the geographical location of where the studies were conducted. Given that technology progresses rapidly and with increased familiarization of AI, it was necessary to restrict the publication year to the period between 2018 and 2023.

Information sources

For the purpose of this scoping review, the following databases were used in conducting the search: Scopus, PubMed, Web of Science, ACM, and Google Scholar. The same parameters (which will be described below) were also used to search MedLine with both MeSh terms and keywords singularly, but it returned zero results. Only the first 100 citations from Google Scholar were revised, because the retrieved sources are sorted based on relevance to the search topic and is further considered more of a supplement. It felt useful to use one database which focuses on technology, and therefore the Association for Computing Machinery (ACM) was used. The search was done between the 16th and 17th of February 2023.

Search

The search terms used was based on three topics, according to the PCC. First was the use of "chatbot*", where additional terms were "conversational agent*", "chatterbot*", "virtual agent*" OR "embodied conversational agent*". Then AND "adolescent*" OR "student*" was added. Finally, it was AND "eating disorder*". The search terms associated to chatbot was derived from previous reviews and in cooperation with peers. The aim of including adolescent* or student* as terms was to meet the desired age group. The search on EDs did not use terms on specific diagnoses, because the focus of the thesis is on eating problems and

illness in general. This procedure was repeated similarly in all databases used in this scoping review.

Selection of sources of evidence

In selecting the studies for this review, the titles and abstracts of the retrieved studies were first screened independently by myself and a second reviewer who was familiar with the overall project but was not associated with my particular study. There was disagreement between the reviewers regarding the relevance of some articles. This issue was resolved by scanning through the full text and reaching agreement according to the PCC criteria. This process is presented in the PRISMA flow diagram (figure 1). In-depth reading of the included articles was reviewed by me and a second reviewer, which resulted in exclusion of articles that did not represent the eligibility criteria. Some relevant studies might have been published in other journals that are not indexed in the databases used in this review´s data search. As the object of this study is user experience of chatbots focusing on EDs, I chose to examine the reference list of the publications that exclusively reported on EDs and chatbots. According to the headings, no further results were retrieved.

Data charting process

Thematic analysis (TA) is a method which is often used within research focusing on health (Braun & Clarke, 2014). It offers flexibility when approaching the analysis and provides a rich interpretation of the data (Braun & Clarke, 2014). Initially, the first step of this method was to get familiar with the articles by reading them multiple times and then transcribing the text that appeared interesting according to the research question. Further on, the relevant text was generated into codes. By colour coding the codes, themes were identified, and made patterns across the articles more visible. This was a recursive process, where moving back and forth in the steps of analysis ensured that the themes represented the data. When doing

this, I always had the research question in mind. This process was is in line with the guidance on thematic analysis, as stated by Braun and Clarke (2006). However, according to the project description of this thesis, which places some demands on the area of focus, some themes were predefined (e.g. effectiveness). Also, the PCC influenced some aspects of interest. In this manner, the TA was slightly modified, as it relied on using both a deductive and inductive approach in synthesizing the data.

According to the guidelines from Joanna Briggs Institute (JBI) for conducting ScR, data extraction of the articles should be coded by a second reviewer (Peters et al., 2017). Due to the time restrictions put on this thesis, coding by a peer was limited to 20% of the dataset. This was decided in consultation with my supervisor.

Data themes

A theme that was specified on beforehand was “effectiveness”, as this is relevant for the objective of the thesis, in assessing the outcome of interaction with the digital intervention. “Usability” is another theme that was predetermined, as there is an interest in the chatbots abilities to secure safety and trust.

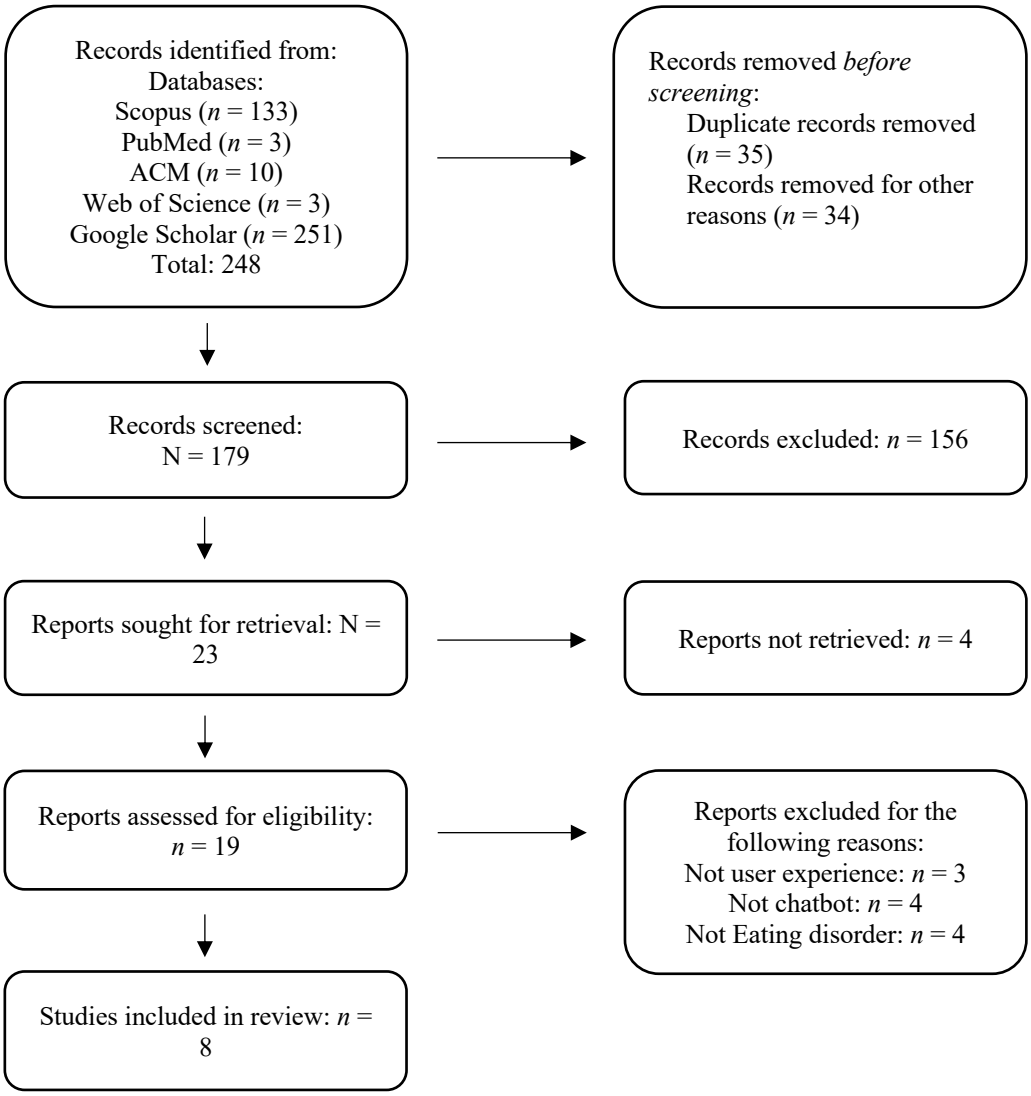
The other themes became visible after data extraction and further detection of patterns. By this, the themes of “treatment”, “accessibility”, “personalization” and “satisfaction” were developed inductively. Themes and associated content will be explained in the results section.

Results

In this section, the process of literature retrieval will be described in detail. The characteristics of each individual article will be presented, clarifying the prevalence of ED chatbots and description of their procedures. Drawn from the thematic analysis, each theme will be explained and elaborate upon the various aspects of user experience.

Figure 1

Flow chart of the study selection process



Selection of sources of evidence

The following process description is shown in figure 1. After searching through five bibliographic databases, 248 citations were retrieved. After removing 35 duplicates and 34 reviews, 179 articles were screened for inclusion. Another 156 citations were removed after reading titles and abstracts. Full text of the remaining 23 publications were read, and this led to the eligibility assessment of 19 articles. 3 articles were withdrawn for not focusing on adolescents, 4 articles for not including chatbots, and another 4 for not looking explicitly on

eating disorders. After the final step, a total of 8 articles were included in the synthesis for this scoping review.

Inter-rater agreement between reviewers was measured using Cohen´s kappa. The value after assessing titles and abstracts was .85, and .83 after reviewing full text in the second step of the selection process. This indicates strong reliability in agreement.

Characteristics of sources of evidence

Table 1. *Characteristics of Sources of Evidence*

Article	Author(s) and year	Country of origin	Study design	Population and sample size	Description and procedure
1	Beilharz et al. (2021)	Australia	Qualitative focus group study	*N* = 17 (excluded 8 in the group for parents/carers) Age: 13-18 Male: 5 Female: 10 Transgender: 1 Other: 1	Embodied, rule-based chatbot. Designed to provide support for people with concerns regarding body image and eating issues.
2	Chan et al. (2022)	USA	Observational study	*N* = 2409 Age: 18-30 Gender: woman	Rule-based chatbot. Intervention for women at risk. Researchers reviewed comments from users, to identify problems and generate possible solutions.
3	Fitzsimmons-Craft et al. (2022)	USA	Randomized clinical trial	*N* = 700 Mean age: 21 Gender: Women	Rule-based chatbot for women at risk for ED. Participants were followed for six months.
4	Jarman et al. (2022)	Australia	Design thinking	*N* = 7 Mean age: 26 Gender: women	Pre-test and co-design of a rule-based chatbot, with end users before randomized controlled trials.
5	Kasson et al. (2021)	USA	Exploratory research design	*N* = 44 Age: 14-17 Gender: Mainly female (100% in	Rule-based chatbot. Enrolling participants from social media.

				discovery group, 74% in testing group)	
6	Matheson et al. (2021)	Brazil	Describes a Randomized controlled trial	*N* = 2800 Age: 13-18 Gender: 56% girls	Rule-based chatbot. Designed to provide population-wide benefits, regardless of risk for ED.
7	Rohrbach et al. (2022)	Netherlands	Randomized controlled trial	*N* = 355 Mean age: 27 Female: 343 Male: 9 Other: 3	Comparison of 4 condition: Rule-based chatbot, chatbot plus expert-patient support, expert-patient support only, and a waiting list control condition. Participants were followed for 12 months.
8	Shah et al. (2022)	USA	Qualitative analysis	*N* = 21 Age: above 18 Female: 20 Male: 1	Usability testing among young adults who screened positive for a clinical or subclinical ED in interaction with a rule-based chatbot.

As presented in table 1, all the studies were published between 2021 and 2022 (*N*=8). Nearly all the studies were journal articles (*n*=7), except for one protocol article. Studies was published across 4 countries, and half of them were published in the USA (*n*=4). The most occurring study design was randomized controlled trial (*n*=3), whilst qualitive approaches was also common (*n*=3). The sample size was below 50 participants in half of the studies (*n*=4) and 355 or more in the remaining (*n*=4). Participants of the studies were in the age ranging from 13 to 30, except one article that included individuals of 33 in age (Jarman et al., 2022). However, the mean age was 25 years old, and therefore it was reasoned for eligibility. In sex, there was a predominance of women, with 3 studies limited to women only, and another 2 with mainly women (*n*=5). Only two of the studies is based upon existing chatbots (Beilharz et al., 2021; Fitzsimmons-Craft et al., 2022), while the remaining studies focuses on acceptability and usability in development of an ED related chatbot (*n*=6). The type of digital intervention in all studies, were based on rule-based computer science (*n*=8). Only one

chatbot was embodied (Beilharz et al., 2021). The most common assessed outcomes were relating usability (*n*=6) and aspects of satisfaction (*n*=7), while 3 studies mainly sought to assess the effectiveness.

Results of individual sources of evidence

Table 2

Purpose and Key Findings of Individual Sources of Evidence

Article	Publication	Purpose	Key findings
1	Beilharz et al., 2021	The aim of this study was to assess the preliminary acceptability and feasibility, of a chatbot´s (KIT) content, structure and design. This was collected via qualitative feedback from young people and parent/carers.	Participants were overall positive regarding the content, structure, and design of the body image chatbot. Specific suggestions were made regarding length and tone of responses to increase interactivity.
2	Chan et al., 2022	Aims to present examples of and solutions to challenges in designing and refining a rule-based prevention chatbot program for EDs. Targeting women at risk.	Rule-based chatbots have the potential to reach a large population, at low cost. In general, the chatbot had limitations in its ability to respond appropriately to unanticipated user responses.
3	Fitzsimmons-Craft et al., 2022	Aims to test whether a chatbot would reduce ED risk factors in women at high risk for an ED, compared to a waitlist control.	There was a significantly greater reduction in intervention versus control. The odds of remaining nonclinical for EDs were higher in intervention versus control, at least in shorter-term.
4	Jarman et al., 2022	Provide the context and a lived experience perspective of the end user, in development of a chatbot.	Practical guidance and recommendations for those developing app-based ED interventions. Highlights the importance of identifying symptoms of eating disorders. Participants emphasize personalization and usability to promote continued engagement and recovery.
5	Kasson et al., 2021	Aimed to identify characteristics of teens with EDs and gather feedback on an mHealth intervention.	There was a high acceptability of and interest in an app to support ED recovery. Use of social media was a feasible method for recruiting teens with EDs.
6	Matheson et al., 2021	Co-creation and evaluation of a body image chatbot, for Brazilian adolescents.	The current research addresses an unmet need for evidence-based body image resources for Brazilian adolescents.
7	Rohrbach et al., 2022	Providing a low-threshold intervention to help bridge the treatment gap. Evaluation of Featback, a fully automated online self-help intervention.	Support from Featback, support from an expert patient and the combination of both were significantly more effective

			compared to a waiting list controller condition. Participants reported higher satisfaction in condition with human contact, but there was no difference in effectiveness.
8	Shah et al., 2022	Development of a chatbot, Alex, as a tool to improve motivation for treatment and self-efficacy among individuals with EDs.	Participants generally reflected positively on interaction with the chatbot, regarding the usability and acceptance, even prior to refinements. Effectiveness in help-seeking was accomplished.

In the following subsections, each of the themes generated from the analysis of the included publications will be elaborated. Each individual study's purpose and key findings is presented in table 2.

Effectiveness

In relation to the research question, the theme of "effectiveness" was mapped first. This includes measuring if users can achieve their objectives in interacting with the chatbot. The theme was emphasized in five out of the eight articles analysed in this scoping review (Beilharz et al., 2021; Fitzsimmons-Craft et al., 2022; Matheson et al., 2021; Rohrbach et al., 2022; Shah et al., 2022). The sample size of each study is listed in table 1. Effectiveness was tested through follow-up interviews and questionnaires in all five studies, apart from Chan et al. (2022), who conducted a review of user comments. Specifically, Fitzsimmons-Craft et al. (2022) and Rohrbach et al. (2022) measured ED symptomatology post-intervention by the Eating Disorder Examination Questionnaire (EDE-Q), which assesses the range and severity of features associated with an ED diagnosis.

Enforcing motivation and help-seeking is the objective in three of the studies, as this attempts to reduce barriers such as stigmatization and neglect of illness (Beilharz et al., 2021; Matheson et al., 2021; Shah et al., 2022). Shah et al. (2022) develops a chatbot that aims to improve motivation for treatment. In this study, effectiveness was accomplished as the

chatbot reinforced help-seeking behaviours among participants. This also applies to Beilharz et al. (2021), as the chatbot provided information that influenced the participants' perception of ED. However, the study stated that further research is required on the topic, and the same holds true in the study protocol by Matheson et al. (2021).

Another aspect of the theme was evaluating the quality of the output that users gained from the intervention. The key findings of each study are listed in table 2. In both Fitzsimmons-Craft et al. (2022) and Rohrbach et al. (2022) the intervention was significantly more effective in reducing ED symptoms, than in the condition for waitlist control, where individuals received no support. Despite this, the difference was no longer significant in the 6-month follow-up, compared to 3-month follow-up (Fitzsimmons-Craft et al., 2022). The same tendency is shown in Rohrbach et al. (2022), as the long-term effect of intervention was not significant. This indicates lack of quality of the effectiveness, in the studies that intended to reduce ED risk factors (Fitzsimmons-Craft et al., 2022; Rohrbach et al., 2022).

Usability

The theme of "usability" refers to the use of the chatbot, in relation to the technology. All chatbots mentioned in the conducted studies are (or will be) programmed based on rules, in cooperation between experts in the fields of psychology and computer science, including individuals with lived experiences of EDs (Beilharz et al., 2021; Jarman et al., 2022). Additionally, Beilharz et al. (2021) designed an embodied character for conversation. Six out of eight of the analysed studies deal with usability (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Jarman et al., 2022; Matheson et al., 2021; Shah et al., 2022). Usability testing was conducted through various ways, such as review of comments from users (Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Matheson et al., 2021), TA of interviews (Beilharz et al., 2021; Jarman et al., 2022), and through the System Usability Scale (SUS) (Shah et al., 2022).

Safety is an essential feature of usability, as it is important verify that the chatbot is trustworthy. The established chatbots from Beilharz et al. (2021) and (Fitzsimmons-Craft et al., 2022) include a protocol for crisis, and Chan et al. (2022) plans to include it in the finished product. If the user responds with phrases such as “hurting myself”, the chatbot will be triggered to activate a crisis module, referring to an emergency hotline (Chan et al., 2022). The same function was incorporated in the chatbot prototype by Shah et al. (2022). In Beilharz et al. (2021), participants found the conversational character safe and unlikely to trigger body image concerns, as it had a non-human like design. The focus of feeling safe during the intervention was also highlighted in Matheson et al. (2021).

Accuracy in response is also a relevant quality for developing trust, which is another important aspect of usability. Users find it frustrating when they are not able to enter their own input to the conversation, (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Jarman et al., 2022; Shah et al., 2022). Chan et al. (2022) stated that this may lead to misinterpretations between the user and the chatbot, and in worst case reinforce harmful behaviours (Fitzsimmons-Craft et al., 2022; Shah et al., 2022). This is a limitation set by the rule-based chatbots, as well as lack of sufficiency in the current level of AI development (Beilharz et al., 2021). The ease of use and accuracy of response was also a topic in Matheson et al. (2021).

Treatment

Across the dataset, a pattern of approaches regarding “treatment” emerged. This theme is relevant for the development of the chatbots, and the content that it wishes to deliver towards the user (purpose of each study are listed in table 2). In order to justify the use of non-human resources in health care, it needs to be reliable and based upon evidence-based science. All the articles included in the review addresses an approach that is related to

cognitive behavioural therapy, which is considered a strong therapeutic approach in treatment of EDs (Beilharz et al., 2021).

The approach of psychoeducation, which is subject to CBT, is incorporated in Beilharz et al. (2021), Fitzsimmons-Craft et al. (2022), Jarman et al. (2022), Rohrbach et al. (2022) and Shah et al. (2022). The aim of this is to communicate more resourceful knowledge about EDs and help overcome barriers towards recovery. Additionally, Shah et al. (2022), included another component of CBT, motivational interviewing, which was reported to enhance help-seeking in participants. Changes in perception of EDs was observed in participants (Fitzsimmons-Craft et al., 2022; Rohrbach et al., 2022; Shah et al., 2022), stated by data collected post-intervention. Kasson et al. (2021) and Matheson et al. (2021) specifically aimed to provide better information about EDs than presented in social media, challenging perceptions related to thin ideals. However, the studies stated that this need further research before results may be presented.

Accessibility

This theme refers to accessibility in relation to eligible participants and how the intervention is exhibited for use. This is an important theme if the chatbot intends to fill the current treatment gap, both in lack of health resources and providing a less intimidating platform. The theme was retrieved out of all articles, except for Jarman et al. (2022).

The purpose of the intervention influences which participants are eligible, regarding the symptom level and severeness of the eating issues. In five out of eight studies, the aim was to provide a supportive chatbot for young adults at risk for ED, that may be a bridge to further treatment if needed (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Matheson et al., 2021; Rohrbach et al., 2022). By this, participants were eligible if they had mild symptoms of, or self-identified being at risk for developing an ED. In Rohrbach et al. (2022) serious cases of ED were eligible, but recommended to seek professional help,

while Fitzsimmons-Craft et al. (2022) excluded participants of clinical level. Kasson et al. (2021) and Shah et al. (2022) intended to included participants that scored positive on the Stanford-Washington University ED Screen (SWED), which is a tool that provides good sensitivity and specificity for most ED diagnoses (Kasson et al., 2021). These two studies had a different intention of the intervention, wanting to map characteristics of teens with ED (Kasson et al., 2021) and improving motivation for treatment (Shah et al., 2021).

Social media is considered a promising avenue for recruitment of young adolescents, as well as providing the chatbot service (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Kasson et al., 2021; Matheson et al., 2021). Four of the studies reported access to the chatbot through SMS or Facebook Messenger (Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Matheson et al., 2021; Shah et al., 2022). Kasson et al. (2021) developed an app, and in Rohrbach et al. (2022) interaction is given through questionnaire and e-mail.

Personalization

Across the articles, there is mentioned a desire for the incorporation of "personalization". This theme introduces elements related to language and text format, which promotes further use if the user finds it engaging. Personalization was addressed in seven out of the eight studies (Beilharz et al., 2021; Chan et al., 2022; Fitzsimmons-Craft et al., 2022; Jarman et al., 2022; Kasson et al., 2021; Rohrbach et al., 2022; Shah et al., 2022). Personalized content is one approach of chatbots that is challenging, since it is not capable of interpretating responses in a human like manner (Beilharz et al., 2021; Jarman et al., 2022). However, Shah et al. (2022) included resources in the intervention that offered tailored recommendations based on treatment preferences in the individual. Rohrbach et al. (2022) based a weekly feedback message on participants self-reported ratings of ED symptoms.

Another method to provide a feeling of personalized content in chatbots, is presenting the text in a familiar format. Beilharz et al. (2022), Chan et al. (2022), Fitzsimmons-Craft et al. (2022) and Kasson et al. (2021) programmed the chatbots responses to be in line with current texting culture by e.g., shortening the length of text and rather sending multiple messages, and including the use of emojis. According to feedback from participants, this felt like receiving a message from a friend (Beilharz et al., 2022). In extension of these adjustments, the language of responses was set to be conveyed in a warm, but neutral tone (Chan et al., 2022; Fitzsimmons-Craft et al., 2022).

Satisfaction

Aside from the effectiveness that can be measured, the subjective reaction of the user is also of importance. This theme reviews to what extent the participants found the experience satisfactory, as in wanting to use it and potentially recommend it to others (Lund, 2001). Feedback related to this theme is needed for further development of chatbots, reassuring its relevance and necessity.

Satisfaction was assessed in three of the studies (Beilharz et al., 2021; Rohrbach et al., 2022; Shah et al., 2022). Satisfaction was measured through TA in Beilharz et al. (2021) and Shah et al. (2022), which also included the Ease of Use Questionnaire looking at satisfaction of users (Lund, 2001). This score was generally above average (approximately equal to 5 out of 7) (Shah et al., 2022). Rohrbach et al. (2022) used a self-developed questionnaire. Focus group interviews post-intervention indicated that participants believed the chatbot would appeal to a wide audience of users (Beilharz et al., 2021), and interview feedback stated that participants found the chatbot enjoyable (Shah et al., 2022).

Rohrbach et al. (2021) compared a chatbot alone condition with a condition including feedback from expert-patients (patients previously suffering from ED and recovered through

treatment). Report from the participants showed that participants were significantly more satisfied with the expert-patient conditions, than chatbot alone.

Discussion

Summary of evidence

In this scoping review I did a thematic analysis that led to six themes which mapped the results of each individual study. Effectiveness and usability were predetermined, as these perspectives were required according to the research question. The most prevalent themes were treatment and personalization, which together with accessibility and satisfaction reflect user experience features.

Effectiveness was evaluated based on whether participants could reach their objective in interaction with the chatbot. This was assessed in five of the studies. The overall results post-intervention indicated improvement of ED symptomology, but this was not significant after a period of time. This suggests that digital interventions may be better suited as a way to initiate help-seeking rather than as a sole treatment method. The effectiveness of chatbots may be influenced by the content and type of treatment, as CBT has shown promising results in reinforcing help-seeking. Therefore, the results stated under “treatment” was promising, as all the conducted studies were based on CBT. Further, this may facilitate more sustainability in effectiveness.

The usability of chatbots was looked upon in six of the articles, where safety and accuracy were prominent aspects. Five of the studies included a module for crisis, that would be triggered by identifying certain keywords. Thus, this can be assumed to be a highly significant protocol for safety. The aspect of accuracy was also considered, as all the retrieved studies presents rule-based chatbots. The feature was assessed to be limited by the current technology. One argument for the use of rule-based chatbots, is that they are easier to

revise and control. However, another reason for this approach is based upon economical aspects. Development of an AI-based chatbot demands a lot of time and resources, which is more costly. Considering ethical aspects, arises a matter if the benefits of AI justify the cost and potential risk associated with their development.

The theme of "treatment" presented evidence-based methods in treatment of EDs, showing that CBT is proven to be efficient. This knowledge is important for developing chatbot content that ensures reliable methods in approaching individuals with eating concerns and ED, which was mapped in all eight studies. Shah et al. (2022) stated that a user felt that their challenges were not severe enough and expressed feelings of guilt in using the chatbot. This represents an overall barrier in help-seeking, as individuals do not recognize their eating problems as a consequential issue (Ali et al., 2017; Fitzsimmons-Craft et al., 2022; Shah et al., 2022). Thus, five of the studies were based upon psychoeducation, reflecting an overall focus across the dataset on communicating knowledge about EDs. Several studies saw strategies for "busting myths" about EDs as a key component of CBT (Beilharz et al., 2021; Fitzsimmons-Craft et al., 2022; Shah et al., 2022). This refers to common concerns about diets, comparison and self-perception, and beauty ideals. According to Kasson et al. (2021), 95% of teens own a smartphone, making them highly exposed to these impressions via social media. Based on findings from seven articles, it was observed that access to the chatbot was mainly provided through mobile applications. Thus, perhaps the content of CBT for EDs can become just as accessible as other content on social media platforms. In this manner, there is a connection between the themes of treatment and accessibility.

Further on, accessibility also referred to who the chatbot is relevant for. In five studies the aim was to provide a supportive chatbot for individuals at risk for EDs while clinically ill individuals were excluded. In relation to usability, concerns have been raised about rule-based chatbots potentially prompting harmful actions. Nevertheless, since the chatbots were

designed for individuals with mild symptoms of eating problems, the risk of inflicting harm is relatively low.

Personalization was stated by seven articles as a wanted feature of chatbots, in regards of giving individual feedback to the user, and having the responses delivered in a familiar tone. In relation to usability, rule-based chatbots cannot provide personalized feedback based on input from the user but generate responses according to keywords. Due to this, evaluation of symptom level when accessing the chatbot is of importance as it may not offer adequate guidance to individuals with severe ED illness. On the other hand, personalization can also be enhanced by incorporating elements of texting with a friend, such as using short responses, emojis, and warm language. This is a way to increase perceived satisfaction with the chatbot, as one study showed that human connection was preferred (Rohrbach et al., 2022).

Satisfaction was reported in three studies, with two reporting positive feedback, whilst in Rohrbach et al. (2022), participants were more satisfied with the expert-patient support, than the condition with chatbot only. However, there was no significant difference in effectiveness when comparing the two conditions. This suggests that intervention does not need to be satisfactory to be effective (Rohrbach et al., 2022). On the other hand, the subjective experience may be important for motivation to further usage. As stated earlier, participants reported frustration in limitations of accuracy, also being able to interact with the chatbot in a natural way appear to be a wanted attribute. Yet, none of the retrieved studies offers AI that can comprehend user input.

Limitations

A scoping review is supposed to map a body of literature on a broader topic, but this field of area is new, and therefore the available evidence is limited so far. This was a limitation which led to change in the eligibility criteria, thus including articles that referred to the development of chatbots on EDs and non-established chatbots.

Another limitation that differs from the standard procedure of a scoping review is that they are often written by multiple authors but given by the thesis terms, this thesis is written by me individually.

The conducted literature search of this thesis is the first attempt of myself and cooperating peers, at doing a comprehensive mapping of available literature. Thus, the results might have been different with some more practice. Guidance from our supervisor and the university library, however, gave us certainty as to how we should embark on the official literature search.

This scoping review intended to assess the user experience of both genders. However, several studies had a limitation in that they only included women, making the results not representative to both genders. Another weakness in population is that one of the studies describes that participants are young adults above the age of 18 years old, but no upper age limit has been stated.

When carrying out a scoping review, recommendations from JBI suggests to conduct an inter-coder reliability testing (Peters et al., 2017). However, due to the time restrictions of this thesis, only 20% of the text was coded by a peer.

Recommendations

Currently, only two of the analysed chatbots are available to users. The remaining studies developed, pre-tested and co-developed with end users based on their feedback. However, as AI technology continues to advance, there is an opportunity to create chatbots that can react to user input, improving usability and satisfaction. Therefore, it is important to evaluate existing literature and gather feedback from users to guide the development of chatbots with extended capabilities.

All retrieved studies origins from developed countries, such as USA and Australia. Hence, there is a lack of literature providing information about the geographical availability

of chatbot interventions for ED support. To address this, the study by Matheson et al. (2021) aims to provide a chatbot for traditionally underserved populations. This chatbot for Brazilian adolescents which is currently under way, may provide insights into the benefits of digital interventions on a global level, given that EDs is a global problem. Therefore, it is important to map the availability of digital support for young adults struggling with ED across cultures to increase the usability and effectiveness.

Conclusions

In this thesis, eight articles were identified for the scoping review of user experience with chatbots targeted at eating disorders. The chatbots were rule-based and conveyed through written language. The most common use of chatbots was promotion of psychoeducation and help-seeking among young adults and adolescents, as this is an age group at risk for ED development. Content of responses were developed to match the current texting culture and was mainly accessed through mobile applications. The review mapped several aspects of user experience, showing satisfaction in some aspects of effectiveness and usability, influenced by treatment approach and personalization. However, the role of chatbots is currently not suited for treatment, as they are limited in reacting to user input. Nevertheless, the literature shows that chatbots can be beneficial in promoting help-seeking and providing scientific knowledge, offering a supportive tool for individuals with eating problems and eating disorders. Overall, this scoping review provides insights into the potential of chatbots in filling the current treatment gap, addressing the problem of eating disorder among adolescents and young adults.

References

Ali, K., Farrer, L., Fassnacht, D. B., Gulliver, A., Bauer, S., & Griffiths, K. M. (2017). Perceived barriers and facilitators towards help-seeking for eating disorders: A systematic review. *International Journal of Eating Disorders*, *50*(1), 9-21. https://doi.org/https://doi.org/10.1002/eat.22598

Beilharz, F., Sukunesan, S., Rossell, S. L., Kulkarni, J., & Sharp, G. (2021). Development of a Positive Body Image Chatbot (KIT) With Young People and Parents/Carers: Qualitative Focus Group Study. *J Med Internet Res*, *23*(6), e27807. https://doi.org/10.2196/27807

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Braun, V., & Clarke, V. (2014). What can "thematic analysis" offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-being*, *9*(1), 26152. https://doi.org/10.3402/qhw.v9.26152

Chan, W. W., Fitzsimmons-Craft, E. E., Smith, A. C., Firebaugh, M.-L., Fowler, L. A., DePietro, B., Topooco, N., Wilfley, D. E., Taylor, C. B., & Jacobson, N. C. (2022). The challenges in designing a prevention chatbot for eating disorders: observational study. *JMIR Formative Research*, *6*(1), e28003.

Fitzsimmons-Craft, E. E., Chan, W. W., Smith, A. C., Firebaugh, M.-L., Fowler, L. A., Topooco, N., DePietro, B., Wilfley, D. E., Taylor, C. B., & Jacobson, N. C. (2022). Effectiveness of a chatbot for eating disorders prevention: A randomized clinical trial. *International Journal of Eating Disorders*, *55*(3), 343-353. https://doi.org/https://doi.org/10.1002/eat.23662

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000-2018 period: a systematic literature review. *American Journal of Clinical Nutrition*, *109*(5), 1402-1413. https://doi.org/10.1093/ajcn/nqy342

Harrer, M., Adam, S. H., Messner, E. M., Baumeister, H., Cuijpers, P., Bruffaerts, R., Auerbach, R. P., Kessler, R. C., Jacobi, C., Taylor, C. B., & Ebert, D. D. (2020). Prevention of eating disorders at universities: A systematic review and meta-analysis. *Int J Eat Disord*, *53*(6), 813-833. https://doi.org/10.1002/eat.23224

Hutton, B., Catala-Lopez, F., & Moher, D. (2016). The PRISMA statement extension for systematic reviews incorporating network meta-analysis: PRISMA-NMA. *Medicina Clínica (English Edition)*, *147*(6), 262-266.

Jarman, H. K., McLean, S. A., Rodgers, R., Fuller-Tyszkiewicz, M., Paxton, S., O'Gorman, B., Harris, E., Shatte, A., Bishop, K., Baumann, T., Mahoney, D., Daugelat, M.-C., & Yager, Z. (2022). Informing mHealth and Web-Based Eating Disorder Interventions: Combining Lived Experience Perspectives With Design Thinking Approaches [Original Paper]. *JMIR Form Res*, *6*(10), e38387. https://doi.org/10.2196/38387

Kasson, E., Vázquez, M. M., Doroshenko, C., Fitzsimmons-Craft, E. E., Wilfley, D. E., Taylor, C. B., & Cavazos-Rehg, P. A. (2021). Exploring social media recruitment strategies and preliminary acceptability of an mHealth tool for teens with eating disorders. *International journal of environmental research and public health*, *18*(15), 7979.

Lund, A. M. (2001). Measuring usability with the use questionnaire12. *Usability interface*, *8*(2), 3-6.

Matheson, E., Smith, H., Amaral, A., Meireles, J., Almeida, M., Mora, G., Leon, C., Gertner, G., Ferrario, N., & Suarez Battan, L. (2021). Improving body image at scale among

Brazilian adolescents: study protocol for the co-creation and randomised trial evaluation of a chatbot intervention. *BMC public health*, *21*(1), 1-14.

Nemesure, M. D., Park, C., Morris, R. R., Chan, W. W., Fitzsimmons-Craft, E. E., Rackoff, G. N., Fowler, L. A., Taylor, C. B., & Jacobson, N. C. (2023). Evaluating change in body image concerns following a single session digital intervention. *Body Image*, *44*, 64-68.

Peters, M., Godfrey, C., Khalil, H., McInerney, P., Soares, C., & Parker, D. (2017). 2017 Guidance for the Conduct of JBI Scoping Reviews. In.

Rohrbach, P. J., Dingemans, A. E., Spinhoven, P., Van Ginkel, J. R., Fokkema, M., Wilderjans, T. F., Bauer, S., & Van Furth, E. F. (2022). Effectiveness of an online self-help program, expert-patient support, and their combination for eating disorders: Results from a randomized controlled trial. *International Journal of Eating Disorders*, *55*(10), 1361-1373. https://doi.org/https://doi.org/10.1002/eat.23785

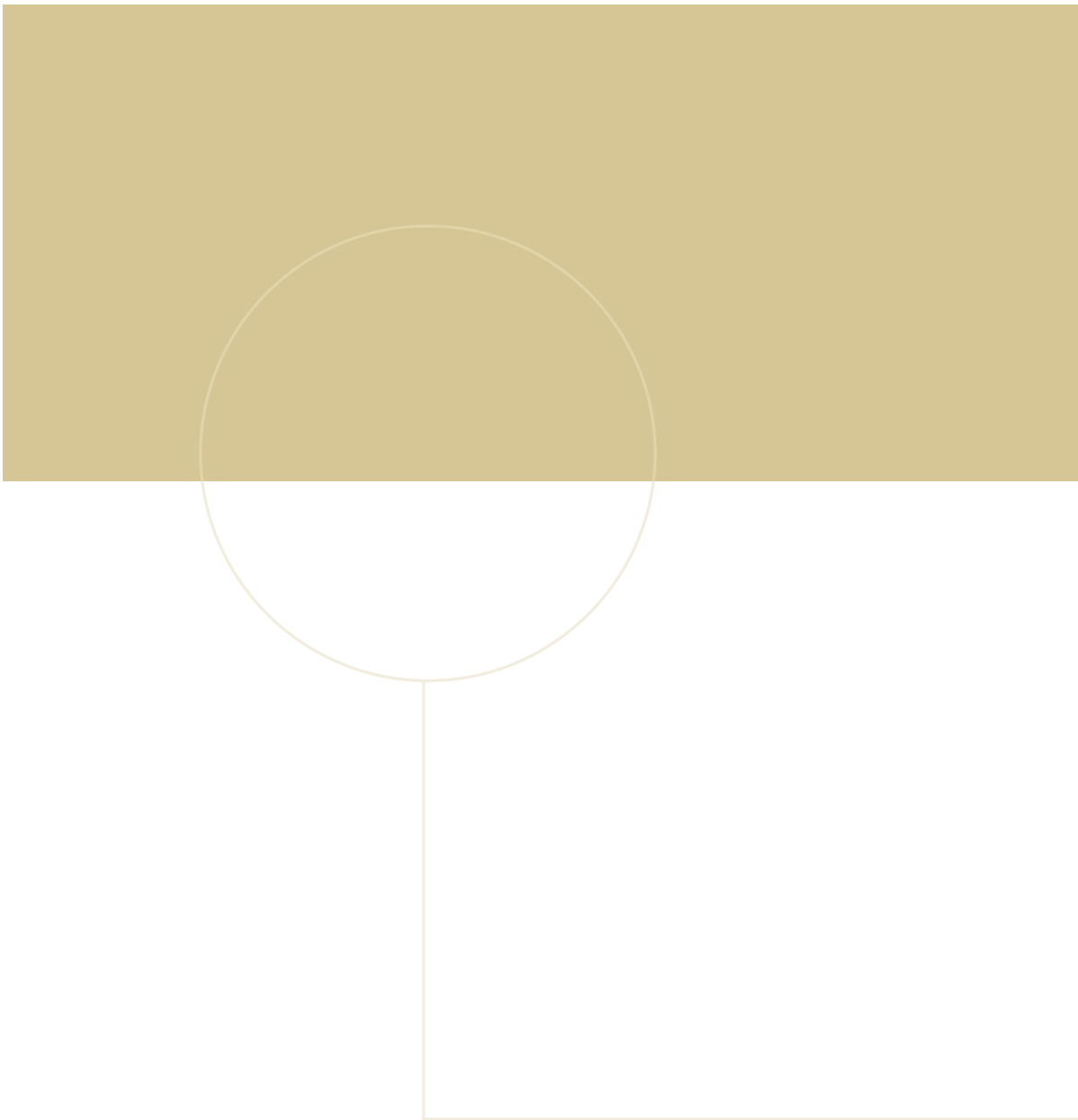
Rojewska, K., Maćkowska, S., Maćkowski, M., Różańska, A., Barańska, K., Dzieciątko, M., & Spinczyk, D. (2022). Natural Language Processing and Machine Learning Supporting the Work of a Psychologist and Its Evaluation on the Example of Support for Psychological Diagnosis of Anorexia. *Applied Sciences*, *12*(9), 4702. https://www.mdpi.com/2076-3417/12/9/4702

Shah, J., DePietro, B., D'Adamo, L., Firebaugh, M.-L., Laing, O., Fowler, L. A., Smolar, L., Sadeh-Sharvit, S., Taylor, C. B., Wilfley, D. E., & Fitzsimmons-Craft, E. E. (2022). Development and usability testing of a chatbot to promote mental health services use among individuals with eating disorders following screening. *International Journal of Eating Disorders*, *55*(9), 1229-1244. https://doi.org/https://doi.org/10.1002/eat.23798

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., & Weeks, L. (2018). PRISMA extension for scoping

reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine*, *169*(7), 467-473.

World Health Organization. (2019). ICD-11: International classification of diseases 11th revision. https://icd.who.int/browse11/l-m/en#/http%253a%252f%252fid.who.int%252ficd%252fentity%252f1412387537



NTNU

Norwegian University of
Science and Technology