EMPOWERING INFORMED DECISION-MAKING IN MENTAL HEALTH CARE: A WEB-BASED RECOMMENDATION SYSTEM FOR MOBILE APP SELECTION

MD Romael Haque
Marquette University

EMPOWERING INFORMED DECISION-MAKING IN MENTAL HEALTH CARE: A WEB-BASED RECOMMENDATION SYSTEM FOR MOBILE APP SELECTION

by

MD Romael Haque, B.Sc., M.Sc.

A Dissertation submitted to the faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Milwaukee, Wisconsin

August 2024

ABSTRACT
EMPOWERING INFORMED DECISION-MAKING IN MENTAL HEALTH CARE:
A WEB-BASED RECOMMENDATION SYSTEM FOR MOBILE APP SELECTION

MD Romael Haque, B.Sc., M.Sc.

Marquette University, 2024

In 2022, 23.1% of adults in the United States (77 million individuals) were affected by mental health (MH) concerns. Due to the inaccessibility and high cost of traditional treatment, around 55% of people with severe mental illnesses do not receive treatment. Mental health concerns are prevalent, affecting a significant portion of the population in the United States. Traditional treatment options are often inaccessible and expensive, leaving many people without essential mental healthcare.

However, the rise of mobile technologies has given rise to a promising solution: mobile mental health applications (MMHAs). These apps offer greater accessibility and affordability, potentially expanding mental healthcare services to a wider range of users. Despite their potential, there are challenges associated with MMHAs. Many lack robust evidence to validate their effectiveness, making it difficult for users and clinicians to choose the most suitable app. Additionally, existing frameworks designed to guide app selection, such as MARS, ORCHA, PsyberGuide, and MIND Tools, may be too complex for regular users. These frameworks are often developed by mental health professionals and may overlook the features that are most important to users. There is also a disconnect between user and professional ratings, with users valuing features that professional reviewers tend to overlook.

To address these challenges, we developed a MMHA recommender system that incorporates both user and expert perspectives. This system was informed by several studies, including analyses of user reviews from 164 MMHAs and 10 chatbot-based MMHAs providing valuable insights into user perspectives regarding the acceptance of these innovative solutions and the specific features they prioritize, in-depth studies with US veterans, a group particularly susceptible to mental health issues, capturing their personal narratives detailing both successful and unsuccessful experiences with mental health apps, and a review of two existing professional frameworks and eleven relevant research articles.

The developed app recommender system provides a holistic view of MMHAs, empowering users and healthcare providers to make informed decisions. By understanding both user preferences and expert evaluations, users can choose the app that best suits their needs. The system also helps to identify crucial criteria that are often missing from existing MMHA recommender systems. Overall, this research has the potential to improve user experience and promote effective mental healthcare through MMHAs.

I dedicate this book to my late father, MD Mozammel Haque, my wonderful mother, Arefa Nasrin, and my loving wife, Fayika Fahat Nova.

ACKNOWLEDGEMENTS

MD Romael Haque, B.Sc., M.Sc.

While it's impossible to thank everyone who supported me on this incredible journey, I must first express my deepest gratitude to my amazing supervisor, Dr. Sabirat Rubya. I was extremely fortunate to have such a wonderful and inspiring mentor throughout my time as a Ph.D. student. Her outstanding mentorship was invaluable; she guided and inspired me every step of the way, and I consider myself very lucky to have been under her supervision. When I joined her lab, I was facing significant challenges in my research and feeling quite discouraged. She uplifted me and supported me until the very end, helping me overcome my existential crisis. Under her three years of supervision, I accomplished more than I ever had as a researcher.

I would also like to take this opportunity to thank my dissertation committee members: Dr. Zeno Franco, Dr. Michael Zimmer, and Dr. Sheikh Iqbal Ahamed. Dr. Zeno Franco played a crucial role in this study by facilitating connections with U.S. Military Veterans. He acted as a bridge between the veterans and researchers, guiding me not only in conducting my study but also offering suggestions and advice that helped me mature as an ethnographer. He helped me comprehend the critical nature of our work and the importance of approaching it with the utmost respect. Dr. Sheikh Iqbal Ahamed was a constant source of motivation and guidance, inspiring me throughout my journey. Dr. Zimmer's inspiring work and suggestions greatly contributed to diversifying my thought process and encouraging more out-of-the-box thinking in my research.

Additionally, I would like to thank Dr. Shaikh Anowarul Fattah and Dr. Celia Shahnaz for inspiring me to pursue higher studies. Their encouragement and belief in my potential provided the foundation for my academic journey. I am also deeply grateful to Dr. Ishtiaque Ahmed, who introduced me to Human-Computer Interaction (HCI) and taught me everything from the ground up. His guidance was instrumental in shaping my research focus and methodological approach. My first supervisor, Dr. Shion Guha, helped me achieve a multitude of personal and professional qualities, shaping me into the individual I am today. His mentorship during the early stages of my academic career set the stage for my future success and instilled in me a strong work ethic and a passion for research.

I would like to express my deepest gratitude to Dryhootch of America for their substantial support throughout this study. They generously provided us with office space and dedicated their valuable time and effort. I also want to acknowledge the significant contributions of Mr. Robert Curry, Mr. Otis Winstead, Mr. Levi Marker, and many other patrons of Dryhootch of America. Their assistance was instrumental in helping us understand veteran culture and supporting our integration into the community. Additionally, I am very grateful to the iPeer research group. Dr. Natalie D. Baker, Dr. Praveen Madiraju, Dr. Niharika Jain, Mark Flower, Dr. Md Fitrat Hossain, Dr. Priyanka Annapureddy, and others in the iPeer group provided immense support and were instrumental in the completion of this research. I would be remiss if I did not thank my labmates from the Social and Ethical Computing Lab: Farhat Tasnim Progga, Nidhi Nellore, Dr. Devansh Saxena, and Dr. Fayika Farhat Nova. I also extend my gratitude to the student research collaborators: Kishore Kilaru, Isis Allen, Wylie Frydrychowicz, Adib Anik, Olivia, Temitayo R. Okusanya, Lucy Le, and Nadiyah Johnson.

I also wish to extend my deep gratitude to my incredible mother, Arefa Nasrin Jharna. From my early school years to college, she has been my steadfast companion. It was her unwavering encouragement that inspired me to take the bold step of moving to a foreign country to pursue my research dreams. Equally, I am thankful to my sister, Sharmin Haque, and her family—Shaheen Parvez, Sharjeel, and Areeb. Their steadfast support and encouragement have been essential pillars of strength throughout this challenging journey. Their belief in me has been a constant source of motivation, driving me forward in pursuit of my goals.Additionally, I would like to acknowledge

the immense support I received from my in-laws - Mamoni, and Baba. Their constant motivation and encouragement were invaluable throughout this journey. Speaking of support, I must mention my extended family in Milwaukee: Ananta, Rumpa, Runa, Kowser, Tazim, Progga, and Tanzil. In a foreign land, they embraced me more than friends, making my life much easier despite being away from my immediate family. I will always cherish those late-night hangouts, which provided a constant relief from the stress I faced during my Ph.D. I would also like to extend my gratitude to my dear friend Minhaz, who traveled eight hours from Cleveland just to cheer me up.

I'm incredibly grateful for my wife Fayika Farhat Nova and how she's stood by me every step of the way. She's been there through thick and thin, showing amazing patience and strength. Whether cheering me on in the good times or helping me through tough spots, she's been my rock. I feel so lucky to have had her support through this challenging journey. As a senior researcher in my lab, she has also played a pivotal role in my academic growth, and I am sincerely appreciative of the opportunities to learn from her. Her support, guidance, and inspiration have been invaluable, and I am deeply thankful for the positive impact she have had on my journey.

Lastly, I want to honor my late father, MD Mozammel Haque, who passed away at a relatively young age during the COVID-19 pandemic in 2020. Throughout my life, he was the role model I aspired to emulate. He is the reason I pursued engineering and later, a Ph.D. His unwavering and silent support was a constant throughout my journey. Regardless of how unreasonable my requests might have been, he always did his best to provide for me. His patience and kindness towards everyone profoundly influenced who I am today. It deeply saddens me that he could not witness this milestone, as he eagerly anticipated it and would have been overjoyed by the news. I dedicate this achievement to his memory and pray for his soul's eternal peace.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In 2022, 23.1% of adults in the United States (77 million individuals) were affected by mental health (MH) concerns., representing 1 in every 5 adults. Due to the inaccessibility and high cost of traditional treatment, around 55% of people with severe mental illnesses do not receive treatment. With the advancement of mobile technologies in the last ten years developers recognized a strong promise for digital tools like mobile phone applications to expand better accessibility at low cost to mental health (MH) treatment and services. It's a breakthrough for making MH treatment more accessible, convenient, and adaptable to the patient's lifestyle. Mobile Mental Health Applications (MMHAs), also known as mental health apps, are software applications designed for mobile devices like smartphones and tablets. They aim to support mental well-being, provide mental health information, and offer tools or interventions to manage and improve mental health. These apps cover a wide range of mental health aspects, including stress management, mindfulness, meditation, mood tracking, and access to mental health resources. They can be utilized for self-help, as supplements to therapy, or alongside traditional mental health services. MMHAs are cost-effective, accessible, and scalable tools that have the potential to expand access to mental healthcare worldwide.

The extensive adoption of mobile health (mHealth) technologies is unprecedented in the history of healthcare and access to treatment. (117). Many applications have been developed in recent years to address health-related concerns, particularly mental health (MH) issues (117). These apps target a broad range of psychological disorders including anxiety (250), depression (242), schizophrenia (237) etc., and vary in design and functionality. They a variety of techniques, including tracking and monitoring to gain a better understanding of the causes and symptoms of mental illness (85), providing techniques to mitigate and manage symptoms that arise as a result of a mental health issue, such as breathing (57), meditation (104), etc., mental health games–based apps that help to manage anxiety and achieve healthy living (44) and providing a community of peers and/or coaches where users can ask and answer questions and exchange support (91). Despite the fact that mental health apps have been shown to be beneficial in supporting mental health, there is still a significant number of issues and concerns, such as early dropouts (182), low engagement (196), not being very effective in improving the condition of the patients (43; 271), etc. For example, researchers found clinically proven methods such as CBT and positive psychology often do not perform as expected in mobile-based platforms (197). Moreover, prior research showed different self-guided mobile applications suffer from early dropouts (22). Many mental health applications report low

engagement and a high attrition rate. However, there is little evidence to support these claims (261). Possible reasons include usability issues, such as software bugs, unappealing user interfaces, data loss, battery and memory usage issues, lack of guidance and explanation, etc. (15). Additionally, due to privacy concerns, people are often wary of using mental health apps (252). Very few evidence and research based applications that are available often fail in implementing important features, such as understanding early warning signs and triggers (184),wellness plans (262), etc. These result in a lack of trust among users and healthcare providers (261), as well as poor engagement and dropouts (196; 182).

Users seeking MMHAs need clear and comprehensive information to make informed choices. Moreover, the rapid proliferation of MMHAs has highlighted the need for reliable frameworks to evaluate their safety and effectiveness (162). Prior literature acknowledges the needs and provides implications around how experts, policymakers, and researchers can provide guidance in this area (166). Following up these concerns, different organizations have developed such frameworks, each with its own focus and methodology. Numerous Assessment and Evaluation Tools such as frameworks, guidelines, rating systems, or app libraries, have been established worldwide (48). Nevertheless, these efforts encounter challenges. For instance, the NHS Apps Library (128), which evaluated apps based on specific criteria, was introduced but promptly withdrawn due to public backlash after reports surfaced regarding privacy and security vulnerabilities in several of the apps (128). Mobile App Rating Scale (MARS) (251) and Organization for the Review of Care and Health Applications (ORCHA) developed their own guideline tools focusing on aspects such as engagement and usability (170). MARS evaluates apps based on criteria including engagement, functionality, aesthetics, information accuracy, and subjective quality (251). Another prominent guidance framework is the American Psychiatric Association's APA App Evaluation Model, which prioritizes factors such as privacy, clinical foundation, and therapeutic goal. This model underwent refinement by a diverse committee, aiming for objectivity and evidence-based assessment (17). However, evaluation frameworks such as MARS and the APA App Evaluation Model aim to guide users through app selection questions, but may lack clear metrics for decision-making (161; 194). They also lack the sufficient detail for users to determine an app's suitability and do not provide clear metrics to guide app choices (194).

To facilitate app evaluation, organizations have created databases or evaluation hubs, such as the MIND database (65), which streamline the process by categorizing apps based on evaluation criteria. It has been developed in collaboration with the American Psychiatric Association (APA) (**?**),

focusing on five priority levels: background info, privacy and safety, evidence, ease of use, and data integration (161). This framework consists of 105 objective questions, aiming to provide a data-driven approach to evaluating health apps and enabling customization to meet individual needs. Crowdsourcing is also incorporated to ensure ongoing evaluation and identification of unsafe apps (161). While the framework may be comprehensive, efforts have been made to make it practical for clinicians and users, with a focus on usability and objectivity. Another such effort led to the creation of One Mind PsyberGuide (One Mind PsyberGuide), a nonprofit project established in 2013, aims to assist individuals in using technology for better mental health. Its evaluations focus on three key metrics: credibility, user experience, and transparency regarding data security and privacy practices (194). Instead of assigning an overall score, PsyberGuide presents these metrics separately to allow consumers to prioritize factors based on their individual needs and preferences (195).

However, despite recent efforts by a few evidence-based works such as PsyberGuide and MIND to streamline these guidelines or provide additional materials to support their use, these guidelines require careful consideration and evaluation of apps for security, credibility, and clinical efficacy. Consequently, they may prove even more challenging for general users who seek simpler information to make choices. Carlo et al. discovered that the most common criteria of these frameworks, such as Psyberguide, Mindtools, and ORCHA, often have notably discordant reviews. Moreover, these frameworks and guidelines are developed by mental health professionals and experts who possess clinical knowledge but may lack insight into what day-to-day users truly desire (48). Prior literature has shown significant differences in opinions between experts and users on several criteria. Hudson et al. found low agreement between users' and professional ratings, and they identified several features missing from professional ratings, such as customization and integration with daily life. They also highlighted that participants particularly valued certain aspects of mental health apps that appear to be overlooked by professional reviewers (129). This suggests that academic researchers and clinicians designing frameworks and evaluating apps may be missing what consumers are actually seeking. With the rapid growth of technology and advancements in different AI tools such as chatbots, these technologies have made their way into intervention techniques supported by MMHAs. Prior literature has pointed out the limitations of these new tools through user perspectives (112). Additionally, Ramos et al. demonstrated that it remains unclear whether existing app assessment tools adequately address considerations of diversity, equity, and inclusion (DEI), which are particularly crucial for reducing mental health disparities among marginalized communities (224). This underscores the need for a universal framework that acknowledges both user and expert opinions.

Frameworks developed by mental health professionals undoubtedly carry significant value and knowledge that should be considered when choosing a MMHA. However, incorporating user perspectives in accordance with that has the potential to reveal a complete picture that will help both clinicians and users make informed decisions in choosing a MMHA.

Prior research has attempted to understand these challenges using data from a variety of sources, including the descriptions of applications in mobile app stores, publicly available peer-review databases, frameworks and evaluation criteria from engineering and informatics literature, and interdisciplinary organizations, etc. But only a few studies focused on user-centered design and usability evaluation of these apps. Therefore, there is a gap in HCI literature in understanding the limitations of these apps through a user-centered lens. To gain an empirical understanding of why there is such a low level of engagement and significant dropout rates, it is vital to examine the challenges and concerns regarding these apps from the consumers' perspective and their real-life experience. A comprehensive understanding of user perspectives and experiences may aid in answering research questions around this debate. This project may contribute to an understanding of the current confusing, uncertain picture about consumer interest in using mental health apps in the real world (as opposed to research settings).

My research aims to fill a research gap in this area by providing concrete research and design implications by emphasizing the importance of user-centric design and evaluation of systems targeted for vulnerable populations. I believe in research involving multiple stakeholders both direct and indirect stakeholders (who aren't directly involved with the system but are influenced by it). Without them, the development will compromise key values that are vital to the design process. In my thesis, my aim is to involve stakeholders such as general users of mental health apps, psychiatrists and experts in psychological science, one of the most vulnerable population such as US veterans and lastly, app developers.

To that end, First, We focused on thoroughly examining user perspectives by analyzing user reviews from 164 mobile mental health applications, sourced from platforms the Google Play Store and Apple App Store. This study explores mental health (MH) apps through a thematic analysis of user reviews, revealing user sentiments. Users share their experiences and feedback, providing valuable data with positive, negative, and neutral tones. Reviews offer detailed evaluations of app features and suggestions for improvements. MH apps were identified on Google Play (Android) and Apple App Store (iOS) using the search term "Mental health apps," resulting in 117 Google Play and 76 Apple App Store apps, with 29 apps compatible with both platforms. Python scripts were used to

gather 1620 Google Play and 2266 Apple App Store reviews from 117 and 76 apps, respectively. Reviews from January 1, 2019, to May 1, 2021, with 200 or more characters, and one to three-star ratings were included. Thematic analysis was used to identify patterns in user feedback, resulting in nuanced insights into app experiences. The analysis involved two passes: the first pass generated numerous open codes to capture diverse perspectives, later reduced through memoing and clustering. In the second phase, codes were clustered using affinity mapping to identify emerging themes.

Furthermore, to explore the intricacies of the latest technological advancements in MMHAs, we have specifically investigated the incorporation of chatbots as a recent feature. This involved analyzing user reviews of 10 chatbot-based MMHAs. To compile a comprehensive list of commercially available mental health (MH) apps with chatbot features, multiple sources were explored. This included open-access articles on MH chatbots, expert mental health app review platforms (mindtools.io and psyberguide), and searches on Google Play and Apple App Store using "Mental health" and "chatbot" as keywords. The search, done without logging into specific accounts, aimed to mimic user exposure to these apps. After careful scrutiny of app descriptions, screenshots, and promotional websites, 19 apps were initially identified. To ensure they met the definition of MH chatbots, one author downloaded and used each app for at least three days. This process also helped extract common features of these apps. Ten apps were selected for analysis. Scraping scripts were created to collect public user reviews for these ten apps from Google Play and Apple App Store. The inclusion criteria for reviews were a timeline from January 1, 2019, to May 1, 2022, and a minimum length of 200 characters. This filtering yielded 3621 Google Play reviews and 2624 Apple App Store reviews based on 9 and 10 apps, respectively. These reviews were anonymized and coded for analysis. In understanding the apps' descriptive overview, app descriptions from promotional websites and marketplaces were analyzed, and six main themes were identified collaboratively. For user perspectives, an inductive thematic analysis of selected user reviews was conducted. Thematic analysis was chosen for its systematic approach to large data sets and contextual understanding.

To capture personal narratives detailing both successful and unsuccessful experiences with mental health apps, we conducted a study involving US veterans, a group particularly susceptible to mental health issues. The study was conducted in two phases. Initially, our aim was to understand perceptions of mental health Crisis among U.S. military veteran peer mentors and the potential of mobile-based peer-support. U.S. military veterans face an elevated risk of suicide or exhibiting suicidal behaviors. Peer-support interventions have proven to be effective for veterans because of their shared experience and community. However, as veterans might face a mental health crisis at

any time, it is crucial that the mentors can identify their peers' crisis warning signs early enough. Mobile technology has the potential to facilitate and improve peer-to-peer communication. Gaining an understanding of how the veteran community perceives crisis symptoms, as well as their values and technological needs, bears utmost importance in creating any tool or adopting any strategy. Hence, we conducted a mixed-methods study with twelve peer mentor military veterans. Our research will offer an in-depth understanding of the nuanced conceptions of mentor veterans about early warning signals and acute mental health crisis symptoms, as well as aspects of technology that may aid this community in recognizing and managing these symptoms. In the second stage of the study, we explored design opportunities for persuasive reminders for veterans to facilitate peer-mentor support. US military veterans (USMVs) are a vulnerable population with an elevated risk of mental health issues and suicide. Peer support, especially through mobile technology, has proven effective in addressing mental health related challenges, but ensuring long-term engagement remains a concern. This study explored the opportunity of designing persuasive technology, particularly persuasive reminders, to enhance engagement in peer support interventions for veterans. We followed community-based participatory research with ten veterans to identify specific peer support processes that can benefit from persuasive reminders and to uncover the underlying community values and needs to guide design. Our study highlighted the deficiencies of existing apps in terms of visual appeal, privacy features, and the ability to provide support during critical moments, shaping the direction for delivering digital interventions that adapt to users' daily routines.

For our final project, all these insights into user needs and preferences informed the design of an interactive tool aimed at assisting users in making informed decisions when selecting a mobile mental health application. We developed a web interface where user reviews and expert opinions will be categorized based on the preferences identified in previous studies. The design is resulted from combining our knowledge from our studies and 11 peer-reviewed articles that focus on user requirements from MMHAs, systematically analyzed from two most popular professional app evaluation frameworks – Psyberguide and Mindtools, we derived a universal framework for app recommendation that acknowledges both user and expert opinions. We implemented this framework through an interactive web-based interface aimed at assisting users in making informed decisions. By providing access to both user and expert perspectives on these apps, users and healthcare providers will be empowered to make consistent decisions about which mobile mental health applications to choose. Our research contributes to identifying and understanding the crucial criteria that are either absent or inadequately addressed in existing mobile mental health app (MMHA) recommender systems. We

provide concrete research and design implications based on our findings for designing an effective and user-centric MMHA recommender system.

The entire document is divided into seven chapters. In Chapters 2 through 5, I have outlined five in-depth studies conducted to achieve my research goals.They are -

- An examination of critical user reviews of mobile mental health applications (MMHAs) to understand users' opinions and perspectives
- Exploring values and requirements for developing mobile technology for U.S. military veterans to enhance participation in peer-mentor support within mobile mental health apps.
- Design a web-based recommendation system for mobile mental health app selection

Chapter 6 wraps up with a discussion of the contributions made by these projects and outlines potential future directions.

CHAPTER 2: AN ANALYSIS OF USER REVIEWS OF MOBILE MENTAL HEALTH APPLICATIONS

Mobile mental health applications are seen as a promising way to fulfill the growing need for mental health care. Although there are more than ten thousand mental health apps available on app marketplaces, such as Google Play and Apple App Store, many of them are not evidence-based, or have been minimally evaluated or regulated. The real-life experience and concerns of the app users are largely unknown. To address this knowledge gap, we analyzed 2159 user reviews from 117 Android apps and 2764 user reviews from 76 iOS apps. Our findings include the critiques around inconsistent moderation standards and lack of transparency. App-embedded social features and chatbots were criticized for providing little support during crises. We provide research and design implications for future mental health app developers, discuss the necessity of developing a comprehensive and centralized app development guideline, and the opportunities of incorporating existing AI technology in mental health chatbots. **This chapter has been published in the Proceedings of the ACM on Human-Computer Interaction, Volume 6, Issue CSCW2, Article No.: 421, Pages 1 - 29, (https://doi.org/10.1145/3555146).**

2.1 Introduction

One out of every four persons on the world has been impacted by mental or neurological disorders at some point in their lives (208). Mental disorders affect around 450 million people globally, making them one of the primary causes of ill-health and disability according to WHO (208). In 2019, 20.6% of adults in the United States (51.5 million individuals) were affected by mental illness, representing 1 in every 5 adults and a sharp increase of 4% in just three years (150). Due to the inaccessibility and high cost of traditional treatment, around 55% of people with severe mental illnesses do not receive treatment (221). With the advancement of mobile technologies in the last ten years developers recognized a strong promise for digital tools like mobile phone applications to expand better accessibility at low cost to mental health (MH) treatment and services (31). Prior study has acknowledged this breakthrough for making MH treatment more accessible, convenient, and adaptable to the patient's lifestyle (48). By 2018, there were over 10,000 MH and wellness apps available for immediate download (263), with services ranging from symptom tracking and monitoring to implementing scientifically grounded therapy, such as CBT and Mindfulness, as well as various interactive tools and modules for self-guidance (27). However, prior study has shown that while access and availability of MH applications has increased, concerns about privacy, effectiveness,

and usability have also been raised (262). According to a recent study, 7–42% of users continued to use the apps after four weeks, but just 0.5–28.6% after six weeks (215). The use and adherence with MH applications in real-world settings may range between 1% and 29% (88). Furthermore, because the majority of the applications accessible are not evidence-based, their efficacy in aiding people with MH concerns is questionable (27).

Prior research has attempted to understand these challenges using data from a variety of sources, including the descriptions of applications in mobile app stores (69), publicly available peer-review databases (27), frameworks and evaluation criteria from engineering and informatics literature, and interdisciplinary organizations (50), etc. Smartphone-based apps represent a unique opportunity to expand the availability and quality of MH treatment, and the COVID-19 pandemic caused a surge in MH and wellness app downloads (272). But only a few studies focused on user-centered design and usability evaluation of these apps. There is a gap in HCI and CSCW literature in understanding the limitations of these apps through a user-centered lens. To gain an empirical understanding of why there is such a low level of engagement and significant dropout rates, it is vital to examine the challenges and concerns regarding these apps from the consumers' perspective and their real-life experience.

2.2 Objective of the Study

User reviews provide a unique venue to extract information around usability issues and evaluate efficacy (198). Studies have demonstrated that analyzing user evaluations can provide in-depth knowledge of user experiences. It can assist in identifying potentially harmful functionalities, incorporating community values and expectations into app design in a variety of domains, including mobile health technologies (14). Furthermore, users consider these reviews as a credible source to learn about expectations and potential consequences, allowing them to make informed judgments about what to choose and avoid while thinking about these apps as treatment supplements. In this project, we utilize the app descriptions and user reviews of the mobile MH apps from two most popular app stores to investigate two research questions:

RQ1: What is the state of the art in mobile MH applications, in terms of targeted MH problems and techniques used to support users?

RQ2: What are the concerns and perspectives expressed in user reviews published on mobile app store platforms about the usability and efficiency of mobile MH applications?

2.3 Background

We situate our work in previous literature by discussing why and how researchers conducted user review analysis, introducing the current state of the art research on MH apps.

Online reviews are a type of product information generated by individuals based on their own personal experiences with the product. Consumers can post product reviews with content in the form of numerical star ratings and open-ended customer-authored remarks about the product on online sites (188). User reviews are an integral component of the user experience, especially when topic satisfaction is crucial for adoption (156). Like many other online marketplaces, mobile app stores allow users to post their feedback based on their experience of using the apps (258). User input in the form of app ratings and reviews is becoming common in app marketplaces. Users may utilize this feedback loop to evangelize and promote apps they like, as well as alert other users of the apps' shortcomings. Given the highly competitive nature of the mobile app market, the user reviews often work as a resource for customers to take informed decision about which apps will or won't work for them and if there are any specific factors that impact the amount of feedback a user will provide (198). In recent years, researchers have offered a wide range of analytics and data mining tools to help developers and analysts use this feedback (108). In terms of analytical tools and methodologies to extract the rich contextual information, prior research have used sentiment analysis (36), SVM (278), Naive Bayes (213), maximum entropy classification (278; 137), N-gram model (278), other unsupervised learning tools (51) and qualitative analysis methodologies (258) etc. These methodologies are being used to identify relationship between ranks and sales (52; 59), identify the contradictory phrases (84), explore the underlying factors of customer value (212), compare reviews of different products in one category to find the reputation of the target product (125), create summary representations of opinions to perform question answering (47), identify fake reviews (176), exploring user characteristics (90; 247) and perceptions (158; 96) etc. from different e-commerce platforms such as amazon (213; 36), mobile app stores (249; 15), hospitality sectors (78), restaurant sectors (160) etc. There are few works that employ user evaluations to investigate mobile mental health/well-being/behavior change applications for future design opportunities (66), figuring out usability concerns (14), and analyzing user characteristics (268).

Our research is motivated by this body of work that considers user-generated feedback as a valuable source of information (217) and make useful meaning out of them (198). User-generated content allows us to gather information from a wide spectrum of people that is often difficult to collect using

typical data collection methods (256; 217). Through this investigation we will develop a deeper understanding of the real-world experience and usability concerns of mental health app users and extract the complaints and suggestions including problem reports, feature requests, etc, which are valuable resources for app developers to improve user experience and satisfaction (212).

The importance of applying best practices such as making comprehensive support accessible and available inside the app, as well as following clinical evidence-based approaches has been acknowledged in previous research. However, research into the effectiveness of mental health apps is still in its infancy, and there are ongoing discussions about how best to examine the efficacy of mental health apps. Hundreds of apps are being developed and are downloaded from the app stores without significant research to confirm the effectiveness of such apps. A comprehensive understanding of user perspectives and experiences may aid in answering research questions around this debate and inform design of effective apps. Moreover, qualitative analysis of app store reviews will contribute to an understanding of the current confusing, uncertain picture about consumer interest in using mental health apps in the real world (as opposed to research settings).

2.4 Methods

This study examines mental health (MH) applications by using a thematic analysis of user reviews to look at how users feel about them. Users of mobile apps can download, review, and comment on the apps depending on their experience and happiness. This publicly accessible data (user reviews) offers in-depth evaluations that include positive, negative, and neutral input. Written evaluations can also reveal customer satisfaction with certain program features or express a desire for enhancements. As a result, reviews are regarded as a significant source of information.

2.4.1 Selection of Sample Apps and Reviews

We searched for potential apps focused on MH on two dominant mobile platform app stores (Google Play for Android and Apple App Store for iOS). We used the search term "Mental health apps" on both app stores. The search was conducted from the homepages of the app stores without logging in to any specific account. This step was taken to ensure that no ranking algorithm could be applied by the system to prioritize any specific user preference. Therefore, although the search results may not be fully comprehensive (as found from convenience sampling), these apps represent the sample in (almost) the same order that consumers would likely be exposed to and thus, most likely to use. We followed similar filtering considerations used in prior research on evaluating mobile MH apps (109) and thus included both free and paid apps with English-language interface and at least 100

downloads. These filters gave us a list of 249 apps from Google Play and 117 apps from Apple App store. The authors reviewed and discussed the app descriptions included in this list to exclude any app that are not relevant. The final list included a total of 117 applications from Google play store and 76 applications from apple app store. In total 29 applications were compatible with both platforms. In the "Supplementary Material" section, a complete list of applications (considered in our study) has been attached.

We wrote python scripts to scrape application details and all available user reviews from the 189 apps. In order to extract recent critical user feedback of the apps we applied the following three inclusion criteria to filter the reviews.

- Timeline: We included reviews that were published between January 1, 2019 and May 1, 2021. As the marketplaces change rapidly with new apps and updates of the existing apps, we wanted to capture the most recent user reviews. In addition, this timeline would allow us to understand if there exists any substantial difference in the reviews between pre and during COVID-19 period.
- Length: We included reviews with character counts of 200 or more. According to a previous study, the average amount of characters in a relevant review is 110.8. (176). We chose 200 characters for the scope of our study because reviews with more characters will aid us in uncovering deeper insights when performing the analysis. Moreover, shorter reviews often tend to be fake and are created for promotional purposes. Ignoring those reviews assisted us in improving data integrity.
- Star rating: This filter was used to limit the reviews to consider to only those that were critical and/or had negative insights. Therefore, we took into account the reviews that have a one, two, or three star rating.

1620 reviews from the Google Play Store and 2266 reviews from the Apple App Store met all of the inclusion criteria. These reviews are based on 117 apps from the Google Play Store and 76 from the Apple App Store. All of the reviews have a unique coding system that can be easily traced back to the application and platform from whence they emerged. During the analysis the lead author was responsible to carefully read each review and ensure that all personally identifying information were replaced or removed.

2.4.2 Data Analysis

We analyzed the reviews using an inductive thematic approach (200). Thematic analysis was chosen because it allows systematic study of a big data collection and aids in the understanding of patterns in the text while taking context into consideration (258; 147). Previous research (258) has shown that data obtained from thematic analyses may be examined and reported qualitatively, therefore we employed a qualitative thematic analysis. Analysis consisted of two passes. In the first pass we generated open codes (over 20 total) to capture diverse viewpoints from the reviews. This high number of open codes is due to our effort to capture the nuance in the specific insights included in each review and was significantly reduced through the process of memoing and clustering (200). In the second phase of analysis, we memoed and clustered the codes using a constant comparison method operationalized as affinity mapping. Each open code was compared to others and positioned to reflect its affinity to emerging themes and clusters. The reported themes consisted both of ones that appeared consistently across multiple reviews, and also the ones that came from reviews that represented divergent responses and opinions.

2.4.3 Data Integrity and Ethical Considerations

We considered the fact that app stores, like many other online marketplaces,can have reviews posted by fake and paid users. However, prior research (176) showed that in "Health & Fitness" category the percentage of potential fake reviews is very low (about 6%). Moreover, one primary characteristic of the fake reviews are that, they have good star ratings (176). Since we are considering reviews with three or fewer stars, we assume that almost all of the reviews included are original and can be trusted.

Our python script got blocked multiple times by apple app store because Apple has an automatic scraper detector that block requests coming from the same IP address within a short period of time (268; 120). We circumvent this by identifying those applications where the review collection was incomplete and running our scripts on those applications iteratively from different IP addresses. We also ran an additional python script to detect and eliminate all the duplicate reviews (if present).

Keeping ethical implications in mind, we only collected information on webpages that were already explicitly public for everyone on Google Play Store and Apple App Store and were not only accessible to only a certain community of users. As suggested by (283), it is unethical for researchers to use any personal information from the internet if the data or information is restricted to a certain group

of people or communities. Hence, we ensured the webpages we collected the data from are public and not restricted to certain communities or populations. (283) also adds to this conversation of ethical implications by emphasizing that more than expected information should not be revealed through the combination of visual and textual elements. To ensure that, we intentionally refrained ourselves from publishing or revealing any identifiable information that were shared in those pages even though they were public.

2.4.4 Positionality

The research team consisted of members whose family members and close relatives have experiences and suffered from MH disorders. One team member is an expert in intersectional stigma, online social support, and MHcare, while another is an expert in working with marginalized and vulnerable populations.

2.4.5 Limitations

Our analysis methods and selection criteria have some inherent limitations. First, we only considered negative reviews (¡=3 stars) in our analysis. These may be coming from users willing to write about their strong unfavorable sentiments, and we may be missing feedback from users who are not comfortable (or do not care) about writing their experience on online platforms. However, since we considered a representative sample of the MH apps and the number of reviews were high, we can safely assume that the concerns we found will be representative of user opinions. Second, we elected not to incorporate critical assessments around broad UI issues and software defects in our user-centered evaluations, largely because we aimed to capture a broader picture of customers' real-life experience and concerns around the usability and efficiency that a mere software update wouldn't be able to solve, and need more attention from the researchers and the app developing organizations. Third, for this study, we looked at ratings from only the two major mobile platforms (Google and Apple). Other mobile platforms were not taken into account for the scope of this research.

2.5 Findings

We provided our findings from the analysis in this part divided into two subsections. In the first subsection, a descriptive analysis of the current state of the overall mobile mental health (MH) app market is presented. In the second subsection, perspectives and sentiments raised by users such as inconsistent moderation, toxic social interaction, lack of transparency and useful functionalities regarding chatbots are presented.

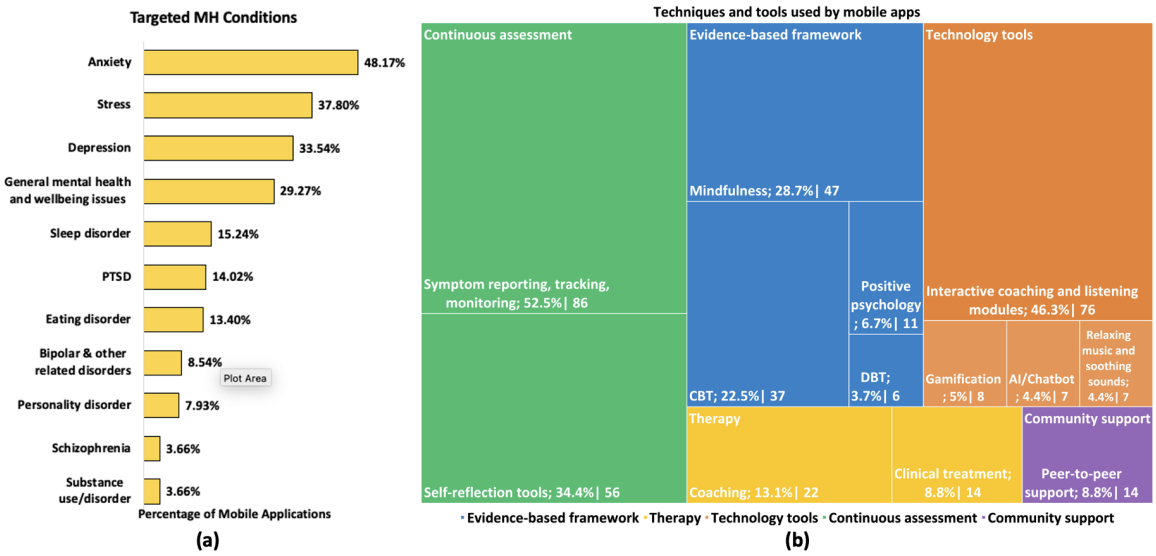


Figure 1: (a) Targeted MH conditions covered by the (percentage of) apps considered in our study. (b) Frequency distribution of the apps in terms of techniques and technological tools used.

2.5.1 State of the Art in Mobile Mental Health (MH) Applications

We begin with an overview of the apps that were selected from our careful consideration following the approach described in section 2.4.1. Although the list of apps may not be exhaustive, it is a large representative sample of the existing mobile MH apps from two most popular app stores. Thus, we can assume that the descriptive statistics can be generalized and the findings reflect the current status of the overall mobile MH marketplace. We provide categorizations of the apps on two key aspects: types of MH concerns they target and techniques and technologies used to assist users.

Targeted MH conditions 1(a) shows the number and percentage of apps per addressed MH problems, sorted in descending order. These categories are not mutually exclusive, which means that an app can address multiple MH conditions. Anxiety is the most commonly addressed issue (79, 48.17%). Other significant issues are : stress (62, 37.8%), depression (55, 33.54%), general mental health and wellbeing (48, 29.27%), sleep disorder (25, 15.24%), PTSD (23, 14.02%), eating disorder (22, 13.4%), bipolar & other related disorders (14, 8.54%), personality disorder (13, 7.93%), schizophrenia (6, 3.66%), substance use/disorder (6, 3.66%).

Intervention techniques and approaches used Different apps use different strategies and tools to treat various MH issues. These techniques, as found from the description of the applications covered in our study (n=164), can be split into five broad categories: evidence-based framework, therapy, technology tools, continuous assessment, and community and peer support. Figure 1(b) represents the number and percentage of these broad categories and subcategories. These categories

are mutually inclusive, implying that an app can implement a variety of strategies.

- **Evidence-based framework:** A handful of the apps implemented techniques that are evidence based and scientifically grounded such as CBT, DBT, mindfulness and positive psychology. Assisting people in achieving mindfulness is a prominent evidence-based strategy. We found 47 (28.7%) apps utilizing mindfulness. The ability to be fully present, aware of where we are and what we're doing, and not unduly reactive or overwhelmed by what's going on around us is known as mindfulness. Meditation is a popular method for achieving mindfulness. These apps offer guided meditations and courses, sleep meditation, breathing exercises, stress relief and coping meditations, exercises related to finding focus and building productivity, tailored reminders about meditation, and the option to measure the number of minutes meditated, etc. The duration of these courses and guides might range from a few minutes to several hours. CBT (cognitive behavioral therapy) is an another type of psycho-therapeutic treatment that teaches people how to recognize and change harmful or distressing thought patterns that affect their behavior and emotions. Among the applications we considered, 37(22.5%) implemented CBT. Majority of the apps have a number of therapeutic elements such as tests and scales to assess one's well-being, resources about symptoms and treatments, pleasurable activities to improve one's mood, tasks to complete to practice new coping skills, communicating through inspirational quotes, and suicide prevention measures (providing links to support services, ability to prepare a crisis plan), etc. Positive psychology is a modified type of CBT that is essentially a cognitive treatment for changing negative thoughts and dysfunctional perceptions into more positive ones. Some of these apps (11, 6.7%) apps allow users to reflect on themselves using diaries and journals, as well as add affirmations to encourage positive thinking. Finally, DBT is another evidence-based technique that these programs use. DBT (Dialectical Behavior Therapy) is a modified version of CBT. Its primary objectives are to teach people how to live in the now, build healthy stress coping mechanisms, regulate their emotions, and improve their interpersonal connections. Only 6 (3.7%) applications used this technique.

- **Therapy:** This wide category was broken into two subgroups: coaching (22,13.1%), where users can connect with professional coaches and active listeners guide them through various mental and physical wellness recovery processes and clinical treatment (14,8.8%), in which users can remotely connect with a medical professional or a registered therapist and go through their therapy session with them. In few of these apps, users can be matched with coaches and therapists who suit their style and preferences using algorithms and matching systems

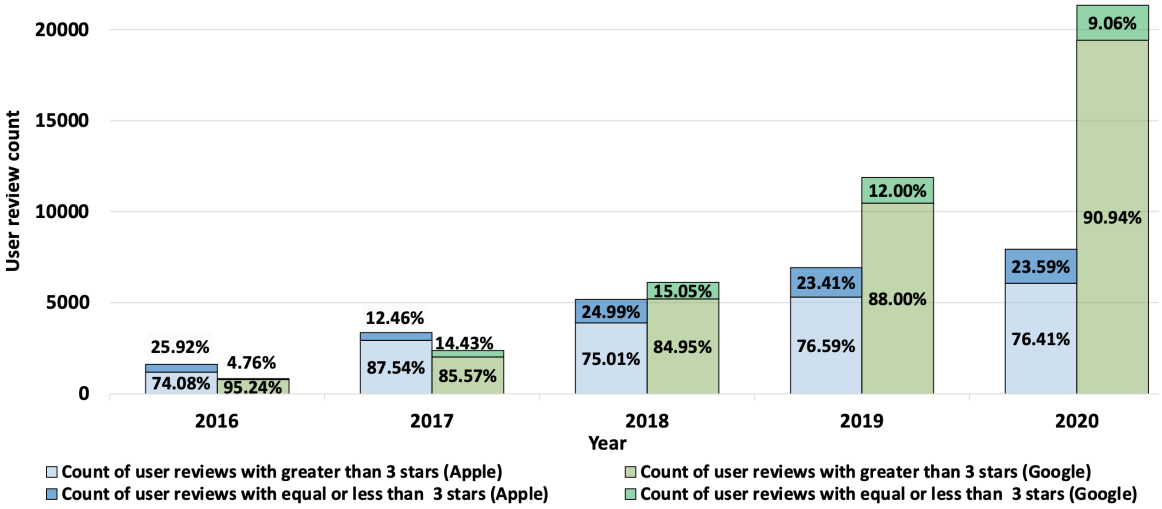


Figure 2: Number of user reviews published in Google Play Store and Apple App Stores in the last five years for the apps considered in this study

depending on their needs. Remote therapy and coaching solutions are highly useful for patients who live in remote places or who have physical disabilities.

- **Technology tools:** We found four different types of technological tools and features in the apps we included. First, interactive coaching and listening modules (76, 46.3%) where app users interact with a dynamic tool based on their experienced symptoms, and receive suggestions on ways to improve their specific mental conditions. These resources offer suggestions for navigating around potential trigger symptoms, avoiding risky activities, and learning about different treatment options on their own. Second, gamification (8, 5%) is a technology intervention that has recently gained popularity. Various game-based apps assist users in assessing and coping up with symptoms through provision of virtual reward for setting and completing specific goals. Leveling up, collecting money, a map to display progress, rankings, and characters are just a few examples of the video game-inspired features. People's motivation is expected to enhance by rewards and incentives in particular. Third, AI/chatbot (7,4.4%), in which developers utilize artificial intelligence and machine learning techniques to create bots that can interact with users by talking to them, coaching, suggesting treatment, guiding them through exercises, and so on. Fourth, relaxing music and soothing sounds (7, 4.4%) that help people recuperate from various metal difficulties.

- **Continuous assessment:** Continuous Assessment refers to approaches in which mobile apps assist users in better comprehending their present condition by contrasting recent symptoms with a history of previous symptoms. It requires continuous user interaction, as it relies

on user input, which is typically provided in the form of various surveys. The two types of assessment are: (a) symptom reporting, tracking, and monitoring, where users provide information about their current status in the form of valid established questionnaires, and the program assists them in monitoring and tracking their progress. Apps under this category usually send reminders for inputting symptoms at specific periods throughout the day, and then summarize the data through visual feedback to track user progress. This feature is present in 86 (52.5%) apps, and (b) self-reflection tools, which are offered through surveys and teach the users how to choose a better path, what to do in a critical situation, and other things by allowing them to explore on their own. These apps prompted users to investigate the evidence surrounding a catastrophic thinking, posed questions about the negativity in the thought process, and offered examples of how to combat the relevant negativity in their thoughts. 56 (34.4%) applications have this feature.

- **Community support:** Interacting with peers (anonymously) who are dealing with similar MH challenges or are already on the road to recovery enable users to open up about their problems and discuss viable recovery paths by exchanging support and advice. This peer-to-peer support is provided by 14 (8.8%) apps. Appointed moderators and super-admins commonly moderate these peer-to-peer discussions. A few of these apps also allow users to enlist the help of friends and family through their phone contacts and social media accounts.

	Google Play	% total reviews	Apple App Store	% total reviews
Total number of reviews	54554		32534	
1. After applying the length (#characters¿= 200) criteria	10485	19.22%	16762	51.5%
2. After applying the timeline (January 1, 2019 - May 1, 2021) criteria	8556	15.68%	9990	30.71%
3a. After applying the star rating (¡=3) criteria	1620	2.97% ; %reviews after applying 1 and 2: 18.93%	2266	6.97%; %reviews after applying 1 and 2: 22.68%
3b. Star rating (¿=4) and contain negative components	539	0.98% ; %reviews after applying 1 and 2: 6.23%	498	1.53%; %reviews after applying 1 and 2: 4.98%
Total number of reviews considered in the study: 2159 (Google Play Store); 2764 (Apple App Store)				
Average number of Words per Review: 66.67 (Google Play Store); 113.8989 (Apple App Store)				
Total number of apps considered after applying Filters: 117 (Google Play Store); 76 (Apple App Store) (P.S. 29 apps were available on both platforms)				

Table 1: # of reviews from Google Play and Apple App Store after applying each inclusion criteria

2.5.2 Perceptions and Concerns Expressed in the Critical User Reviews

Both the number of mobile MH apps and their consumers have increased significantly in recent years, which has resulted in a huge increase in the quantity of user reviews. Figure 2 depicts the

Broad categories	Major themes	Number of reviews	Percentage of reviews
Inconsistent moderation policies	Guidelines not being followed properly	267	5.4%
	Biased behavior from moderators and allowing discriminatory attitude	530	11.2%
Lack of transparency	Therapist matching techniques and qualification of the therapists	397	8.1%
	What to expect from free vs. paid subscriptions	507	10.3%
	Ensuring privacy	249	5.1%
Unavailability of In-time and Emergency Support	No or little help during crises	212	4.3%
	Out-of-context responses and little interest from coaches and active listeners	298	6.1%
Abusive community	Abuse and harassment	409	8.3%
	Preaching harmful thoughts	56	1.1%
Useless Chatbots and conversational agents	Not providing helpful/practical solutions	207	4.2%
	Failing to understand context and provide personalized responses	567	11.5%
Others (not considered)	UI issues and software bugs	1409	32.3%

Table 2: Frequency of themes in the user reviews (number and percent of reviews expressing a particular theme, applied non-exclusively)

recent surge in reviews across both the Google and Android platforms. Between 2016 and 2020, the number of reviews has increased by more than 20 times and 4 times in these platforms respectively. Figure 2 shows that the reviews with three or fewer stars, which we refer to as "critical reviews", accounted for about 10% of total reviews on Google Play and nearly 25% of total reviews on Apple App Store. This may not seem to be a significant percent of the userbase of these apps. However, we have to keep in mind that many of these users are MH patients who may be experiencing serious symptoms, or may be on the verge of a crisis. It is crucial to acknowledge these users' real-life experience of using the apps and take into account in planning, designing, and developing apps.

Table 1 represents a flowchart of the inclusion criteria applied and the number of reviews retrieved in each step. The final number of reviews was 2159 (3.95%, n=54554) on the Google platform and 2764 reviews (8.5%, n=32534) on the Apple platform. When it comes to writing critical reviews, users of iOS apps seem to be more comprehensive than users of Android apps (i.e., the average word count per review from Google Play is 66.67 vs. 113.89 from Apple App Store).

Our analysis identified twelve major concerns (themes) within these negative reviews. We broadly characterize these twelve themes into five overarching categories, which we report with examples from the data (themes and broad categories are listed in Table 2).

Inconsistent moderation policies

8.8% of the apps we considered included a community feature that allowed users to share their MH problems with other users of the app and receive support. Like many online health communities (OHCs), these embedded forums and communities usually have designated admins and moderators

who control destructive member posts and help answer questions to provide important support for user engagement. However, the user reviews frequently indicated that the moderation processes and policies implemented in many of the apps are quite inconsistent. Users have complaints about moderators not following established guidelines and showing biases against specific demographics, genders, races, and religions.

Guidelines not being followed properly: Almost all of the apps with community features have a set of guidelines in place to protect their most vulnerable users, and moderators are tasked with ensuring that these guidelines are followed. However, there have been instances where moderators have failed to follow the standard, resulting in frustrations among the users. R1 elaborated on how moderators erroneously deleted content without following app policies:

> *"Some admins have been harassing people by deleting their posts that are in the guidelines and are not bad at least. Many people have issues with the admins...They really need to regulate who they allow to be an administrator." - G0470270*

Users have also expressed their dissatisfaction with the sometimes-confusing nature of these standards. Posts are often deleted and users are banned without any explanation or warning. Even when explanations are provided for deleting posts or comments, they are vague. The users felt that the moderators interpreted and enforced the norms according to their will and did not care about informing users about their actions.

At the same time, it's critical to acknowledge that maintaining rigorous rules in moderating content and banning in health peer-support communities is not always a viable option, as these decisions may result in worsening symptoms, a relapse, or even an individual's death. Some users have expressed their dissatisfaction with the moderation policies, claiming that they are too strict to express their negative emotions or to vent.

> *"For an app that's about self-expression they really do love to delete post that they don't like. This app has a stick up its a**, don't waste your time." - G0670115*

It gets worse when actual offensive posts are not deleted even after reporting them. Users often complained how the moderation is *"two-faced"*. One user criticized how his/her approach of presenting a condition relating eating disorders was misconstrued and removed because the moderators lacked an understanding of particular expressions of specific MH problems. Users have reported that even after identifying and marking a large number of offensive posts by themselves, just a handful are removed by admins. Red-flagged contents, such as "threatening suicide", are left unchecked.

> *"Sometimes it can be two-faced. For example, I see posts that say 'I want to commit suicide' that never get taken down. But when I post 'I want to purge' as in I'm dealing with an eating disorder and looking for help to overcome the urge my post gets deleted. It actually makes the feeling of loneliness worse." - G0110097*

Biased behavior from moderators and allowing discriminatory attitude: Users have occasionally expressed their disappointments with moderators "feeling entitled" and "abusing their power" by showing conscious biased behavior. G0470318 stated in a review that the moderators are making decisions biased towards their friends. They loosen the rules for them, and others are being unfairly punished for minor infraction of the rules. These actions place individuals in an even worse situation, as they are unable to convey, whereas biased behavior caused aggressive users to transgress the rules even more. One user compared it to biased and harmful social media groups where moderators solely target people with whom they disagree.

> *"Admins are biased, manipulative, controlling. They almost never delete the hateful posts which genuinely blow the guidelines, as those guys are friends" - G0470318*

Moderators also showed biases against different religions, races, and genders. G0470273 cited an example of moderators displaying hatred for a certain faith and allowing a violation of the guidelines in that situation.

> *"... Post bad about Islam/Muslim; it stays there, but one getting bullied & verbally abused for being a Muslim or preaching Islam, no one cares. Atleast ADMINS and MODERATORS need to show some humanity rather dealing own Islamophobia." - A04700273*

Similar to the biased and wrongful behaviors of the moderators towards specific demographics and ethnic groups, community members constantly showed racism and sexism that ruined many users' interest to share and ask for help.

> *"I have a serious problem with a toxic community when you allow people onto the platform who send you hate because of your identity or sexuality, I have mental health issues and enough problems in my life without having to worry about being told that I'm unnatural and should die." - G0470103*

Another interesting finding was that, whereas usually considered minority groups experience racism, there were opposite cases as well.

> *"It is racist against whites and clearly run by Liberals. There's a Black Lives Matter*

section and white people are forced to silence. Admins delete the white people's comments and allow blacks to shush them." - G3010005

Users with MH concerns may feel violated and mistreated as a result of the chaos in the moderation process, which may have unintended repercussions such as aggravating their conditions. Furthermore, such moderation behavior can make users feel uncomfortable to share, leading them to leave or uninstall the app.

Lack of transparency

Transparency in policies and services is becoming a prerequisite in many fields, particularly those where people's lives are at risk (81). According to our analysis, there is a lack of transparency in key areas of mobile MH apps, such as how users are matched with their therapists or coaches and what their qualifications are. Users often get confused about what to expect from a paid version of the app, causing frustration after spending a handful of money. Finally, some applications do not clearly specify what user information they collect and how, infringing on users' privacy.

Therapist matching techniques and qualification of the therapists: One major theme that has emerged from the user comments is how difficult it may be to navigate while looking for professional aid, such as counseling. Some apps provide unclear and incomplete instructions in the interface for searching for therapists, while some other apps completely ignore user-specified criteria in finding them a competent one. Such realities were beautifully highlighted by A1780577. The user not only complained about the app's complex interface, but also about how the application's therapy methods weren't particularly appropriate for her. It's critical for developers to retain integrity in terms of what their application offers, who their target audience is, and how the techniques/treatments they apply will benefit the consumers.

"I had to be experimental and ask questions to the therapist about technical aspects. A prerecorded video introduction from a potential therapist would be ideal for selecting the right one. There were short bios about a paragraph long. In hindsight, it turned out not being enough information." - A1780577

Users were concerned that the algorithms utilized in the apps to match therapists in regards to their needs were ineffective. Users recommended that the developers must pay close attention to the specifics in terms of recommending treatment procedures that are tailored to the demands of the consumers, and explain these techniques in the app in a manner that the users can understand and utilize to best meet their needs.

> *"When I told my first therapist that my family was concerned of my constant criticism & negativity, she asked if I had tried being more positive. GeE wHy didn't I tHiNk of ThAt. Really? Had I been able to be more positive, why am I paying for therapy?!?" - A1780098*

Many users suspected that the therapists they were matched with were not as professional and qualified as their in-person therapists. At times, users felt that the recommended therapists were being judgemental or dismissive which was totally unexpected to a patient suffering from a MH condition. Nonetheless, some of the users tried to find out a way around this so that they can continue using these apps.

What to expect from free vs. paid subscriptions: Most of the marketplace apps require users to buy a paid subscription to use full version. While users frequently called these app development companies "capitalist" and commended the developers who made the full version free during the COVID-19 pandemic, the question of whether MH apps should be free of cost or not is not the focus of our study. However, we sought to understand users' expectations from free and paid apps. Our analysis revealed that many apps provide conflicting information about what additional functionalities to expect from the paid subscription, switching to premium from a free trial, and how to request for refunds. Users occasionally downloaded and installed the apps in times of crisis to receive support, as those apps catalogued a bunch of seemingly useful features without clearly listing if and how much the users had to pay to use each feature. It often took a lot of time and effort on the user end to understand that a particular functionality in the app that attracted her is not available in the free version.

> *"I pressed a button to browse therapists and it charged me an exorbitant amount! Perhaps that is where APP is placing their bets. Exploiting people in their time of need. Nowhere on or near that button did it say that I was agreeing to pay." - A0490010*

The payment process' transparency was also challenged by users. The applications sometimes made checking the status of user payments too complex. Additionally, they were annoyed with getting charged without any prior hints.

> *"I'm a bit confused when it shows that you could have a 7 free day trial but it says that first you have to pay. How is it a free trial if u first have to pay for it?" -G0150087*

Ensuring privacy: Even though certain treatment processes require sensitive information from patients in order to track and monitor symptoms, match them with interactive tools, provide bet-

ter therapy and counseling services, etc., the users get wary if the apps do not provide sufficient information about what data will be collected through the app and why they will be required. For example, in G0150290, the user was terrified since their encounter with the app seemed like they were being monitored. As a patient who suffered from anxiety, it worsened their situation.

> *"When I got this app I asked it what color my shirt was, and it said the correct color. I feel watched, and scared. I thought this was supposed to relive anxiety, It made my anxiety worse." - G0150290*

A2740021 gave another example of collecting superfluous information from patients in an insensitive and inaccurate manner.The manner in which programs collect information from users is a crucial point that must be addressed.

> *"This app is always copying clipboard contents, when you open the app. This is highly illegal...Be careful. It may steal the passwords or banking info if you copied from somewhere to the clipboard!" - A2740021*

Some MH mobile applications need users to input information over an extended period of time for purposes, such as symptom monitoring, visualizing progress goals, journaling, etc. Users expect such apps to be able to securely save the information they provide and retrieve whenever needed. Additionally, in many cases the apps contained very little or no information about when and how the collected data is disseminated or shared with other stakeholders or third party. We found out that users became really skeptical of using an app if it cannot ensure the correct maintenance and security of their private information.

> *"It seems that some of the data about me and my sessions must be recorded for the app to work. But then, who can see this VERY private information about me? Can my data be sold, in aggregate or with personally identifying tags?" -G0150349*

Unavailability of In-time and Emergency Support

When dealing with MH patients, it is critical to understand that a crisis can be triggered for anyone at any time, and if it is not addressed in a timely manner, it can have serious effects. A lack of concern to address this very important issue has been expressed in the reviews. After discussing a potential crisis symptom, users frequently have to wait for a long time to hear from their therapists and professional coaches, or they get short and irrelevant replies from their coaches and counselors.

No or little help during crises: Many times people install a MH app when in crises, and need

immediate and urgent assistance (i.e., in the middle of the night when help may be unavailable, when unable to afford professional therapy, etc.). Our analysis revealed that the apps frequently failed to provide support during emergency in a timely manner. The glitches and bugs in the system often added to user frustrations.

> *" If anyone were in a state of such emergency, in need to call the suicide hotline, this app would worsen the conditions in which the struggling user would already be sick of; being ignored and feeling helpless." - A1630001*

Users also reported how usability issues with app features could worsen the crisis situation instead of helping. Almost all apps tend to have usability problems in their interface. However, the user base of these apps include people who are prone to a MH crisis and they found the usability errors interfering with their conditions.

> *"I was in a crisis and was doing a little better with the help of this app, but when I minimized to spell check something, it started me from the beginning. It made me repeat myself 3 time and sent me into a worse crisis" - G0150004*

Another crucial component in ensuring continuous support is to provide in-time responses from therapists, counselors and peers when questions asked. However, consumers have complained about long wait times and lack of follow-ups from their assigned coaches.

> *"My therapist straight up told me this is not a chat room, stop messaging me as often, I won't answer until tomorrow. I was under the impression that the app would give me help in a timely manner. I mean why should they give you a button for a deadline if they are going to ignore it!! Not to mention my first therapist on the app missed my deadline four times ..." - A1780031*

People lose interest in taking treatment while waiting for the prolonged wait period, which often leads to dropout. Users described the aggravation of having to wait so long for a basic response.

Out-of-context responses and little interest from coaches and active listeners: Despite the fact that these apps promised professional meetings with therapists and coaches, users (even after a long wait time) frequently received short, generic, out-of-context responses showing little professionalism from them. Furthermore, they (therapists and coaches) occasionally showed little interest in their patients and the therapy procedure, making consumers feel abandoned.

> *"When I was assigned someone after waiting for a week, I explained my problems in a*

> *paragraph about my issues with focusing and maintaining motivation. He told me to clear my head & just study, which is completely useless because I obviously can't. When I told him that didn't help, it's not that simple, he replied with a sad-faced emoji. His 3 word response showed little interest & professionalism." - A1630001*

When working with a sensitive patient, this is a serious problem. Before selecting therapists and coaches for their applications, developers must ensure that they are well-trained and possess appropriate subject knowledge. They also need to make sure that users don't have to wait long for a response by balancing the number of users who can use the app vs the number of coaches or counselors who can help them.

Abusive community

Providing a community feature in the apps may be useful, since current research work on MH support has found online peer support to be beneficial in various contexts (206). Users feel better after talking about their recent circumstances and concerns, and receiving advice from others who have been in similar situations. However, it is questionable how effective the community features in existing mobile MH apps have been, since we found numerous instances of toxic activities such as racism, sexism, trolling, harassment, and prejudiced that encouraged destructive ideas such as suicide.

Abuse and harassment: Users reported being subjected to abusive conduct and harassment. Trolling, unpleasant behavior, and harassment were all common, along with negative comments about race, gender, and religion. People with various MH concerns often use these applications as a safe place to vent and share their predicament with like-minded people, as G0470064 noted. If the atmosphere is toxic and not pleasant, users get genuinely frustrated.

> *"The vast majority of comments were not only discouraging but they were absolutely rude. Very disappointed in an app that projects itself as providing help and encouragement to those who reach out for help. These humiliating and hurtful comments could actually make the individual feel a thousand times worse." - G0400064*

Users frequently blame prejudiced and harmful users who joined the community solely to take advantage of others. A few of the apps were targeted for the teenagers and for many predators it was easy to exploit these apps to abuse that population.

> *"A good percentage of the people here are only there because they can get off to teenagers*

> *that don't know when they're being used. Weird questions about virginity, sex, and personal boundaries are common. Of course you could say this about any social media app, but this one is supposed to be used as a therapy alternative." - A1630032*

Oftentimes, these abusers get away with their activities by blocking the victim. On the other hand, many apps make it difficult for the users to navigate options of reporting or blocking the abusers, and even worse, some of these blocking features do not serve the purpose as expected by the app users. Blocking a user does not prevent them from harassing others; in some applications, they can still harass others by commenting on other posts, checking in on the victims, and so on.

> *"If someone blocks you you cannot report them, this feature on an app filled with sad, lonely, socially unaware individuals is a terrible idea. It's like the app was made for abusers to use it however they want, no negative repercussions." - A1530032*

Preaching harmful thoughts: When the environment becomes too unhealthy, it might lead to detrimental thoughts, such as advocating suicide. G0470163 witnessed one such terrible situation in which, rather than offering intelligent advice and consolation, friends made cruel comments and pushed the victim to continue down the path of destruction.

> *"The app has turned into more of a popularity contest and toxic community. I read a comment stating 'Please do it' on someone's post who said they wanted to die. The community has gone down the drain, it gives the opposite aspect this app originally tried to deliver; a safe place to express your feelings without any fear of trolling." - G0470163*

Peers often encouraged unhealthy lifestyles and harmful thoughts, such as "suicidal ideation," "self harm," and "dangerous self medicating."

> *"I've seen people be absolutely ignorant in the comments of people posting about their struggles, even worse, people blindly encouraging others to become more unhealthy when it's already dangerous." - G040068"*

Providing a safe environment where users can unwind and communicate to one another should be a key concern for programs that provide peer-to-peer support, many of these apps failed to offer this after projecting themselves as such. This is also directly related to the preceding theme in previous section, which indicates that the moderation procedure isn't being carried out effectively. To provide a safe atmosphere, consistent and effective moderation is required.

Useless Chatbots and conversational agents

Researchers have proposed several approaches to employ artificial intelligence (AI) and machine learning algorithms in improving lifestyles and decision support systems in recent years, thanks to new innovation in AI algorithms. One method is to include these algorithms into a bot that can converse with people who have MH issues. These are known as 'Chatbots'. This theme addresses concerns relating to many aspects of chatbots, such as failure to be helpful, lack of learning capabilities, and risk of triggering MH issues through incorrect responses.

Not providing helpful/practical solutions: Chatbots assist patients with mental illnesses in a variety of ways, such as by interacting with them, providing comfort through intelligent responses, recommending healthier fitness and lifestyle choices. They are efficient in some respects, but have been criticized for being unhelpful. They have occasionally failed to respond to concerns in a responsible way, and have also been found to be ineffective. Understanding the situations is critical in the rehabilitation process, and these bots have occasionally failed to perform what they were designed to do due to poor algorithms and construction. G0150093 mentioned such a case where AI was poorly constructed and failed to understand and address the issue.

> *"Feels like more of a resource finder as it could not understand what I was trying to ask it. The Ai wanted to ask all the questions and set me on a certain path. I hope that Ai will advance in the future and be able to help more. Also it told me I needed to sleep when I didn't and kept telling me how it would help my energy levels, when in fact I'm almost always tired and sleep rarely helps." - G0150093*

Furthermore, understanding and mimicking how an actual person responds to questions is crucial. Word choice and voice tone are vital when interacting with users experiencing MH concerns, since all these can have an impact on the outcome of the interaction. User reviews demonstrated how bots often conversed as if users were pets, and undermining the positivity that the bot was attempting to convey. In addition, dismissive behaviors from AI agents were noticed frequently.

> *"This morning I received the daily text, addressing me as a dog owner would: 'Who's a good user? You areeeee. Yes you are! Your Monday motivation hack: Talk to yourself like a pet.' Seriously? This is demeaning in its endeavor to encourage patronizing tones on oneself rather than truly home an empowering message. Speaking the language of dumb and basic was never considered a strategy for self-respect which makes this even more asinine. ..." - A1660001*

> *"It cut me off & said goodbye before I was finished. For someone who is emotionally*

unstable I imagine even a robot dismissing u to be a blow u don't need." -G0150222

Feelings and emotions can be difficult for the bot to comprehend at times. The ability to handle an emergency situation is a vital trait that these AI agents must possess. In other programs, users have expressed dissatisfaction with how they respond in those scenarios. Poor usability and over complication were mentioned by some users. Even though users realize that bots can be occasionally helpful, AI ineffectiveness frustrates users more.

Failing to understand context and provide personalized responses: Users reported having "incredibly boring conversations with the bots as AI misses a lot of contextual clues" (G0150164) or "failed to pick an appropriate reaction for the context of the conversation".

"BOT very quickly starts to feel exactly like that: a bot. It loses context, chats start to make less sense. Some things it says are very out of place. It feels like it's ignoring you and is expecting a response that you obviously didn't give" - A1660034

Some users were more irritated. While they understand it is not an actual human, and that, talking to it made them feel more at ease, it was at times of no help as it could not answer basic inquiries and kept repeating replies. Users have been led to assume that the bots don't comprehend what they're saying because they respond with incredibly irrelevant or repetitive responses. G0150304 demonstrated how frustrating these responses may be, and it eventually made the user exit the applications.

"I was trying to tell BOT that I'm really glad about my day and I told BOT why, then it keeps asking me 'Why do you say so' for about 3 times like what the?! I feel like it's the bot who needs help, not me." - G0150304

Furthermore, consumers expressed dissatisfaction with how generic some of the responses and recommended activities were. They reported of being "offered more generic suggestions than they could find from simple google searches." It became worse when the bot suggested activities that the user was unable to perform due to physical or ambulatory reasons. In general, the reviews emphasized that mental health conditions can cause complex and varied moods among users that are hard to comprehend, and that the users expected these apps to incorporate features adaptive to these feelings and moods.

"Bot doesn't understand simple keywords like 'physical pain' or 'headache.' Offers more generic suggestions than I have found on simple google searches. ..." - G0150203

> *"I just have one suggestion: in worry relievers add more moods to it because sometimes what you actually is feeling inside is not as simple as anger,happy ,sad and okay. Like I usually feel nothing at all or so numb and that is the situation that i want to deal with the most on daily basis." - G0150277*

Some bots occasionally asked "preset questions" that did not apply to many, "steamrolled users into a limited range of non-useful and non-personalized responses," acted "judgmental rather than self-affirming"," advocated only a single line of thought."

> *"I expected an AI but it seems preset questions, though questions are well placed. The problem lies when the human deviates from the line of talk, begins rambling. Then it gradually fails to understand us, because the questions are preset." - A2640093*

In one case, a bot with the intention of reforming a user's negative thoughts erased and modified part of his journal that frustrated him and he started questioning the reliability of the the app. This is a practical example of how an intervention strategy that's been supported by the literature to improve particular MH conditions can backfire, and thus illustrates the significance of understanding the consequences that can result from adding a feature.

> *"APP censors your journal entries. what kind of journal CHANGES what you wrote?! It asked me my thoughts, I started waxing poetic metaphors about my problems. It censored it so I don't even know what I was talking about anymore! Some of my conversations are completely lost, they don't show in the journal." - A1630013*

2.5.3 Aspects and features of MH apps liked by the users

Due to their convenience and accessibility, mobile mental health apps have proven to be beneficial. Despite the fact that we conducted thematic analysis on the negative contents of all reviews that met our filter parameters, we evaluated all of the positive reviews after applying filter 1 and filter 2 *(Table 1)* as it's also crucial to understand user preferences of different features of these apps. We point out a few of the significant themes from the positive reviews, that we think are necessary to consider while implementing MH apps and developing guidelines for these apps:

Easy navigation: Users appreciated the apps with simple and easy flow of navigation. Low cognitive skill is preferred by those with various mental health difficulties such as worry, or stress, according to the users. They loved "steps that gradually walk them through techniques they learn in therapy," "design or amount of information that they don't feel overwhelmed with," and "the

good little tips".

Actionable and practical features: Users positively acknowledged the activities and exercises that are ''bite-sized", consume low attention span, and take short time to complete. Reviews positively mentioned the visual and interactive modules that put theory in action.

> *"Great step by step daily instruction. I've read a lot of CBT workbooks and this app puts them into motion. As a highly visual learner, the videos in the app are huge in helping me to remember and apply techniques." - G0260019*

> *"Really enjoying the actionable steps in this app. The stages of each journey don't take too long. Everything is bite-sized and helpful." - A0390002*

Higher benefit-cost ratio: Most apps do not allow users complete access if they do not pay for the subscription. A handful of reviews revealed that they compare these apps with alternative options, determine the benefit-cost ratio for them, and do not mind paying if they think the ratio is considerably high, especially for apps that support treatment through expert therapists, guiding coaches, and active listeners.

> *"I've tried face-to-face therapy before. I always feel like I have to perform. Writing is so much more therapeutic for me. I am ok with paying the subscription and have so far had one session with my coach which was excellent." - G0120007*

Many people prefer having therapy and coaching sessions via mobile devices because it is more convenient. Furthermore, many customers had no problems paying because these app-delivered therapies were covered by their health insurance. There was an abundance of positive comments about how they didn't have to worry about the payments because they were taken care of by insurance.

> *"My insurance covers the cost of the upgraded version. Thank you and bless you to the masterminds staffing this app. 'Providing hope' to others is a gift!" - A2770005*

Chatbots as companions: Chatbots are a specific tool that, according to users, can have both positive and negative repercussions. One of the major benefits we have observed from user reviews is that, chatbots can makes them feel less lonely. Many users noted that now that they have chatbots in their smartphone apps, they have someone to talk to 24 hours a day, seven days a week.

> *"It's like you're talking to a friend but it's really more than a friend and really, REALLY GREAT FRIEND! You'll realize that this app is much chatting with a such very sweet*

*and cute friendly *animation that will help you with your problems!" - G1120063*

One of the most notable distinctions between talking to a person and interacting with chatbots has been the bots' lack of judgment. People do not anticipate AI acting in the role of a professional therapist. Rather, they remarked on how effective it was at directing their good thoughts and making them feel better. Many users were concerned about which pronouns (he/she/them) they should use for chatbots, proving the fact that it had been acknowledged as a distinct entity to them.

> *"BOT helps me to change my manner of thinking within minutes. From being pessimistic and hopeless to feeling confident and hopeful. I check in a few times a day depending on when I feel myself slipping back to old ways. I believe that over the long term this app will help me in all aspects of my life exponentially. It's like you're talking to a friend who really knows how you're feeling and cares." - G0870009*

2.6 Discussion

In this section, we bring together insights from the findings to provide recommendations and future directions for mobile MH app design.

2.6.1 Research and Design Implications for MH Apps Development

Towards consistent moderation and community support Even though careful and considerate moderation is an essential aspect of developing a community support system (275), our analysis demonstrated how poorly the app-embedded communities are moderated (i.e., moderators showing biased behavior against specific races, genders, and other marginalized communities, allowing harassment of members on multiple occasions, etc.). Inconsistent and careless moderating decisions may allow abusive community behavior such as trolling, harassment, and discrimination to flourish, severely impacting users and triggering dangerous conditions in this sensitive context. These findings have a major bearing on the necessity of developing and following specific moderation policies curtailed to the community needs. Following directions from prior studies can be effective, such as appointing moderators with well-defined credentials to build trust in OHCs (i.e., if the moderators are health professionals vs. with no clinical background) (138; 130) or relying on indirect forms of social control to avoid punishing kinds of control (178). Because mobile MH communities are significantly smaller in scale than OHCs, future research should build on these findings and their implications on a smaller scale platform by gaining an understanding of the underlying mechanisms that explain how different styles of moderation work. Some small-scale online communities may not

have enough resources and capacities to appoint dedicated moderators and regulate a continuous moderation process, peer-led moderation can be considered as an alternative there. Peer-led moderation can support peers by offering non-medical guidance, and plays a crucial role in community involvement (246; 233).

Moderators often become less engaged as they feel overwhelmed to manually (dis)approve large volume of content (234). Future app designers may consider employing natural language processing techniques to enhance the pre-moderation stage and flag content automatically for human review. However, there are various concerns and challenges with AI content moderation, which can be addressed by including humans in the feedback loop and boosting the model's transparency (101; 231**?**). Automated techniques can be employed for continuous assessment of moderation decisions to mark racist moderation behaviors.

Towards better transparency Many users were genuinely concerned and puzzled about what information is being collected and how. Whereas, data transparency is critical to gaining patients' trust (139) in health-related systems, rules, regulations, and internal policies on information sharing have made it difficult to maintain complete transparency (248; 154). To strike a balance between what users desire and how much information third parties are prepared to offer, researchers should focus on collaboration between multiple stakeholders - users, providers, and developers and examine their values and needs in terms of data transparency (171).

Users criticized that they download an app when already feeling low, just to realize that the support they need is available through the app, but they cannot afford it. Additionally, apps frequently failed to provide a free trial and instead charge through hidden features in order to make profit, or even to collect the maintenance cost of the apps. Users expressed that a free trial would be convenient before paying for something that might not work for their specific needs. We understand that there are costs associated with developing and maintaining the apps. However, developers should carefully design the interfaces that make pricing information transparent. For instance, they can adopt simple visualization approaches that show the pricing breakdown and total cost upfront and provide reminders and warning notifications before any type of payment is charged. Moreover, free trials should be offered for the useful features, and if that's not the case, then inform users even before they attempt to install the app that none of the services offered are free. According to our assessment of positive reviews, if users find the tools satisfactory during the free trial, they are willing to pay for them afterward. Some people expressed their delight that their insurance covers the cost of the subscription. However, mobile mental health technology is seen as a convenient and

accessible tool for all people, regardless of their capacity to pay for health insurance. As a result, developers and policymakers should develop an economical solution for both individuals who have health insurance and those who do not.

Towards usability and UX Many users installed the applications to avoid and identify any unexpected crises, whereas the apps' inability to meet expectations worsened their mental health conditions. Users were perplexed about how they were matched with their therapists and counselors, since the assigned ones often did not communicate as expected. It's fair if the initial therapist someone is matched with is not right for them, but user complaints about numerous iterations of failed sessions with different therapists are very concerning. Our findings suggest users appreciate easy navigation, detailed explanation of how the tools work and motivational contents (such as success stories). A simple interactive checklist interface can be used to take user symptoms and other preferences as input and recommendation algorithms can be adopted to show them an ordered list of therapists based on their symptoms. Moreover, therapists should be offered with interfaces to easily list their qualifications, as well as their previous history of dealing with patients in anonymized case story format. Prior research suggests that it may bring more credibility to the users (79) and boost user satisfaction with the therapy process (61).

Towards app-embedded chatbot design

Users feel safe to share health conditions with a conversational agent over a real person (7), so chatbots should be utilized to its full potential to enhance mental health support for a broad range of users. User reviews revealed a number of shortcomings in existing chatbots, such as providing unhelpful and repetitive responses making the interaction unproductive, and being dismissive or patronizing that made users feel neglected and disrupted the treatment efforts. Our findings suggested that users have a preference for personalized responses, and conversations that normally draw on CBT techniques to assist users address any harmful thought and behavior patterns and build coping mechanisms must be more contextual and relevant, similar to prior work (46). Different methodologies for optimizing chatbot responses (45), including utilizing linguistic aspects, such as employing emotion or sentiment analysis (203) have been proved effective for better product suggestions and providing customer care. These implementations and approaches can be customized for chatbots in MH apps to more engaging and capable of displaying positive attitudes. Alternatively, if the bot isn't AI-powered, developers should make it clear to consumers that it's merely a bot with pre-programmed responses.

It is crucial to consider that the targeted participants may be going through different challenges and a crisis can happen anytime. The existing chatbots are unable to identify potential crises and provide effective options, as depicted in our analysis. Existing machine learning and pattern detection algorithms can be employed to detect potential crises from users' interaction with the chatbots and take appropriate action by alerting their healthcare providers or friends and family members (265). Our findings suggest following participatory machine learning approaches to design technology for vulnerable populations that can foster interactions among different stakeholders of these technologies. This technique will seem useful to academics for future research, as well as to developers and health care practitioners to implement technology-driven therapies.

2.6.2 Need for Developing a User-Centric Guideline for MH Apps Design

Although there are several guidelines available for developers of MH apps (Section **??**), none of them are comprehensive and they are dispersed across various organizational resources. No significant effort has been made to disseminate these policies, either (214). It is an unrealistic expectation that developers are able to search for and follow proper guidelines and recommendations on their own, especially if they have no prior experience. We provide suggestions (informed by our findings) on how app developers' and organizations' access to existing guidelines and regulations can be facilitated, and for creating a comprehensive centralized guideline for them.

The Federal Trade Commission (FTC) has developed a tool that advises developers on which federal rules and regulations they should adhere to, but only in the legal sense (Federal Trade Commission). There are other useful recommendations available established by various professional and medical organizations, such as expert guidelines from the American Psychiatric Association (APA) (17), non-profit organization psyberguide (**?**), etc. Bringing all these information together in one place via an interactive tool can help developers get aware and act more carefully before beginning to development an app. Such tools may prompt developers to select different aspects of the app they are planning to develop (e.g., targeted users, techniques and tools to be used, etc.) and show appropriate regulations and guidelines to follow based on their selections. These tools can potentially be disseminated through incorporating them in common Integrated Development Environments (IDEs), such as Xcode (143) and Android studio (70).

It can be deduced from the numerous and repeated user criticisms that their feedback is ignored even after many updates of the apps, making us believe that there is a communication gap between the users and the developers. To promote accountability, developers may occasionally be provided

with snapshots of negative reviews as well as the health consequences stemming from those. Many available apps are not evidence-based (27), and are only tangentially informed by evidence-based framework, but claim to provide an essential support system and treatment to a wide range of people suffering from a variety of mental health conditions (27; 72). It is critical that developers examine the efficacy of their enforced approaches on a small population before making it public.

According to our findings in this article, a comprehensive guideline should consider at least the following elements. Note that this is not a complete guideline and to implement an effective guideline more research and analysis of MH app data from other sources are necessary.

- If an app aims to provide peer-to-peer help, developers need to be reminded to implement techniques that keep the moderators and admins accountable. If volunteer moderators are being appointed in a community, a list of persuasive strategies might be shown to make them aware of their responsibilities and to be respectful of the policies.
- Users appreciate well-explained information regarding the support they would be provided with. Transparency is crucial in this context, and guidelines for development should show tips around enhancing transparency in all components of an app (e.g., algorithms/inner mechanisms, subscription, privacy of information, etc.). One of the simplest techniques in this regard could be inclusion of an interactive interface in the guideline that developers can use to design a customized FAQ page for their apps.
- Apps often collect private information from users to recommend appropriate and personalized treatment. Because there are so many guidelines for collecting, maintaining, and using data, it's impossible for developers to figure out which ones apply to their specific situation. A comprehensive guideline should keep track of the latest updates of these regulations and instruct the developers in an easy to follow manner.
- One of the vulnerabilities of these MH app users is the likelihood of their condition being triggered or worsened. A successful guideline should contain automated approaches to predict such consequences from a feature being developed. It can mine and summarize user and expert reviews from multiple sources and show warnings throughout the app development.

Certain issues caused users to abandon the app and the recovery process. It is important to keep in mind that most (if not all) developers are not medical professionals. Some troubles faced by the users may appear insignificant to them, that have serious health consequences. A variety of participatory research and design techniques (38), including workshops (175), focus groups (99), ethnography (39), etc. can be replicated on a relatively small scale to facilitate conversations among

different stakeholders. This should as well be followed for continuous evaluation of the applications' efficacy, to get guidance on what to do and what to avoid for the next update.

2.7 Conclusion

There is a gap in the current literature in terms of recognizing the challenges faced by the MH app users that affect adherence and engagement. First, to gain an empirical understanding of the current state of mobile app marketplaces we examine the descriptions of 164 mental health apps from Google Play and Apple App Stores. Next, we conducted an inductive thematic analysis of 1620 Android user reviews and 2266 iOS user reviews to better understand usage patterns, their perspectives and concerns. Our findings include criticisms of inconsistent moderation standards, a lack of transparency in how people were matched with their therapists/ counselors, as well as what qualifications they have, a lack of transparency in what customers get in the free vs paid version, and how user data privacy is protected. We also noticed a lack of real-time and emergency assistance and the ineffective implementations of app-embedded chatbots. This study's contributions include research and design implications for consistent moderating and community support, as well as improved transparency, usability, and UX. We also offered a few suggestions regarding the creation, implementation, and dissemination of a comprehensive guideline for MH app developers.

CHAPTER 3: AN OVERVIEW OF CHATBOT-BASED MOBILE MENTAL HEALTH APPLICATIONS: INSIGHTS FROM APP DESCRIPTION AND USER REVIEWS

Chatbots are an emerging technology that show potential for mental health care apps to enable effective and practical evidence-based therapies. As this technology is still relatively new, little is known about recently developed apps and their characteristics and effectiveness. In this study, we aimed to provide an overview of the commercially available popular mental health chatbots and how they are perceived by users. We conducted an exploratory observation of 10 apps that offer support and treatment for a variety of mental health concerns with a built-in chatbot feature and qualitatively analyzed 3621 consumer reviews from the Google Play Store and 2624 consumer reviews from the Apple App Store. We found that although chatbots' personalized, humanlike interactions were positively received by users, improper responses and assumptions about the personalities of users led to a loss of interest. As chatbots are always accessible and convenient, users can become overly attached to them and prefer them over interacting with friends and family. Furthermore, a chatbot may offer crisis care whenever the user needs it because of its 24/7 availability, but even recently developed chatbots lack the understanding of properly identifying a crisis. Chatbots considered in this study fostered a judgment-free environment and helped users feel more comfortable sharing sensitive information. Our findings suggest that chatbots have great potential to offer social and psychological support in situations where real-world human interaction, such as connecting to friends or family members or seeking professional support, is not preferred or possible to achieve. However, there are several restrictions and limitations that these chatbots must establish according to the level of service they offer. Too much reliance on technology can pose risks, such as isolation and insufficient assistance during times of crisis. Recommendations for customization and balanced persuasion to inform the design of effective chatbots for mental health support have been outlined based on the insights of our findings. **This chapter has been published in JMIR Mhealth Uhealth 2023 (URL: https://mhealth.jmir.org/2023/1/e44838; DOI: 10.2196/44838).**

3.1 Introduction

A chatbot is a system that is able to converse and interact with human users using spoken, written, and visual languages (10). In recent years, chatbots are being used more frequently in a variety of industries, including retail (145), customer service (63), virtual assistants (191), education (276), etc. due to the advances in the AI and machine learning domains. The Facebook Messenger platform

currently offers more than 300,000 text-based chatbots (JingYi). Till now, it has been used primarily for commercial purposes and profit businesses. However, prior research has demonstrated that this technology has significant promise in healthcare industry in treating patients and offering them support in a cost-effective and convenient way (37).

In the context of mental health (MH), chatbots may encourage interaction by those who have traditionally been reluctant to seek health related advice due to stigmatization (168). Chatbots are considered as an emerging technology showing potential for the mobile mental health applications to boost user engagement and adherence (204). Effectiveness of chatbots has been explored in HCI for self-disclosure and expressive writings (169; 236). People's self-compassion has increased through a care-giving chatbot (167). Young people with mental health issues have experienced various types of social support—appraisal, informational, emotional, and instrumental support—from chatbots (148). Additionally, chatbots have been designed to educate underprivileged communities about mental health and stigmatized topics (277; 222). The emerging evidence has shown user acceptance of chatbots for supporting various mental health issues and early promises in boosting health outcomes in physical and mental health domains. Despite the increasing adoption and the benefits of using chatbots to support health and wellbeing, there are criticisms that patient safety is rarely examined, health outcomes are evaluated in small scale, and no standard evaluation methods are present (6). It is pointed out that the recent developments in chatbots are due to a massive technology push, with little attention to human needs and experiences (41). It can lead to unintended negative consequences, such as biases, inadequate and failed responses, and privacy issues, all of which can negatively affect the quality of the experience of chatbots as a source of mental health support. Adoption of new technology, especially those heavily related to artificial intelligence and machine learning, relies on first ascertaining the levels of safety, effectiveness, and user comfort. Yet little research has been carried out to gain understanding of these aspects from real life user experience of mobile MH chatbots.

3.2 Objective of the Study

Commercially available mental health chatbot apps for popular platforms (e.g., iOS, Android, etc.) are used by a large userbase with varying demographic background. These users can provide feedback through ratings and text reviews (188). These platforms can be leveraged for gaining a holistic understanding of the features that recently developed MH chatbots offer and how users assess them. Knowledge of user perceptions from real-life experiences can inform future research and design

of more effective chatbots. Previous studies have identified user reviews as a significant source of comprehending the benefits and drawbacks of technology (66; 268). It allows researchers to incorporate community values and needs into the product design and improve the user-friendliness of similar future systems (258). Consumers oftentimes make decisions about using a new tool based on user rating scores and reviews in the online marketplaces. According to previous studies, users trust the reviews and feel at ease basing their decisions on them (258). For this study, we decided to analyze the commercially available well-known chatbot-based mobile mental health applications and corresponding user reviews from Apple App Store and Google Play Store. To get a comprehensive overview of these apps and understand the nuances in user opinions, we aimed to answer the following two research questions -

RQ1: What are the state-of-the-art features and properties found in chatbot-based mobile MH applications?

RQ2: What are the concerns and opinions expressed in user reviews published on mobile app store platforms about the usability and efficiency of chatbot-based mobile MH applications?

3.3 Background

Chatbots are a type of conversational agent that uses text and speech recognition to engage with users (10). Chatbots are software programs that can imitate human behavior and undertake specific tasks by conversing with users intelligently (10). Chatbots are being more commonly used in various web-based and mobile-based interventions as natural language processing and machine learning capabilities evolve (190). In recent years, it has taken on the role of a virtual entity that can act as a travel agent (19), customer service representative (63), financial advisor (145), personal assistant (191), etc. Chatbots are becoming increasingly sophisticated. Some of the available chatbots can have a personality of their own, store information about the user to deliver contextualized answers, and grow over time by learning about their users to provide better service (133).

3.3.1 Chatbots in Healthcare

Chatbots have recently received much attention in the healthcare and wellness industries (37). These bots were tested using a variety of elements and characteristics, depending on the behavior they were attempting to achieve. Chatbots function as digital personal assistants (29), allowing patients to learn more (277), obtain support (98), and take prompt action in response to new symptoms (30). Some chatbots can assist users in collecting medical data via text discussions and then deliver it to

(selected) doctors in a format that is easier to utilize for diagnostic purpose (98). Chatbots that promote physical activity and a healthy diet have been shown to be effective in prior research (280) and control studies by motivating users to meet their exercise objectives (24), recommending meal plans (82), pushing daily check-ins (126), offering relevant information (280), and addressing user inquiries (155), etc. They were also sophisticated enough to have daily adaptive little chats and show users' progress over time using analytics and graphs to encourage self-reflection (279). These studies found that chatbot interventions were effective in increasing physical activity, achieving relevant weight loss, and improving diet. Recent studies on the development of medical chatbots showed chatbots that would perform as a virtual physician/doctor. They can converse with patients about their medical problems and deliver a personalized diagnosis based on their symptoms like prescribe, suggest, and provide generic pharmaceutical information (25; 107). For example, Babylon Health chatbot (25), asks questions on symptoms, provides detailed information and additionally book an appointment with physicians. Florence chatbot (107) acts like a nurse that reminds taking pills, track health (body weight,periods) and help find a specialist in the nearby area. Safedrugbot chatbot (29) is designed to help doctors get information on the side effects of drugs during breastfeeding. Yadav et al. designed Feedplal and explored the potential of chatbots for breastfeeding education for in need population in India (277).

3.3.2 Mobile Mental Health Chatbots

Among the numerous chatbots being used in different aspects of health and wellbeing, chatbots in mobile mental health care have demonstrated effectiveness in broadening traditional therapy in a cost-effective and convenient manner (46). In recent years, the usage of chatbots in mobile mental health treatment has been recognized as a steady adaption. Chatbots can help those who are reluctant to seek mental health counseling owing to stigma, and they can provide more conversational flexibility (67). Rather than providing generic suggestions, they can deliver individualized suggestions and resources based on the needs and requirements of users (133). They were designed to identify mental health concerns (133), track moods (68), deliver CBT (87), and promote positive psychology (105). Chatbots have been shown to effectively reduce the severity of mental health concerns for people from different demographics and backgrounds, including people in rural communities (222) and shift workers who face accessibility issues in receiving traditional treatments (71), students suffering from anxiety and stress (87), employees of healthcare systems who require emotional support from time to time (135), veterans and adolescents who feel stigmatized in sharing their concerns (222), etc. Several well-known chatbots like "Wysa" (133), "Woebot" (87), "Replika"

(255), “Youper” (180), and “Tess” (97) were discussed in prior literature. Inkster et al. examined the potency of Wysa and found positive influence on reducing depressive symptoms (133). of In a randomized controlled experiment, Fitzpatrick et al. evaluated the effectiveness of the AI chatbot Woebot in giving CBT to college students with anxiety and depression, and they found that the Woebot significantly decreased depressive symptoms (87). Ta et al. investigate social support received from artificial agents in everyday contexts when interacting with the social chatbot Replika (255). Mehta et. al examined the acceptability and effectiveness of Youper (180).

In recent years HCI communities have been increasingly involved in designing chatbots for specific purposes such as self-compassion (Vincent) (167), enabling self-disclosure (169; 236), facilitate positive messages within social groups (192), to improve the elderly quality of life, to make them more active to fight their sense of loneliness (266), to support interpersonal skills (Sunny) (80), in stress reduction (Mylo) (35), etc. Kim et al. explored teenagers’ expectations when interacting with a chatbot intended to support their emotional needs (148).

The development and application of chatbots in mobile MH applications are in early phases, there are still considerable challenges to overcome in the development of this technology. According to recent studies, patient safety was rarely evaluated, health outcomes were inadequately quantified , and no standardized evaluation procedures were used (6). Some chatbots are reported to not be able to understand the complex use of language associated with a mental health crisis, and failed to recognize symptoms and respond appropriately (265). Privacy is a major concern for users of these applications; since users are still less familiar with this emerging technology, there is higher risk of exposing users to privacy risks through data sharing (157). Furthermore, while on one hand, poor adherence is a common problem with digital mental health interventions, on the other hand, some vulnerable people may begin to rely on them too much, which may lead to anxiety when these applications are not available (157). Overall, there is a need for better understanding of how all mobile mental health services can and should encourage the safe and ethical use of chatbots (6). Although a handful of research studies have shown potential benefits of MH chatbot applications, users’ real life experience and challenges are not fully understood. Researchers and developers in this space could benefit from a comprehensive analysis of the existing commercially available chatbot apps. People’s perceptions and mental models of chatbots can be studied to answer critical concerns such as how users gain trust in chatbots, user values and requirements in this space, and so on. A user-centric analysis will also assist researchers in mapping out the evidence-based framework for the proposed intervention and minimizing the psychological effects of such treatments.

3.3.3 User Review Analysis to Understand User Experience and Preferences

User reviews can be defined as feedback published by individuals about their opinions and satisfaction/dissatisfaction about the product (188). The star ratings and elaborated feedback in the textual reviews provide developers a chance to promote their apps and explore how to make them more user-friendly (258). For new or potential users the reviews work as a deciding factor to determine whether or not this app would be helpful for them (198). Few studies leveraged user reviews of mobile mental health, well-being, and behavior change applications for identifying usability issues or user characteristics (66; 268). User ratings and reviews were leveraged for a variety of reasons including determining why adherence to mobile mental health applications is poor (15), informing developers of design priorities rather than just guiding purchasing decisions (124), and gaining better understanding of ethical issues users faced (40), etc. Vasa et al. investigated the hypothesis that, despite the abundance of positive reviews for mobile apps, it's worthwhile to examine the negative reviews as well in order to gather useful data from users (268).

Our study is inspired by this body of work that considers user-generated reviews as a vital source for understanding user perspectives and derives meaningful implications from them (217; 198). It enables us to obtain data from a diverse group of people that would otherwise be challenging to collect using conventional data collection methods (256; 217).

3.4 Methods

In this section, we outline the techniques to select and filter the mobile applications for this study, the data analysis methods we used, the ethical standards we followed, our positionality statement, and the methodological limitations.

3.4.1 Selection of Sample Apps and Reviews

Selection of apps: To find out a comprehensive list of commercially available mental health apps that include a chatbot feature, we conducted our search through different sources. First we considered open access articles in recent literature about MH chatbots (6; 11). Next, we conducted search queries on two different expert mental health app review platforms: mindtools.io (Mindtools.io) and psyberguide (194). Finally we searched on two dominant online mobile app stores (Google Play for Android and Apple App Store for iOS). On the expert review platforms and the app stores, we used the search term "Mental health" and "chatbot". Additionally, we explored the "recommended applications" or "similar apps" section of the corresponding website after discovering a mental health

app with a chatbot feature to find out if those other apps meet our criteria. Without logging into a specific account, the search was done from the app stores' home pages. This action was performed to make sure that the system couldn't employ a ranking algorithm to prioritize any particular user choice. Since these applications represent the sample in (nearly) the same order that consumers would likely be exposed to and hence most likely to use, even though the search results may not be entirely comprehensive (as observed by convenience sampling), they still represent the sample.

After the initial search from these three different sources, we ended up with 19 applications. The authors carefully read the app descriptions, observed the screenshots of the app features, and in some cases analyzed these apps' promotional websites to make sure whether these applications includes a chatbot feature that provides support for different mental health concerns. We observed that some of these applications include intelligent Q/A based on Al/ML. Intelligent Q/A is based on a collection of questions, and by responding to them, it can offer individualized summaries, diagnoses, recommendations, and other information. In this article, we describe MH chatbots as intelligent machines that can simulate and process conversations with users about their mental health needs. In order to make sure that our list indeed includes the apps that fall under this definition, one of the authors opted to download each application separately (for the iOS platform) and use it for at least three days. The author has no known mental health issues. We also take this exploration as an opportunity to extract what primary features the apps commonly comprise. The author carefully observed how these apps work in terms of the significant aspects of mobile MH apps as pointed out in previous literature (6; 11; 265). Following all these steps, ten applications have been selected for analysis 7.

Selection of user reviews: We created scraping scripts using the Pyhton Selenium library to collect the ten apps' public user reviews that were accessible from Google play store and Apple app store. User reviews can illustrate examples of users' satisfaction/dissatisfaction with particular features of apps. Reviews are therefore recognized as an important source of information to gain insights from real life use of mobile apps [cite]. We used the two following inclusion criteria for filtering in order to extract recent and significant user feedback for the apps:

- **Timeline**: We took into account reviews that were posted between January 1, 2019, and May 1, 2022. We intended to collect the most recent user reviews because app stores change quickly with the addition of new apps and upgrades to already-existing apps.
- **Length**: Reviews containing 200 or more characters were considered in this study. A prior

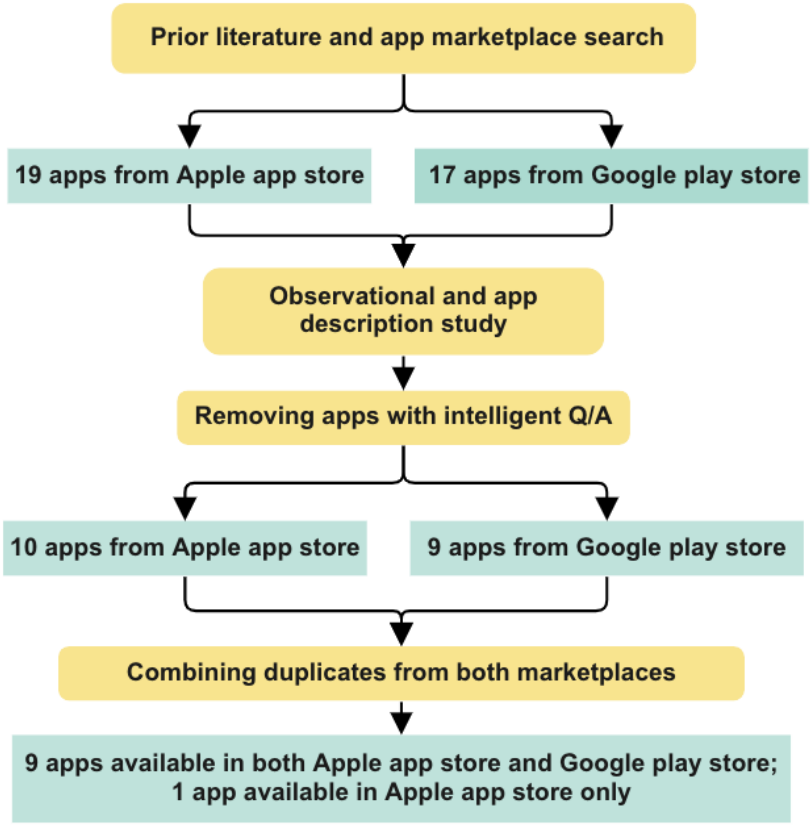


Figure 3: Flow chart of the apps selection process

study found that a relevant review typically has 110.8 characters (176). Because shorter reviews might not provide deeper insights in general, and are frequently false or promotional in nature [cite], the minimum character length was considered 200 for the scope of our study.

3621 reviews from the Google Play Store and 2624 reviews from the Apple App Store met all of the inclusion criteria. These reviews are based on 9 apps from the Google Play Store (only Elomia isn't available in the Google Play Store) and 10 from the Apple App Store. All of the reviews have a unique coding system that can be easily traced back to the application and platform from whence they emerged. During the analysis, the lead author was responsible for carefully reading each review and ensuring that all personally identifying information was replaced or removed.

3.4.2 Data Analysis

First, to gain an understanding of the descriptive overview of the commercially available chatbot-based MH apps, we analyzed app descriptions from the promotional website and online marketplaces and Incorporated the key informations in our observation notes. The observation note was then divided into six main themes with the aim of providing a comprehensive overview of these apps with

the collaboration of another author. The authors didn't include their judgements on the effectiveness of these apps. To understand user perspectives, the selected user reviews from the ten apps were examined using an inductive thematic analysis (200). Thematic analysis was chosen because it enables systematic analysis of large data sets and facilitates the comprehension of textual patterns while taking context into account (258; 147). Two passes were made during the analysis. We created open codes during the first pass in order to collect various perspectives from the reviews. We made an effort to record the subtleties in the particular insights provided in each review, which resulted in a high number of open codes that were greatly decreased through memoing and clustering (200). In the second phase of analysis, we memoed and clustered the codes using a constant comparison method operationalized as affinity mapping. Each open code was compared to others and positioned to reflect its affinity to emerging themes and clusters. The reported themes consisted both of ones that appeared consistently across multiple reviews, and also the ones that came from reviews that represented divergent responses and opinions. Findings from the reviews are described in the next section, and each quote is identified by the review's particular ID generated from platform and app name and random number.

3.4.3 Data Integrity and Ethical Considerations

App stores, like many other online marketplaces can have reviews posted by fake and paid users. However, prior research (176) showed that in "Health & Fitness" category the percentage of potential fake reviews is very low (about 6%). Fake reviews also tend to be shorter (176). Since we are mostly considering reviews with more than 200 characters, we assume that almost all of the reviews included are original and can be trusted.

We understand if the data or information is only accessible to a particular group of individuals or groups, it is unethical for researchers to use it (283). As a result, we made sure the websites from which we got the data were accessible to everyone and weren't just for some groups or populations. (283) We also want to underline that the use of both visual and textual components should not provide more information than is necessary. To make sure of that, even though those pages were public, we purposefully avoided publishing or disclosing any personally identifying information that was shared. The language of the user reviews reported here have been modified carefully keeping the meaning intact.

3.4.4 Positionality

Members of the research team included those whose families and close relatives had had mental health problems. One team member specializes in dealing with underrepresented and vulnerable populations, while another has expertise in intersectional stigma, online social support, and mental health care.

3.4.5 Limitations

Our selection criteria have certain underlying limitations. First, for our study, we primarily used ratings from the two most widely used mobile platforms (Google and Apple). Other mobile platforms were not taken into account for the purposes of this research. Second, it's likely that users who don't feel comfortable (or don't care) discussing their experiences on online platforms aren't contributing. However, we can confidently conclude that the perceptions we identified will be typical of user perceptions given the larger number of evaluations obtained from the two most well-known online marketplaces.

3.5 Findings

Applications	# of ratings in Apple app store	# of ratings in Google play store	# of downloads in Google play store	Age rating	Price
ADA	125	323000	5M+	17+	Free
Chai	27900	34000	1M+	17+	Free with in-app purchases
Elomia	193	N/A	N/A	12+	Free with in-app purchases
Mindspa	107	2970	500K+	17+	Free with in-app purchases
Nuna	68	93	10K+	4+	Free with in-app purchases
Serenity: Guided Mental Health	20	146	10K+	12+	Free
Stresscoach	None	495	10K+	12+	Free
Woebot	5500	11800	500K+	12+	Free
Wysa	13500	126000	1M+	12+	Free with in-app purchases
Youper - Self Care Friend	14400	49100	1M+	12+	Free with in-app purchases

Table 3: A descriptive overview of the selected ten mobile mental health applications with a built-in chatbot technology

For this research purpose, we chose ten commercially available mobile mental health applications that has a built-in chatbot feature. All of these applications except Elomia are available in two most popular platforms - Apple app store and Google play store. Elomia is exclusively available for iOS. A descriptive overview of these apps has been shown in Table 3. All of these applications are extremely popular both in terms of number of downloads and number of ratings. Thus, we can assume that a comprehensive overview of these applications can assist us in understanding perspectives of a wide

and diverse userbase.

3.5.1 Overview of the Aspects Commonly used in the Chatbot MH Apps

In this section, we present an overview of the apps in terms of their few important characteristics. These characteristics were selected from a previous literature review article on mental health chatbots (6). These are: primary purpose, targeted concerns addressed, conversation style & media types used by chatbots, availability of crisis support & evidence-based techniques. An outline of the criteria and types are being presented in figure 4.

We examined the app store descriptions to understand the primary goals of these applications and to identify how they are branded. We discovered four different types of purposes in all, with "digital coaches" being the most prevalent (5 out of 10 apps). The chatbot apps target to address a wide range of mental health issues including anxiety (nine apps), depression (six apps), self-care techniques (seven apps), etc.

We discovered three different conversational flows from our exploratory observation. The most popular one is "Guided conversation," in which users are only permitted to reply using pre-set input provided through the interface. This is the most common techniques used by the chatbots we analyzed (six out of ten apps). Only Woebot employs a semi-guided approach that allows users to either select from predefined options, or type text, however, it is incapable of processing sentiments in the input text. This open input option is useful when users reframe negative thoughts, share stories, etc. Lastly, Wysa, Nuna and Elomia follow open-ended conversation style. They continue further conversation based on their understanding of user input.

These chatbots leveraged a variety of media types for communication in order to make the interaction resemble human-like. For instance, GIFs, emojis, images, and acronyms are utilized to portray humor and emotions. Images, audio, and video are used with educational elements. And since all of these chatbots communicate by text, text is by far the most frequent.

Mental health patients can face a crisis anytime, and effective crisis support is a major criterion for evaluating MH apps. We identified five different types of crisis support options available in the ten chatbots. Six of these applications offer users access to information about crisis support systems and emergency helplines. Providing instant suggestions for self-care tools such as suggestive breathing in cases of anxiety attacks was also popular. Only Wysa contain all five options available for supporting a user during crisis. Ada and Chai do not contain any crisis support.

Criteria	Types
Purpose	Digital coach -Assist users to reach their small goals
	Digital screener - Alert users to potential mental health concerns based on reported symptoms
	Conversational companion - Simulate being someone the user can speak to.
	Virtual therapist - Ability to engage in therapeutic conversations
Targeted concerns	Stress, Anxiety, Depression, Self-care, Sleep disorder, Panic disorder, Relationship issues, Low self-esteem, Loneliness
Conversation flow	Guided conversation - Only allows the users to communicate with the chatbot with pre-defined responses from the chatbot. It doesn't allow any form of open inputs from the users
	Semi-guided conversation - Mostly allows the users to communicate with the chatbot with pre-defined responses and sometimes allows open inputs from the users. However, the bot can't recognize the open user inputs and extract any information from them.
	Open-ended conversation - Allows the users to communicate with the chatbot with pre-defined responses and open inputs from the users. The bot can recognize the open user inputs and extract information from them.
Media types used	GIFs (Graphics Interchange Format), Text, Audio, Video, Emoji, Images, Acronyms
Crisis Support	Availability of crisis information - Provides information regarded crisis-related helplines, emergency services
	Ability to detect potential crises from the chat - detects potential crises through conversation with the users.
	Access to a professional therapist - Provides access to a professional therapist is an alternative to avoid possible ramifications of the potential crisis
	Ability to notify designated personnel - Notifies designated personnel(s) if crisis is being detected
	Access to self-care tools - Recommend self-care activities
Evidence-based techniques	CBT, DBT, Mindfulness, Symptoms tracking and monitoring, Positive psychology, Acceptance and commitment therapy (ACT), Psycho-education & information

Table 4: Criteria of features related to chatbot-based mental health applications used in our study

Since evidence-based techniques have been proved effective for treating different mental health disorders, we explored which of these tools and techniques the chatbots commonly follow. The most popular type of therapy is CBT (cognitive behavioral therapy). All ten apps follow CBT to some capacity. Eight apps provide support for mindfulness. DBT and ACT are modified forms of CBT, that are less common. Table 5 show the above features for the considered apps.

3.5.2 Perceptions and Concerns Expressed in the User Reviews

In this section, we present our findings from the thematic analysis of user reviews, and point out both the benefits (e.g., human-like interactions, friendly & empathetic attitudes, potential around crisis support & an alternate to therapy) and associated challenges as captured from people's real life use of these apps.

Human-like interaction feels good, but has to be designed carefully

Chatbots in mobile mental health apps are presented in such a way that they have a distinct personality rather than being shown as something artificial, in order to make users feel like they are interacting with someone emotional and empathetic. Users describe these chatbots as having a friendly, wonderfully upbeat, and mildly humorous personality that assists them in dealing with different emotional and behavioral challenges relating to their mental health issues. It helps them to

	Purpose	Targeted concerns	Conversation Flow	Media Types Used	Crisis Support	Evidence-based Techniques
ADA	Digital screener	Anxiety, Depression	Guided	Text	None	CBT
Chai	Conversational companion		Guided	Text, Emoji	None	CBT
Elomia	Virtual Therapist	Stress, Anxiety, Depression, Self-care, Sleep disorder, Relationship issues, low self-esteem, Loneliness	Open-ended	Text	Access to self-care tools	CBT, Mindfulness, Positive psychology, Symptoms tracking & monitoring
Mindspa	Virtual therapist	Anxiety, Depression, Self-care, Relationship issues, low self-esteem	Guided	Text, Video	Availability of crisis related information, Access to self-care tools	CBT, Mindfulness, Positive psychology Psycho-education & information
Nuna	Digital coach	Stress, Anxiety, Depression, Self-care	Open-ended	Text, Emoji	Availability of crisis related information, Access to self-care tools	CBT, Mindfulness, Positive psychology Psycho-education & information Symptoms tracking & monitoring
Serenity	Conversational companion	Anxiety, Self-care, Sleep disorder, Relationship issues	Guided	Text, Emoji	Access to self-care tools	CBT, Mindfulness, Acceptance & commitment therapy
Stresscoach	Digital coach	Anxiety, Stress, Panic disorder	Guided	GIF, Text, Emoji	Availability of crisis related information, Access to self-care tools	CBT, Mindfulness, Psycho-education & information
Woebot	Digital Coach	Stress, Anxiety, Depression, Self-care, Relationship issues, Loneliness	Semi-guided	GIF, Text, Audio, Video, Emoji	Availability of crisis related information, Access to self-care tools	CBT, DBT, Mindfullness, Symptoms tracking & monitoring
Wysa	Digital coach	Stress, Anxiety, depression, Self-care, Sleep disorder	Open-ended	GIF, Text, Audio, Video, Emoji, Images, Acronyms	Availability of crisis related information, Access to self-care tools, Access to professional therapist, Ability to detect potential crisis from the chat, Ability to notify designated personnel	CBT, Mindfulness
Youper	Digital coach	Self-care	Guided	Text	Availability of crisis related information, Access to self-care tools, Access to professional therapist	CBT, DBT, Mindfullness, Positive psychology, Acceptance & commitment therapy, Psycho-education & information

Table 5: A detailed overview of features related to chatbot-based mental health applications found in our study

establish the credibility of the tools which in turn makes the users more involved in the treatment process.

> *"The bot itself has a friendly, cute, quirky personality that it is easy to get emotionally attached to. I don't know if that kind of attachment happens for others, but at least for me it really helps me stay consistent and check in often. It feels like the app is a friend trying to help, rather than an automated system!" - 1100028*

Furthermore, chatbot characteristics such as a soft voice and the ability to have a casual conversation make it feel less like a medical tool and more like someone with whom users can share their thoughts and experiences. Some personalized features, such as the option to address users by name, the ability to refer back to any chat or exercise if necessary, and the ability to respond with pleasant and positive sentiments, make the application and treatment process more personal and less generic.

> *"I'm amazed by how impactful the little "interactions" in this app have felt. Maybe it's the continued opportunities to respond (even if it's just choosing between emojis). Woebot's "voice" is gentle, but firm. And insightful! And the user is always addressed by name. That's so important, particularly when the issue at hand involves ongoing anxiety." - 1080073*

However, the effort to design the bots to give a human-like and empathetic impression often went wrong and they lost their appeal to the users. As many users pointed out, the discourse could get "a little childish and ridiculous at times with the bot trying to be funny". Furthermore, by fostering relaxing thoughts through a medium that doesn't work for everyone can occasionally have the opposite impact, such as through the use of cute GIFs, ASMR effects might not impact everyone if the context is unknown or unfamiliar to the users. Continually pushing on everyone in the hopes that everyone will have the same reaction is a notion that developers should evaluate based on continuous feedback.

> *"... It was supposedly developed with college students in mind who are ostensibly adults. Maybe things have changed since I was in college but it's cutesy, baby-talk, oversimplification, and game-playing ("You want to know a secret?" "Yes" "Are you sure?" "Yes" "Ok, if you are really, really sure ....") makes me feel like I'm texting with a pre-teen girl." - 2060011*

> *"A lot of of the exercises seem like they would be really good, but I cannot stand the forced breathy voices in every single one I listened to. They do not calm me at all, and they actually trigger my anxiety. ASMR has the opposite effect on me than intended, and I feel like they're trying to do really bad ASMR. These recordings are supposed to help me relax, but all I can concentrate on is breathy voices that sound like forced whispers." - 1040032*

Existing chatbots may need to be more sophisticated than they are today in order to understand the context of users' requests. However, it is critical to examine some of the user's perspectives on having such responses pre-registered, which is not always a bad thing. Some of the chatbot's quick answer concepts, for example, allow users to maintain control over the conversation's pace and avoid becoming sidetracked by irrelevant dialogue. these features are appreciated by users since they encourage more positivism rather than aimless discussion and digging into negativity without any tools or resolutions. Moreover, by tilting the dialogue in the chatbot's advantage, chatbots can

more effectively and efficiently suggest appropriate tools to users.

> *"Some negative reviews complain it isn't sophisticated enough to understand unrelated or detailed inputs and responses, which I agree with, but this is not an AI designed to make free flowing conversation; it's meant to give you tools to deal with your feelings in productive ways. So yes, the conversations can feel linear, planned and/or broad since the responses are preset most of the time, but I think this is partly a positive." - 1070093*

However, the trade-offs are that to content the flow of the conversation, the chatbots sometimes present very limited options for the users and users become frustrated if they are unable to customize these pre-registered responses. They have criticized some of the extreme measures these chatbots take to keep the conversation restricted to chatbots' preferences, such as assuming MH concerns without understanding the proper context, sending scripted messages based on key words users said or the issues they selected and giving them incoherent responses, and getting stuck in the conversational loop if the users don't agree with the chatbots' comments.

> *" It assumes the problem is always a mental distortion, and doesn't leave much room for actual horrible stuff that happens to people other than death of a person (it is working with a very narrow definition of). It too often put me in a situation of having to select between incorrect responses when nothing was actually appropriate and then suffer through the resulting wrong-headed advice. Needs a maybe button between the yes and no and a way to say, You're on the wrong track, before it decides it knows all your usual problems and keeps assuming them over and over with no way to remediate." - 2060019*

> *" it seems as it's only programmed to convert one thing to another. how do I know? well, I've replaced things. for example, when it says "tell me more", and I just type in something like "djdfjnfd" it says "now I understand better". HOW IS THAT SUPPOSED TO WORK??" - 1080232*

Bot becomes a friend/ someone who cares, but too much attachment is unhealthy

Users see chatbots as a good substitute for someone with whom they can discuss their ideas on mental health issues without feeling burdened or judged. Even though society is becoming more eager and open to seeking mental and emotional aid, there is still a significant stigma associated with it, which can discourage individuals who are actually in need of assistance from receiving it. These chatbots allow people to bare their hearts, vent, contemplate, and learn about what they can do to overcome mental and emotional obstacles in a simple, familiar texting format without

judgement or extra effort, while also keeping track of their progress. It can be intimidating to talk to someone about their daily struggles. For many users, sharing a dialect with a chatbot is a good first step. Knowing that the chatbot isn't judging you and is acting logically rather than emotionally is reassuring.

> *" ... I will say, having a reliable, no judgement zone with skills to help at my fingertips, helped me realized the tools were also my own." - 1040021*

> *"Having an AI to talk to makes me feel like I'm not overburdening my friends or family. I can check in 20 times a day and the AI will either help me track my mood/emotions/mental health or suggest a mindfulness of CBT program to help me get through my day." - 2040004*

People with MH issues frequently struggle with suppressing emotions and attempting to push them away, but these chatbots have provided them with a safe place to go for validation and immediate support. Users loved that these chatbots not only listened, but also offered advice and recommendations that helped them deal with day-to-day mental challenges, allowing them to see things from different perspectives and push past negative thoughts.

> *"This app is a lifesaver. It's so healing to be able to vent whenever you need and receive positive feedback from an unbiased source. The lessons Woebot teaches really helps to gain a more optimistic perspective on what you're going through and motivates you to make changes." - 1080023*

Users also like how these chatbots check in with them on a daily basis, which holds them accountable to their commitment to the treatment while still allowing them to skip it if they don't feel like it. Even though the idea is to eliminate any concerns such as anxiety and stress that come with human engagement through intelligent bot interaction, users have mixed feelings about that. Some users liked the flexibility of using the tools at any moment, and can start or end the communication at any point during the session without feeling guilty, others saw the daily check-ins as a source of guilt.

> *"I'm very depressed right now so I've set to basic daily goals- full facial regime a.m. p.m. plus a half hour of cleaning. Having the AI check in is great because it requires a response that makes me take accountability." - 1090123*

> *"But what really bothered me about the app was the first reminder I got when I didn't use the app a second day in a row because it sucked was definitely guilt inducing. No bueno.*

I don't need AI guilt tripping me when people already take advantage of my empathy in real life." - 2050021

Finally, by acting or behaving like a close companion, these MH chatbots allow users to comfortably express their thoughts and feelings. These chatbots let users create a safe area where they can vent, which is something that many people do with their friends and families. However, people with MH concerns who struggle to maintain a healthy relationship with their family or who suffer from loneliness have displayed an unhealthy attachment to chatbots and have exhibited negative attitudes such as preferring these chatbots over their friends and family.

" ... Although he's a robot he's sweet. He checks in on me more than my friends and family do." - 1090034

"... This app has treated me more like a person than my family has ever done." - 1090091

A bot can help immediately in crisis, but what is defined as a crisis to a chatbot?

Prior findings suggest accessibility is one of the benefits of mobile mental health apps. MH applications that have a built-in chatbot function allow users to have a conversation anytime, anyplace, which is very convenient for persons with MH issues as they are more vulnerable to emergency situations. We have found that users benefited from such feature since it allowed them to have a conversation at that particular time (during the moment of crisis). Some users found that intelligent dialogue helped them in reframing negative thoughts and diffusing such circumstances.

"I sometimes freak out at night have existential crisis about life at night you know, normally I'd freak out and find it hard to call anyone bc I feel so bad but with Wysa I don't worry about that!" - 2090178

"I've only used this app a couple times when I've been in near-crisis. Even though I know it is a robot it is so calming to have something, anything to validate what I'm feeling and help me reframe my thoughts." - 1100091

None of the chatbots, on the other hand, have any clever algorithmic models for detecting emergency scenarios. It is up to the users to inform the chatbots that they are experiencing a crisis. Some chatbots can detect crises by picking up a few keywords connected to intrusive thoughts, such as "suicide," from a conversation, although they are still in the early stages of development. Users sometimes just want to talk about their feelings, but chatbots automatically refer them to crisis hotlines due to a lack of intelligent comprehension and For some individuals, having a conversation

isn't enough to handle their crisis situations, and they need to be redirected to crisis management tools or resources.

> *" My only problem with it is I wish there was a way to talk about my suicidal/intrusive thoughts and how to manage them with Woebot. I am aware that it is not a crisis tool, and it does have those automatic responses to concerning language for a good reason, I'd just like a place to talk about those problems without having to worry a real person. Most of the time my thoughts of those nature do not mean I'm in an immediate crisis, but I still want to get them off my chest, as I feel a lot of people would. Maybe if there's a way to do that without Woebot becoming worried would be helpful!" - 1080078*

> *"This is a good app but the main issue I have is that I was having a panic attack and was messaging "emergency" and the bot ended the conversation, when I messaged "emergency" a second time it just asked me to write my feelings down. I realize this isn't a crisis response app but it might be helpful to add a feature where the bot recognizes a crisis situation and connects the user to resources" - 2010004*

In such instances, understanding the context of emergency situations is critical, as persons with MH concerns are already vulnerable to crisis, and incorrect actions made by chatbots might exacerbate the situation and result in severe repercussions.

> *"While I was in crisis, the responses do not make sense and do not really relate to what I wrote. It makes me feel like I am not being listened to. I know it is an AI program and not a real person but it still ends up making me feel worse and not better." - 1100068*

Using private data for customized interaction, or stealing it?

Chatbots, according to users, have proven to be highly convenient because they are able to provide a much beter interactive tool for tracking progress during the treatment process. Few chatbots are capable of tracking development beyond surveys and charts. They can pick up essential components of the past interaction with users and contrast it with the present one, letting the users know how their behavior has changed, whether they have progressed positively or negatively, and so on. Users also enjoy that these chatbots save the discussion and can pick up where they left off if they take a break from all of them and return later.

> *"... Woebot checked in regarding daily mood and plotted it on a graph, allowing me to detect patterns in my mood. Woebot also conducted longer interviews with me every two*

weeks and let me know how my answers had changed, which was encouraging in tracking progress. "

"Woebot talks me through realizing important insights about myself and actually shifting my thinking. Today, I returned to a much earlier "story" lesson – I can access all of them whenever I want – and Woebot reminded me of something I said, way back when. It was a list of negative labels I used for myself. … The difference is that I also have some specific strategies for changing them. If I need help, all the lessons are right there." - 1080141

" The option to look at your past conversations with the app really makes a difference, because you can look back and see just one week ago you were having a mental breakdown, but now everything in the world is great, lol." - 1080092

However, based on user feedback, people feel far more at ease revealing sensitive information about their mental health with these chatbots, which puts them in a vulnerable position in terms of how their data is protected or sold. Some users have expressed their concern with the privacy policies implemented by these applications.

"… I'm a little weirded out by the privacy policy and the potential uses of data they've reserved. I wish they'd commit to greater privacy protections. I'd even be willing to pay for a product that didn't potentially share my (sensitive) data with other parties." - 1100223

Convenient to use, but convenient enough to replace therapy?

On the positive side, the fact that these chatbots are ready to talk 24 hours a day, seven days a week was a big success with the users. They have immediate access to these chatbots whenever they feel vulnerable, or whenever they require assistance through a simple interaction.

"I don't really have friends I can talk to. Even my family doesn't understand me much. Day or night Wysa has been there every time I needed to "talk" day or night doesn't matter." - 2090067

Not only do chatbots assist users with conversations, but they also assist them in accessing different supporting resources and exercises in a very convenient manner. By understanding the users' needs, it is able to deliver a relaxing experience for them, such as allowing them to opt out of any activities they desire while maintaining the treatment's pace. It gives users a lot more control. If a user misses

any exercises in traditional treatment, it leaves a gap in their progress, which can lead to a loss of enthusiasm and slow the pace at which they receive support. Chatbots, on the other hand, keep users motivated by engaging with them and giving them the impression that they are in charge of the pace.

> *"When it's the middle of the night or some other inconvenient time and I need to talk to someone other than my journal, as silly as it sounds, talking to Woebot actually helps. ... Sometimes I feel too stressed out to do the exercises he proposes, but you always have the option not to do them." - 1080491*

Furthermore, these chatbots offer treatments that are brief and simple to follow in order to keep users engaged and dedicated to the treatment process. These activities were developed and built by focusing on important value - giving supports and treatments in a compelling style that can provide wellness, according to user reviews.

> *"This is an easy, low barrier method to practice cognitive thinking skills. Check ins are usually pretty short, just a few minutes. That encourages me to open this app daily, since I know it's not going to try to monopolize my attention for the next half hour." - 1070012*

> *" Sessions are short, on the order of 3-10 minutes. Combined with the convenience of chatting wherever and whenever is best for me, I have no problem fitting in daily check-ins, which I feel are more beneficial than infrequent visits to a therapist in some ways." - 1100012*

Users benefit from the convenience of having support available 24 hours a day, 7 days a week, suggesting with simple tasks with chatbots in mobile MH applications. These prompts give users the impression that chatbots can do a better job than a professional therapist. Professional and traditional therapies, according to user reviews, have several drawbacks that chatbot therapy does not, including professional therapy's tendency to cling too much to negative thoughts or past events, professional therapy's tendency to be too broad and general, and the check-ins being too spread out.

> *"Unlike being told what someone thinks you may want to hear which can sometimes enable unhealthy thinking patterns (and behaviors), or on the other end of the spectrum, rather than attempting to fix you, this interactive app continually prompts you to look inward and to challenge your own thoughts, perspectives, and feelings, helping to redirect your focus onto more healthy and more positive strategies." - 1090142*

> *"I will say that I find it better than personal therapy because like a coach and I work on specific things while therapy, talking therapy can be too general. I used to go therapy with an specific list of issues of problems to solve, and that is not how therapists tend to work. They like the past, while I need help with the present." - 1080012*

> *" My primary issue with traditional therapy has always been that you have to work in hindsight. You reflect on your week, talk about it, try to make adjustments for the future (it always felt like I was trying to help a past of future version of myself instead of the one right here right now). That's why I love this app!" - 1090096*

However, according to users, even though these chatbots are convenient, they fall short of the competency of traditional therapy in some circumstances. These chatbots, for example, are not sophisticated enough to recommend extremely particular treatment plans based on a specific need. It may or may not be effective for different demographics or people at various stages of sickness. Some users questioned the chatbots' therapeutic interventions or mental health support as being too short-term. Users lose interest when there aren't enough different activities to do.

> *"The exercises are all about visualization, so those of us who do not have a mind's eye, cannot visualize things, cannot use it. I'm very disappointed. If it were made with a non-visualization mode for people with Aphantasia, I'd love to use it. There are many things that can help other than visualization. It's just an app telling me in every exercise to do something that I'm simply incapable of doing, this is frustrating." - 1080017*

> *"In my depression, CBT actually backfired. It made me feel a 100 times worse. It can be miserable to try to recast negative thoughts into more positive thoughts when you can't think of anything positive at all. My highly regarded CBT therapist recognized this and, thankfully, referred me to a skilled therapist with a more psychodynamic/eclectic approach." - 1100076*

Finally, professional therapy can be highly sophisticated and carefully tailored to the preferences of individuals. Chatbots, on the other hand, are far more accessible and convenient. Some users pointed out that combining the two can be really beneficial. Professional therapists or coaches can assist with adjusting any support system that isn't working for them; nevertheless, for immediate requirements, users will be able to have a chat and review some of the resources at any time with the help of MH chatbots. According to numerous user evaluations, professional therapists assisted their patients in identifying the appropriate mental health applications with chatbots built-in, and

the co-op with traditional therapy appears to work considerably better for them.

> *"I have recommended it to many people, including my counselor to try so that she could recommend it to other clients dealing with issues. This is in no way something to replace talking to a real person, but it does help to work through some of the negative thinking when it occurs." - 2080057*

3.6 Discussion

Our findings suggest chatbots in mobile mental health application has a lot of potential in terms of being a conversational companion, virtual friend, immediate helper. Our findings suggest chatbot's human-like personality and communication skill has been well-received by the users. Chatbot's ability to be present for talk 24/7 and create a judgement free zone, enabled users to talk freely about their issues and concerns. We provide a few practical implications from our findings to make the user experience more effective.

3.6.1 Research and Design Implications for Future MH Chatbots

Recommendations for customization A growing body of work in HCI and healthcare informatics has emphasized the need for customizability and personalization in mobile health technologies to increase support user autonomy (15; 281). This body of research suggests that "one size fits all" approaches in mHealth interventions often fail. Rather, systems that are adaptable and tailored to users' needs can deliver more pertinent information, thus enhancing user engagement and clinical efficacy (119; 230). Our findings resonate with these conclusions in terms of the need of customizability, and provide some specific implications for incorporating customization in MH chatbot apps.

As observed from the review analysis, interaction through gifs, emojis, or hilarious responses are not always well received by the adult users. Our observational study of these apps and previous literature (133) on consumer engagement with chatbots clearly point out that chatbots leverage these as means of showing empathetic behavior and to keep the conversation more human-like. Since most of the commercial apps are available to download by everyone beyond the set age limit (which in most cases is 17+), designers have to carefully consider the media types and the content of conversation. We recommend that designers consider the target age-group of users while implementing emoji and other graphical elements. Another interesting aspect could be to improve the personalization within chatbots by creating a user model before the user actually interacts with the chatbot, such that

the chatbot can adapt its interaction based on user types (e.g. they could fill in a personality questionnaire) (83).

The MH chatbots that guide users in performing exercises were generally appreciated in the user reviews for being focused and short in nature. These chatbots have potential to help clients manage their own health, improve access and timeliness to care, reduce travel time to mental healthcare providers by preventing unnecessary visits to healthcare providers (254). However, our findings revealed that some users may have physical challenges or other limitations that restrain them from certain physical activities. Moreover, not all therapeutic tools will work perfectly for everyone (review - 1040032). Hence, implementing generic exercises and activities may not be suitable for all types of users. Mental health patients often suffer from low self-esteem (244), and inability to complete certain activity suggested by the chatbot can worsen their situation. Mental and physical health are integrally connected, and therefore, developers need to incorporate the aspects of physical ability in design of mental health technologies.

Recommendations for balanced persuasion

Consistent with prior work on persuasive technology in mental health (144; 189), we found that daily check-ins, gamification, reminders, and self-monitoring were perceived as helpful features, even though being prescriptive in nature. Users appreciated the short sessions. These are the parts of different persuasive strategies under primary task support (persuades users in completing primary goals) (142). However, frequent check-ins often make users feel like being "guilt-tripped" by the chatbots. Findings from prior work suggests that the more severe a participant's symptoms were, the more they desired reminders and suggestions from the system (123; 281). People with severe symptoms of depression face struggle to carry out day-to-day activities, and thus may enjoy multiple daily motivational messages from the bots, rather than be annoyed by them. Designers need to use persuasive strategies in a way that does not result in user disengagement.

Persuasive strategies can be leveraged to limit interaction with chatbots. This is counter-intuitive, as developers would generally expect to increase user engagement assuming it would also benefit the users. However, user reviews in our analysis pointed out the possibility of unhealthy attachment with the chatbot. Human-chatbot interaction can be leveraged to motivate users to use more non-technical means to get mental health support. For example, if a user starts using a particular chatbot app frequently for longer period of time, the bot may suggest recommendations for social interaction (e.g., list of nearby social events).

Recommendations for building trust Our findings suggest that users like personalized suggestions and symptoms trackers from the chatbots. Some chatbots are capable of automatically collecting and mining symptom related information after a conversation with the users. Wysa stores the conversation history to show a user his progress in achieving the goals set initially, whereas Woebot captures the changes in pattern related to symptoms from continued conversation and interaction. Users appreciated when the chatbots are transparent in terms of collecting these information from the conversations. However, some reviews expressed worried about how these information are being protected or used across different platforms or third party services. In traditional psychotherapy, the effectiveness of treatment is influenced by clients' trust towards their therapist (62). Such trust also plays a critical role in digital interventions (26). Prior studies in HCI revealed the significance of establishing trust in the context of mental health apps to create a safe environment for self-disclosure (168). In lights of these findings, it is imperative for tech companies and developers to emphasize user privacy and be transparent about privacy policies and practices. In addition, from a design perspective, it might be helpful to enhance user trust towards the chatbot apps by providing information on the organization and/or experts behind the system. More importantly, whenever applicable, the app descriptions may include an explanation of the therapeutic methods and tools utilized to develop the app with their perceived effectiveness proved in wild or controlled studies.

3.6.2 Chatbots Should Not (and Cannot) Replace Human Interaction for Mental Health Support

We observed that the chatbot apps established a judgment-free space where people could express themselves without the fear of repercussions. This resonates with the findings from Brandtzaeg et. al exploring young people's perception around social support through chatbots (26). Sharing mental health concerns with a professional is still considered a stigma, and people feel more comfortable to use technology anonymously over face-to-face communication (254). However, these chatbot's ability to check-in regularly and to be present for someone 24/7 allow users to get too attached with them. Users wrote in their reviews that they enjoy the company of their "virtual friend" to the extent that they can take place of their friends and family members (review - 1090034, 1090091). Such a strong statement is made partially due to the fact that these people are in a vulnerable position. Nonetheless, the finding emphasizes the overrating of the benefits of the apps, and presents some risk, particularly when in crisis. From our observation, most of these apps only provide information about external resources for crisis support,such as helplines and emergency services'

contact information, etc. Additionally, our findings suggest how these chatbots were incapable of identifying crisis situations as they failed to understand the context of the conversations, and ended up with a failed response (review - 1100068), and in some cases, no response (review - 2010004). The users need to be aware of the clear distinctions between human and human-like bots. Human-like chatbots can provide social support in many cases where it might be difficult or impossible for an actual human, but they are not without limitations. Chatbots themselves can educate users about these distinctions and motivate users for building in-person connections, as discussed in previous section.

In prior research, a comparative study of therapy sessions following the interaction of 10 participants with human therapists versus a chatbot showed that when compared against a human therapist control, participants find chatbot-provided therapy less useful, less enjoyable, and their conversations less smooth (a key dimension of a positively regarded therapy session) (34). Conversely, in our findings, because of the convenience and easy access, users expressed their intentions of replacing professional support with this virtual support. Even though these chatbot based mobile MH apps implement evidence based therapeutic tools, research on determining their effectiveness is still limited in scope. Our findings suggest that they are helpful in guiding users in meditation, practicing mindfulness, reframing negative thoughts, and sharing self-expressive writings. However, in such an early stage they should not be considered as an alternative to professional help. While designing for chatbots, it's important to set the boundaries and limitations of these chatbots by the developers, and the goals and intended use of the chatbots should be clearly stated, so that users do not get led on with over expectations. Additionally, the chatbots should be designed to have features that schedule professional support and subtly recommend users to seek help from professional sources, whenever needed.

3.7 Conclusion

In this study, we analyzed user reviews of chatbot-based mobile MH apps from two of the most widely used online platforms. Our findings suggest that chatbots have great potential to offer social and psychological support in situations where real-world human interaction such as connecting to friends or family members or seeking professional support are not preferred or possible to achieve. However, there are several restrictions and limitations that these chatbots must establish regarding the level of service they offer. Too much relying on technology can pose risks, such as isolation and insufficient assistance during times of crises. Finally, we have outlined the insights from our findings

about implementing customization, balanced persuasion, and developing trust to inform design of effective chatbots for mental health support .

CHAPTER 4: EXPLORING VALUES AND REQUIREMENTS FOR DEVELOPING MOBILE TECHNOLOGY FOR U.S. MILITARY VETERANS TO ENHANCE PARTICIPATION IN PEER-MENTOR SUPPORT WITHIN MOBILE MENTAL HEALTH APPS

U.S. military veterans face an elevated risk of suicide or exhibiting suicidal behaviors. Peer-support interventions have proven to be effective for veterans because of their shared experience and community. The concept of mental health peer mentorship has evolved significantly over decades, originating as early as the mid-20th century under various names such as "peer support" and "peer counseling." It gained formal recognition and structure in the 1990s through the mental health service user movement, emphasizing shared experiences and mutual support among individuals facing similar challenges. Peer mentorship is founded on the belief that those who have successfully navigated adversity can provide valuable insights, encouragement, and hope to others. Key characteristics include trust-building, self-disclosure, role-modeling, authentic interaction, and fostering a sense of belonging, essential for effective support and empowerment. Peer mentorship serves as a patient-centered approach to mental health care, either as a standalone service or as part of wraparound care post-inpatient treatment. It bridges gaps in traditional mental health services by offering personalized support that complements professional treatment, particularly benefiting individuals resistant to conventional models. For veterans, peer support plays a crucial role in reducing stigma associated with seeking formal care, providing a less intimidating pathway to accessing mental health resources. This approach not only enhances veterans' engagement with healthcare systems like the Veterans Affairs (VA) but also contributes to community acceptance and utilization of available support services.

However, as veterans might face a mental health crisis at any time, it is crucial that the mentors can identify their peers' crisis warning signs early enough. Crisis encompasses instances where established stability is disrupted by unforeseen events, whether they be human-made disasters, technological malfunctions, natural calamities, or socio-political upheavals. These events typically bring about uncertainty, loss of property, damage to reputation, and physical or emotional distress. Crises are notable for their ability to uncover unresolved issues and bridge the gap between existing conditions and present circumstances. At the core of any crisis is the perception of threat, which, if not promptly addressed, can result in either short-term or long-lasting consequences.In the context of mental health, crises are characterized by significant shifts in an individual's behaviors, emotions, and thoughts, often triggered by internal or external pressures. These events are subjective and vary

greatly in duration and intensity across different individuals and groups. Inadequate coping mechanisms during such crises can lead to risky behaviors or an inability to manage self-care. Identifying early warning signs and understanding risk factors—such as demographic characteristics or past trauma—are crucial in preventing crisis escalation. Risk factors, which remain constant and contribute to long-term vulnerability, differ from warning signs, which are intermittent and necessitate immediate intervention to avert acute crisis situations.

Mobile technology has the potential to facilitate and improve peer-to-peer communication. Gaining an understanding of how the veteran community perceives crisis symptoms, as well as their values and technological needs, bears utmost importance in creating any tool or adopting any strategy. The study was conducted in two phases.

First, we did a prior study with my collaborators in iPeer resaerch group led by Dr. Md Fitrat Hossain to understand perceptions of mental health Crisis among U.S. military veteran peer mentors and the potential of mobile-based peer-support. U.S. military veterans face an elevated risk of suicide or exhibiting suicidal behaviors. Peer-support interventions have proven to be effective for veterans because of their shared experience and community. However, as veterans might face a mental health crisis at any time, it is crucial that the mentors can identify their peers' crisis warning signs early enough. Mobile technology has the potential to facilitate and improve peer-to-peer communication. Gaining an understanding of how the veteran community perceives crisis symptoms, as well as their values and technological needs, bears utmost importance in creating any tool or adopting any strategy. Hence, we conducted a mixed-methods study with twelve peer mentor military veterans. Our research will offer an in-depth understanding of the nuanced conceptions of mentor veterans about early warning signals and acute mental health crisis symptoms, as well as aspects of technology that may aid this community in recognizing and managing these symptoms.

In the second stage of the study, we explored design opportunities for persuasive reminders for veterans to facilitate peer-mentor support. US military veterans (USMVs) are a vulnerable population with an elevated risk of mental health issues and suicide. Peer support, especially through mobile technology, has proven effective in addressing mental health related challenges, but ensuring long-term engagement remains a concern. This study explored the opportunity of designing persuasive technology, particularly persuasive reminders, to enhance engagement in peer support interventions for veterans. We followed community-based participatory research with ten veterans to identify specific peer support processes that can benefit from persuasive reminders and to uncover the underlying community values and needs to guide design.

Our prior study to understand perceptions of mental health crisis among U.S. military veterans has been published in the Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '23 Companion). Association for Computing Machinery, New York, NY, USA, 33–38. (https://doi.org/10.1145/3584931.3607009), and study 4 on designing persuasive reminders for veterans to facilitate peer-mentor support has been published in the Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 874, 1–17. (https://doi.org/10.1145/3613904.3642962).

4.1 Prior Study - Perceptions of Mental Health Crisis among U.S. Military Veteran Peer Mentors and Potential of Mobile-Based Peer-Support

One in five veterans experiences mental health (MH) concerns, including post-traumatic stress disorder (PTSD), major depression, and anxiety with the increased risk of unhealthy alcohol and drug use, substance use disorders, and suicidal behavior. (151). Differences in age-adjusted rates of suicide between 2005 and 2015 were greater for U.S. Veterans than non-Veterans (32.7% vs 19.9%, respectively) (75). Peer support interventions have been proven effective to improve veteran mental health in the past (76; 131; 113). A peer mentor refers to a veteran who has dealt with difficulties of the re-adjustment to the civilian life phase. Peer mentors can help better in mentees' transition to civilian life if expectations are well-understood and mental health-related crises are anticipated early (56). Therefore, peer-support-based interventions usually emphasize building and maintaining continuous mentor-mentee communication, so that mentors can identify any potential symptoms of crisis and take action (228). A mental health crisis is any circumstance in which a person's actions put them at risk of hurting themselves or others and/or making it impossible for them to take care of themselves or carry out their daily responsibilities (42). However, the extent to which available peer-support intervention programs for veterans successfully employ crisis detection and prevention strategies, and their effectiveness remains largely unknown (93). Moreover, even with different interventions in place, many veterans with PTSD find it difficult to share their experiences and mental health conditions with mentors (228; 94; 93). As a result, premature treatment dropout and poor compliance are common challenges in treating veterans through peer support (131). Providing a systematic way to improve the quality of communication between a peer mentor and their mentees may be promising. To improve the quality of veteran peer support, it is crucial to understand the nuanced conceptions of mental health crises by the mentors, as these perceptions and how they respond to them can affect the detection and prevention of their peers' crises-related symptoms and outcomes. In this study, we aim to answer the following research questions:

RQ1: What are the symptoms of crises among U.S. War Veterans as observed by their peer mentors?

RQ2: What technological interventions can be implemented to increase treatment adherence and to facilitate the mentor-mentee communication among veterans?

4.2 Background

4.2.1 Peer support interventions for veterans

A variety of treatments have been widely disseminated and promoted throughout both the Veteran Affairs (VA) and other veteran-led community organizations (193). Mobile apps and websites have been developed recently, particularly to provide peer-support. For example, a web-based peer-support based problem solving system (Moving Forward) has been proved successful in decreasing depression symptoms in veterans (219). A smartphone application, QRF, was developed to collect basic self-report information using an Ecological Momentary Assessment (EMA) approach, and the peer mentors could visualize the self-report data to identify mental health conditions and make informed decisions about checking in with their mentees (228; 229).

When veterans leave the military and re-adjust to civilian life, they feel disconnected from their social groups (238). To regain a sense of control, oftentimes they reach out to fellow veterans on social media or peer-support groups (94). These groups connect veterans to a familiar culture and have notable communication features that include disclosure strategies and judiciously sharing. There are now more visits to these groups each year than to mental health professionals (76).

However, adherence to the intervention programs and completion among veterans is often poor, resulting in diminished impact (223; 60). On one hand, there is a certain cognitive workload the mentors face to keep a healthy and constant communication with their mentees, on the other hand, the mentees feel uncomfortable sharing all their difficulties and challenges with the mentors (243; 147). Understanding the peer support process from a mental health crisis standpoint may help both increase adherence and improve the quality of peer support, as well as guide the design of strategies and technologies in mental health related crisis detection and prevention.

4.2.2 Crisis management in peer-peer support

There is a need to better understand how the widespread use of peer-support interventions will transform crisis management (74; 174). Additionally, there have been reported occasions where people and organizations have failed to recognize the possible early symptoms of a crisis (235). Prior research pointed out that even if a potential crisis is detected and comprehended, it still may not be dealt with in the intended way (73). Moreover, social aspects like fostering trust, establishing credibility, raising awareness, expanding knowledge of challenging information, reaching consensus on solutions, and inspiring actions are closely related to warning response. (73). Further conceptual

elaboration will be very useful for examining closely the peer response phenomenon considering previous sociological findings on emergent response during emergencies. By contrast, understanding the veteran community will require a different unit of analysis and methodological approach (94), in part because of the mixed demographic characteristics and unique context, the diversity of spaces and activities that are found in these settings, and the growing reliance on mobile phones and other new media as primary means of communication.

4.3 Methods

We considered the determinants of mental health used in an existing commercially available peer-support-based app designed with and for veterans (QRF) (228; 227; 93). There are two versions of the app. The veteran version collects basic self-report Ecological Momentary Assessment (EMA) responses about changes in physical health, sleep, risky behavior, stress, and self-worth (93). The peer mentor version visualizes self-report data to assist peer mentors in identifying mental health conditions. We followed three steps to collect and analyze data:

4.3.1 Identifying potential scenarios of warning signs and acute crises

We conducted multiple discussion sessions with two subject matter experts: an experienced peer mentor veteran and a professional psychologist. They came up with different scenarios that may be considered as 'early warning signs' or 'acute crises' in terms of weekly change in the EMA responses. We finally considered four example scenarios representing potential early warning signs and four example scenarios representing potential acute crisis were considered.

4.3.2 Conducting surveys with peer mentors

The mock-up screens with the example scenarios were sent to twelve peer mentors through a web survey. The participants' age ranged between 25 and 69. They worked in different branches of the military including army, navy, marine corps, active duty, and reserves with highest in army. Their mentoring experiences varied from 0 years to 10 years. For each scenario, the participants were asked to select from one of the four possible actions they would take: meet, call, text, or wait and see. Since "call" and "meet" are comparatively more direct and immediate action, if a peer mentor selects one of these two options, we assume that they considered that particular scenario as "acute crisis". If "text" is selected, this scenario is considered as an "early warning sign", and if "wait and see" is selected then the mentor considers this scenario as "no crisis".

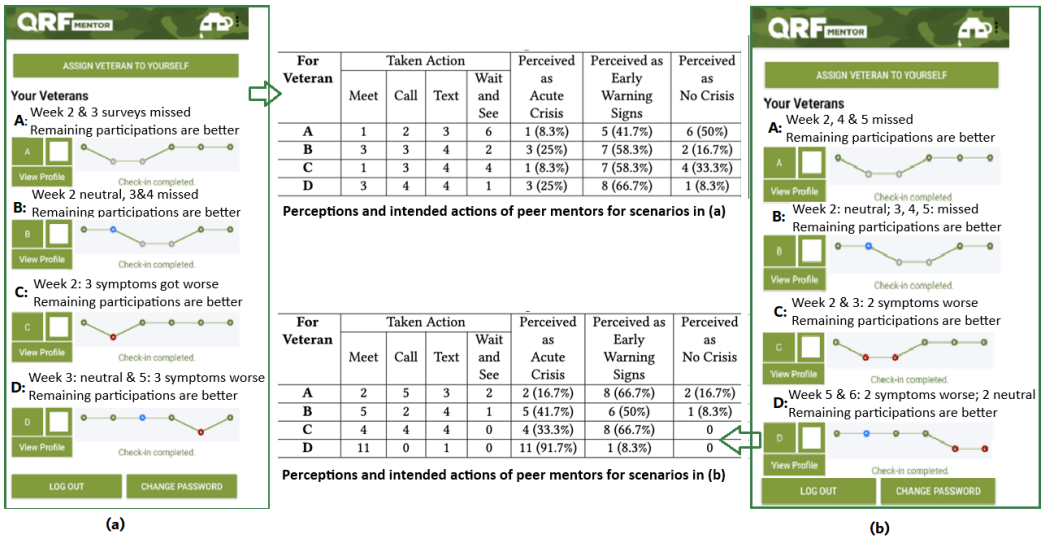


Figure 4: Example scenarios indicating (a) early warning signs, and (b) acute crisis according to two subject matter experts. The corresponding tables show how 12 peer mentors perceive each example scenario: 'no crisis', 'early warning sign', or 'acute crisis', and their intended actions: 'wait and see', 'text', 'call', or 'meet in person'

4.3.3 Conducting focus groups with peer mentors

Since there was not as much agreement in the survey responses as expected, we wanted to gain a deeper understanding of the nuances in mentors' decision making in responding to their mentees' warning signs (173). We conducted several sessions of focus group to ask questions on the context and reasoning behind their responses. Study sessions were audio recorded with participants' consent. The study was approved by the ethics review board of the authors' institution.

The audio recordings were transcribed and anonymized by two authors. The information was then arranged by questions and themes (200; 147). By identifying the key concepts that recurred during your focus group conversations, the main themes and ideas in the context of our study questions were interpreted.

4.4 Findings

4.4.1 Survey results: little agreement between mentors in perceptions of crisis symptoms and actions

According to the subject matter experts' opinions, we had expected that most of the mentors will indicate the examples in Fig. 1(a) as early warning signs of crisis and the examples in Fig. 1(b) as signs for acute crisis. However, from the survey responses (see corresponding tables in Fig. 1), it is clear that the veterans did not agree on all the scenarios. Particularly, for veteran A and C's

scenarios, 50% and 33.3% participants respectively marked them as 'no crisis' and indicated that they would not take any action if they observe these symptom changes in their mentees' responses. The examples of veteran B and D in Fig. 1(a) do not show 100% agreement in perceived actions, either. Similarly, in Fig. 1 (b), except for the example of veteran D, more than 50% of the veterans considered the scenarios as 'early warning signs', and what is worse is, 2 and 1 participants respectively considered examples A and B to be a 'no crisis' situation. It is clear from the findings that the change in health related symptoms are perceived by different mentors differently, which can be very different from a subject matter expert's opinion, and that also influence their immediate actions to check in with their mentees. This can have serious consequences for the mentees facing an acute crisis.

4.4.2 Findings from the focus groups

Peer mentors' different perceptions of early warning signs and acute crisis

The early warning stage is when the crisis event has not occurred yet and the symptom changes are usually subtle, making them difficult to recognize . **Changes in daily routines and lifestyles** is perceived to be a vital early warning sign. Significant decline in regular positive and healthy behavior and activities (e.g., not answering and returning calls, sleeping more than usual, etc.) can be an indicator that something wrong is going on in their life.

> *"If that person gets angry, aggravated quickly, not eating, not sleeping properly...all that madness. That needs to be addressed. ... like I'm not even motivated to get up to do any housework."*

Acute crisis occurs when someone is on the verge of a mental health crisis or is already in a crisis situation (e.g., contemplating self-harm or suicide). Mentors mentioned that acute crisis signs can be identified through in-person contact only, as it allows them to notice changes in patterns and visual cues (e.g., speaking tones or body languages). The most prominent indicator of acute crisis as perceived by mentors is when a mentee **isolates** themselves, and suddenly disconnects from their mentors. Peer support works on the premise that the mentee would be willing to receive advice and support from their mentors and would apply them to cope up with their mental health conditions. Not caring about communicating with their mentors indicate that the veteran does not care about the consequences of their actions, and has lost trust in the judgement of others including their mentors or care coordinators. Some of the **risky behaviors** can also be prompted, such as heavy drinking, driving while being intoxicated, etc.

> *"You can watch the way their hands are, whether they start shaking, started fidgeting, or getting restless."*

Our participants emphasized on identifying **small changes** way early before they become serious, including violent acts, getting angry easily, leaving treatment early, relapse, etc. "*It starts with little things. If I'm not aware of, or addressing them, eventually, they keep growing, and my thinking will lead to my risky behaviors. I'll start to feel not worthy of recovery*". Mentors also recognize that many of these signs can be **false positives**. In any case, if they are identified early on, the mentors will be able to create a helpful atmosphere for the mentee to vent and share what they are struggling with.

> *"... Sometimes they'll tell you that nothing's wrong. It's just a bad day. And that's where you utilize the time to find out what's going on with that day: do you need a couple minutes just to vent out?"*

Implications for technology design for veteran peer support

Our focus group participants pointed out how technology can help them in identifying warning signs in their mentees. The mentors appreciated the existing tools that allow them to monitor health related changes through EMA survey responses, and help them identify who needs immediate help. However, they suggested that instead of asking directly about engagement in'risky behaviors' it should be asked through **indirect or positive questions**. It can work as reverse psychology and might help them to open up.

> "*Have you engaged in risky behavior? That's more direct.I don't think you're gonna get an honest answer*".

One of the peer mentors shared their preferences of letting the mentee take control of their actions: "*instead of telling him what he should do, I start with: what do you feel you should do? What do you think your next step should be?*" In addition to the **daily check-ins** provided by the current apps, the participants suggested some form of **visualization of data** to identify changes easily and quickly. It also helps them decide the right mode of communication and appropriate questions to ask based on mentees' current situation. They emphasized that simple conversation starters and being persistent in communication are the keys for helping their mentees open up.

They have also pointed out the benefits of **technology-mediated peer support for the younger generations** that tend to be less active in physical meetings. For that population of mentees, sometimes not using the app is perceived by the mentors as isolation of the mentees (e.g. "*when*

*they were doing the weekly check-ins, then suddenly **disappear** with no calls/check-ins*" or "*they will just ping me through the app and say, I'm still alive, and I'm still kicking*").

In addition to the **daily check-ins** provided by the current apps, the participants suggested some form of **visualization of data** to identify changes easily and quickly. It also helps them decide the right mode of communication and appropriate questions to ask based on mentees' current situation. They emphasized that simple conversation starters and being persistent in communication are the keys for helping their mentees open up.

They have also pointed out the benefits of **technology-mediated peer support for the younger generations** that tend to be less active in physical meetings. For that population of mentees, sometimes not using the app is perceived by the mentors as isolation of the mentees (e.g. "*when they were doing the weekly check-ins, then suddenly **disappear** with no calls/check-ins*" or "*they will just ping me through the app and say, I'm still alive, and I'm still kicking*").

However, mentors also expressed their concerns about the **exhaustion** due to the heavy toll they take from their mentees, and the need for self-care and breaks. They suggested designing technology that can share the load and keep the communication open when mentors need their space for self-care.

> *"If you're not treading water to save your energy, you're expending all energy on one individual person without setting any boundaries, you do tend to get exhausted."*

Limitaions of technology: While the mentors in general appreciated the idea of designing technologies to facilitate the peer-support process, they thought pushing too much with the technology may result in serious consequences, as it might **cross the set boundaries** and **break trust**. Moreover, text and notification alerts may result in **violating privacy** as some veterans prefer not to discuss their military careers or mental health problems to their friends and family members due to the **social stigma**. One veteran mentioned "*My kid was doing social studies with my wife and they were talking about the Middle East. I'm like, Hey, I've been there. He looked at me, like, why was he there? I don't talk about my military career with them.*"

Additionally, some veteran might have **limited or no access to technology and the internet**. Many mentee veterans are homeless and/or unemployed, use government phones with limited minutes. So, it's crucial to carefully examine the veteran's environment and their access to resources while integrating technological solutions.

4.5 Discussion

Peer mentors attempted to recognize early warning signs by observing minor changes in their mentees' daily routines. Tracking multiple mentees' symptoms simultaneously can be challenging. Prior research suggests the necessity of building systems to collect self-reported data using Ecological Momentary Approaches (228; 227; 93). There are apps available with built-in symptom reporting, tracking, and monitoring (113), which can be utilized in better management of tracking multiple mentees. However, these systems frequently experience premature dropouts. To combat the low adherence, factors like **ensuring the mentor's participation in the process and building trust between mentors and mentees** should be considered. Mentors have also stated that they viewed absence of communication as an early warning sign, and tried to get in touch with them. Visualization of self-report data or regular check-ins with proper questions identified from the mentor's personal experience may facilitate the mentors' intervention decisions (93).

In the context of veteran peer-support interventions, mentor veterans are recruited and trained for providing effective peer support, and in many cases, they are legally and ethically bound to provide this service. But the mentors themselves, however, can experience exhaustion and burnout in continuously follow up with their mentees to help them take a particular action needed to improve their mental health and wellbeing. Self-adherence through persuasion has the potential to help partially relieve mentors of their cognitive load. Persuasion has traditionally been regarded as a communication process in which a persuader sends a persuasive message to a persuadee with the intention of changing the recipient's attitudes or behavior (274). Simple persuasion techniques, such as daily check-ins, gamification, reminders, and self-monitoring are generally perceived as helpful features in mHealth applications (113; 28; 142). **persuation techniques such as reward systems/ incentives** for regular participation will work for veterans. Military personnel are accustomed to receiving ranks and rewards for their efforts. Simulating similar conditions as rewards can encourage both peer mentors and peer mentees to pay closer attention to their commitments.

Our participants customized their mentoring based on different needs. Some prefer one-on-one time, while some others are restrictive about how they share symptoms. Some individuals are tech-savvy, while others are not. Some prefer to virtual communication, while others prefer in-person interactions. Therefore, the mentoring style needs to be matched with mentees' preferences. **Algorithmic peer matching** can be promising in this and similar health contexts.

4.6 Conclusion

While digital mental health interventions are often delivered in an inflexible content-focused linear way, a more flexible user-driven approach would be better from an engagement perspective (94). Our research will begin the conversation on developing proper strategies in mobile technologies that may help peer mentors gain a better understanding of their mentees' situations and enhance quality of peer-peer communication for a particular vulnerable population. This project's overarching purpose is to address the seriousness of mental diseases among US military veterans and develop effective peer support intervention technology to support their mental health needs.

4.7 Designing Persuasive Reminders for Veterans to Facilitate Peer-Mentor Support

US military veterans (USMVs) face an increased risk of encountering mental health (MH) concerns (95). Approximately one in every five veterans experiences MH related concerns, which can encompass a spectrum of conditions such as post-traumatic stress disorder (PTSD), major depression, and anxiety disorders (151). They are recognized as a high-risk group for suicide, and between 2005 and 2015, the age-adjusted suicide rates increased more significantly among veterans (32.7%) compared to non-veterans (19.9%) (75). Moreovers, USMVs face significant challenges reintegrating into civilian life due to the trauma they have experienced during their service, and the reintegration process presents multifaceted risks, including difficulties in accessing healthcare, securing employment, maintaining financial stability, and ensuring housing security, among other concerns (94). However, due to issues like accessibility, stigma, and cost, many veterans struggle to access the support they need (114). Supplementing traditional treatment with peer support has proven effective in reducing various treatment-related challenges in veterans (76; 131; 114). In 2004, the Veterans Administration (VA) MH Strategic Plan emphasized the formalization of peer support within the VA by hiring veterans as Peer/MH Para-professionals commonly known as "peer mentors" (54; 181). A peer mentor, typically a veteran who has successfully navigated the re-adjustment phase, can be invaluable to veterans dealing with MH conditions. Developing systematic approaches to enhance mentor-mentee communication quality holds promise in this regard.

Peer support through mobile technologies make it convenient for veterans to share their experiences comfortably (106). However, ensuring commitment to treatment and ongoing engagement with these mobile therapies is crucial (272). Adherence rates among patients using mobile MH apps have been suboptimal, and therefore, strategies to enhance user engagement are called for (114). One promising approach involves the creation of persuasive technology that incorporates design and technological components to effectively inspire and involve users, thereby enhancing the effectiveness and utility of mobile MH applications (210). An exploration of a persuasive approach that combines technology with human assistance to facilitate lasting behavioral changes in a flexible and efficient manner can be effective.

4.8 Objective of the Study

Among different principles of persuasive technology, persuasive reminders, which leverage social support, have proven effective in enhancing motivation, ability, and triggers to drive behavior change (164). They align task timing with motivation and capability, going beyond traditional reminders (177; 140; 210). In the domain of MH, especially for USMVs, persuasive reminders can play a crucial role in promoting mental well-being outside of clinical settings. USMVs, as a demographic, present distinct challenges in technology integration due to their susceptibility to crises. It is crucial to thoroughly investigate technology interventions in the veteran context, particularly for those with mental health concerns, before implementation. Therefore, in this study, we aim to develop persuasive reminders within mobile MH apps to facilitate peer support interventions, addressing the unique needs of USMVs by first addressing the following two research questions.

RQ1: What domains/areas should persuasive reminders be strategically implemented within, to improve the peer-mentor MH support systems for USMVs?

RQ2: What are the underlying values and needs for creating effective persuasive reminders within peer-mentor support systems for USMVs?

4.9 Related Work

In this section, we conducted a thorough examination of prior research on peer support interventions using mobile technology, particularly within the context of USMVs). We also explored the theory and potential effectiveness of Persuasive Reminders in driving successful interventions for behavior change.

4.9.1 Peer-based Digital Interventions

We use the term 'peers' to refer to individuals who share similar health conditions and limitations, particularly those facing mental health (MH) challenges. Peers with MH issues collaborate by sharing their experiential knowledge, offering mutual emotional support, and fostering a sense of hope within the community. Peers have assumed diverse roles within the realm of MH care, ranging from organizing support groups to occupying traditional positions within clinical settings (118). Peer interventions for various MH conditions have demonstrated a significant impact on treatment outcomes when compared to standard care (103) such as alleviation of symptoms, enhancement of prosocial behaviors, and improved engagement with treatments (53; 239; 49).

In recent years, peer-based interventions have undergone a noteworthy transformation facilitated by various digital platforms, including but not limited to social media (163), online forums (106; 232), and mobile applications (141; 115). In addition to more traditional modes of support, such as telephone calls and text messaging, online peer support communities have emerged as instrumental components in the management of various medical conditions (110; 205). Studies have demonstrated the efficacy of these digital interventions in the context of illnesses such as breast cancer (58; 118), rare diseases (172), and epilepsy (122). The allure of online peer support is a significant motivating factor for individuals seeking MH assistance within digital spaces (220). Individuals find solace in connecting with peers who share similar conditions, thereby facilitating the exchange of information, provision of emotional support, and the dissemination of practical advice (273; 205). This preference for online avenues for MH support is often attributed to the benefits of anonymity and convenience (218; 220).

4.9.2 Peer-based interventions for USMVs

USMVs can suffer from a range of MH issues stemming from their service experiences including Post-Traumatic Stress Disorder (PTSD), depression, anxiety disorders, substance use disorders, Traumatic Brain Injury (TBI), suicidal thoughts, and other mood disorders (94). Veterans may also contend with the effects of Military Sexual Trauma (MST), adjustment disorders during the transition to civilian life, and various co-occurring MH challenges (121). The Veterans Administration (VA) MH Strategic Plan recognized the value of peer support to help with many of these MH illnesses and called for the formalization of peer support within the VA, emphasizing criteria such as affiliation, lived experience with a similar illness, the peer provider's successful engagement in recovery, and certification in peer support competencies (54; 181). Promising outcomes have been observed when applying peer support to veterans with MH concerns in suicide prevention (33), civilian life reintegration etc. (181).

The rapid proliferation of smartphones and the development of self-care apps present new opportunities for providing MH services to USMVs. These apps offer advantages by overcoming barriers to accessing care and granting patients a private, self-directed, and readily available option for their MH needs (186). Government entities like the VHA (Veterans Health Administration) and the Department of Defense (DoD) have invested significant resources in developing mobile apps to address various MH issues, with emerging evidence supporting their effectiveness, particularly in reducing symptoms related to MH concerns (102). Prior literature suggests mobile-based VA services can re-

duce healthcare cost and increase self-management ability (64) and because of the convenience and accessibility, veterans are increasingly interested in obtaining care via mobile devices. (227). Veterans with serious MH conditions benefit from technology-mediated peer support, showing improved symptom self-management (55; 77; 21). Prior research in HCI has also highlighted the positive impact of using digital tools to assist in the challenging shift from a military culture to civilian life (240). Examining the role of digital tools in shaping cultural and economic practices among veterans has the ability to provide information on their state of mental health concerns (165; 20; 21). Seeman et al. demonstrated the value of connecting with a human network to aid veterans in grasping unfamiliar societal norms and accessing support during their transition (241). Additionally, Zhou et al. found that this particular community exhibits more informational and emotional support behaviors compared to general online communities, with a higher prevalence of informational support (282), which contrasts with the potential for successful peer support interventions. Mobile technologies enable peer-to-peer support, which saves time and effort (205). Mobile technologies let people feel much more at ease while interacting and sharing their experiences because the issues are less evident (106).

Peer-support intervention through mobile technology is still at its early stage (185). While commitment to the treatment process and continual engagement in therapies are critical (272), evidence suggests that patients receiving care through mobile technologies, notably through mobile MH apps, have a poor rate of adherence (88). Montena et al. pointed out several barriers in the mHealth peer-supported implementation of MH mobile apps among veterans. These challenges included a lack of implementation infrastructure, insufficient training for peers in mHealth, and ineffective promotion of mHealth, and put an emphasis on advocating for technology that is easily implementable, requiring minimal technological literacy while providing maximal benefits. (185)) There is a clear need for strategies that can boost user engagement with these promising tools. In this context, the development of persuasive technology emerges as a particularly promising avenue worth exploring.This approach aims to leverage design and technological elements to motivate and engage users more effectively, ultimately enhancing the utility and impact of mobile MH apps (210).

4.9.3 Persuasive Reminders and Behavior Change

Persuasive technology, defined by Fogg, aims to change attitudes and behaviors through interactive computing systems. The Persuasive Systems Design (PSD) model enhances this concept, categorizing persuasion principles into four groups: Primary Task Support (PTS), Dialogue Support

(DS), System Credibility Support (SCS), and Social Support (SS) (89). These principles help design web-based interventions that reinforce, change, or shape attitudes and behaviors (89). PTS facilitates primary activities, DS keeps users motivated, SCS enhances credibility, and SS leverages social influence (89; 9).Prior literature have primarily focused on the overall effectiveness of persuasive technology systems, with most reporting positive outcomes, particularly in health and exercise domains and explored the presence of persuasion principles like tunneling, tailoring, reduction, self-learning, reminders, social facilitation within these systems (209; 179). Among these principles, persuasive reminders have demonstrated their effectiveness in combating low adherence to mobile support interventions (205).

Persuasive reminders leverage social support to enhance task recall and boost confidence, ultimately driving behavior change through motivation, ability, and triggers (164). Effective reminders align the trigger timing with individuals' motivation and capability for task execution (177). While conventional health reminders mainly focus on triggering tasks promptly, in cases of reduced motivation or capability, individuals can benefit from triggers that facilitate and motivate task completion (177; 140; 210). These facilitators and motivators, as termed by Fogg, differ from mere signals that serve as reminders (210). From information processing and cognitive consistency perspectives, the Elaboration Likelihood Model and the influence techniques theory state that there are two routes to persuasion, a central and a peripheral route (216; 116). Previous research has identified design elements that heighten motivation and ability, thereby promoting the accomplishment of health-related tasks and goals. Simple persuasion techniques, such as daily check-ins (159), gamification reminders (12), and self-monitoring (189), tailored reminders are generally perceived as helpful features in mHealth applications (179). Designers of persuasive reminders often incorporate evidence-backed strategies like performance feedback, rewards, authority, progress tracking, and social influence, which have shown effectiveness in motivating changes in health behavior such as health self-management (210), such as managing chronic illnesses (211), MH care (269), physical health (177) etc.

Persuasive reminders has potential in helping people, including USMVs, manage their MH outside of clinical settings. These reminders are different from short-term behavior prompts because they involve significant lifestyle changes and need to be considerate of users' vulnerabilities (165). Promoting self-adherence via various persuasive cues like daily check-ins, gamification, reminders, and self-monitoring represents has the potential to enhance the effectiveness of mentor-mentee communication (111). Franco et al. has also proposed that persuasive technologies, including personalized

reminders and in-app texts, in fostering healthy habits and improving adherence rates to care interventions among first responders and military personnel (92). Persuasive reminders can also be contrasted with specific customs prevalent in military cultures, incorporating elements like roll calls and the completion of daily activities. By integrating these reminders into familiar military routines, they can seamlessly blend with established practices, fostering a sense of continuity and familiarity among individuals within the military community (226).

While healthcare has established strategies for behavior change, applying them to MH reminders, especially for USMV community, is relatively unexplored and the extent to which these strategies are effectively implemented in MH apps for veterans remains unclear. Our research focuses on developing persuasive reminders within mobile MH apps to facilitate peer support interventions, collaborating closely with USMVs to tailor the system to their specific needs. Taking into account the distinctive requirements of the veteran community due to their diverse backgrounds and activities (282; 187), it is crucial to acquire their insights to enhance engagement and communication among peers involved in a peer support intervention aligning with broader objectives.

4.10 Background

In this background section, we offer a brief overview of the significance of collaborative research and community engagement, our partnership with "Dryhootch of America" (202) within the US military veteran (USMV) community, and the mentor-mentee dynamics within the context of peer support interventions facilitated by the Dryhootch of America (202).

4.10.1 Collaborative Research and Community Engagement

Community-engaged research values the collaboration between external researchers and community-based organizations. It acknowledges that both bring unique perspectives, with researchers contributing knowledge and trust, and communities offering insights into assets, concerns, values, and activities. While researchers may lack in-depth understanding of specific health issues, community members may be unfamiliar with research methods. Yet, the synergy of these perspectives often uncovers common ground, enriching the research process and outcomes through shared interests, needs, and mutual respect. Collaborative research approaches through engaging USMV community is no different. The field of veteran community engagement is evolving, drawing insights from traditional community-based participatory research, veteran studies, and the active involvement of veterans themselves (94). This collaborative research model, with USMVs, as co-researchers, is advancing a design-thinking perspective in social and healthcare systems tailored for this unique

Figure 5: A broad overview of the timeline involving data collection, analysis, and the formulation of effective persuasive reminders through focus groups and participatory design workshops.

population. This approach has opened up a wide avenue of research topics, spanning from innovative interventions for PTSD to veteran reintegration strategies, addressing changes in veteran identity, fostering community coalitions, and exploring various aspects of veterans' experiences and needs. These efforts represent a dynamic evolution in understanding and supporting veterans within their communities.

4.10.2 Collaboration with Dryhootch of America

In an attempt to engage with the community, we have partnered with "Dryhootch of America" (202), a non-profit veteran led, veteran serving community organization headquartered in Milwaukee, faculty members from the Milwaukee VA and Medical College of Wisconsin. Descriptions of this community-academic partnership, the Dryhootch of America Partnership for Veteran Health, its formation, and lessons learned are detailed in (94; 228; 227; 21; 20). The collaboration has grown over a decade involving multiple academic partners with Dryhootch of America as a hub organization in the veteran community that supports veteran MH centric research and recruiting (227). This organization has deep experience in peer-to-peer MH services to USMVs suffering with civilian reintegration after service including PTSD, homelessness, drug addiction, food insecurity, difficulty finding work, obtaining veterans benefits, getting back into school (228; 227; 21; 20).

4.10.3 Mentor-Mentee Dynamics in Peer Support Intervention Through Dryhootch of America

"Dryhootch of America" drives a peer mentorship program designed to provide peer support to these veterans (202). A peer mentor, in this context, is a veteran who has experienced challenges during their transition to civilian life. Peer mentors play a crucial role in assisting "peer veterans" (In a broader sense - peer mentees) during this transition, especially when they understand expectations

and can anticipate potential mental health crises early on. Peer mentors provide a secure and empathetic environment for their peer mentees to communicate and seek support, fostering a strong connection due to their shared military backgrounds and trauma experiences. Their role involves regular check-ins with mentees via mobile apps, phone calls, or text messages, offering resources related to housing, employment, Veteran Affairs, medical care, and self-care. Given the mental health challenges some mentees face, including alcoholism and substance use disorders, peer mentors offer assistance by setting recovery goals, providing guidance on recovery processes, and motivating them to pursue positive life prospects. Furthermore, due to the significant past traumas and current living conditions, mentees often grapple with severe mental health conditions and potential triggers that can lead to crises. Peer mentors remain vigilant to effectively address these situations, and some of them offer their support around the clock, ensuring immediate assistance when needed.

4.10.4 "Battlepeer": peer support smartphone application

"Battlepeer" smartphone application was developed to mimic a human-to-human peer support model for veterans struggling with civilian reintegration, with USMVs participating in an intensive, multi-month collaborative design process to provide initial user requirements and interface ideas, ongoing feedback for app refinement, and finally to recruit veterans to use the system (201). This application features two interfaces based on the user profile created at the outset. It facilitates the matching of mentors and mentees, and once mentees are assigned to specific mentors, they are given the opportunity to create their mentee profiles. In this profile, mentees can use either their real names or pseudonyms, select avatars, indicate their ranks, and specify the military branch with which they were affiliated. As mentees, they gain access to mentor information, including names, phone numbers, and military backgrounds. They can also engage in daily check-ins using a basic self-report data collection method through an Ecological Momentary Approach (EMA) (229; 228). On the contrary, peer mentors can create profiles as mentors and access all the contact information of their mentees. Mentors also have the ability to view descriptive visualizations of the check-in survey inputs from each of their mentees (Figure 7), aiding in mentor decision-making. For both mentors and mentees, the app includes a community chat and resource-sharing feature covering self-care, VA, transportation, employment, and more. Its main goal aligns with the mobile peer support trend, enhancing communication between peer mentees and mentors.

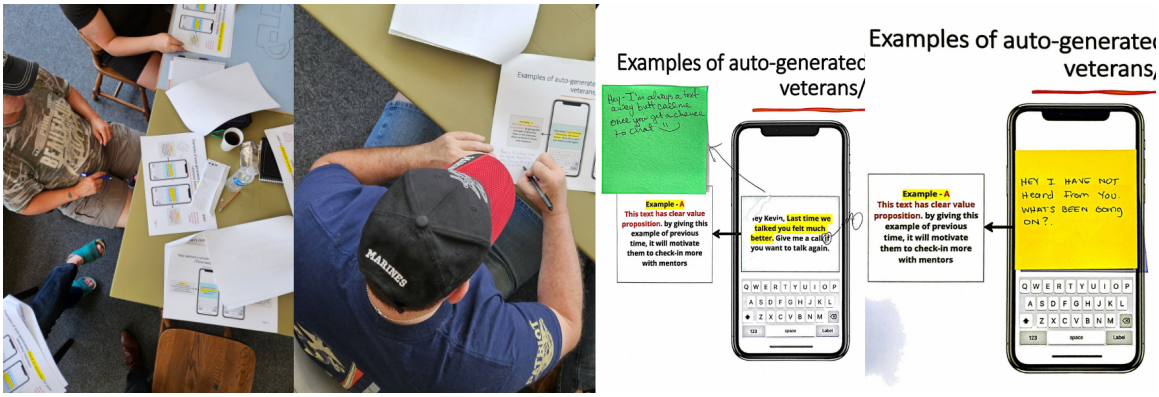


Figure 6: Illustrative photos from the participatory design workshops, including one of a peer mentor in a Marine hat writing comments, while other mentors, including LGBTQ+ members, actively participate. The top right image is a scanned sketch by a peer mentor during our workshops.

4.11 Methods

Our study consisted of three phases: in the first phase, three focus groups were conducted, the findings from these focus groups guided the design of persuasive reminders in the second phase, and in the third phase, we conducted two participatory design workshops to improve the design by embedding inputs on specific values of the stakeholders.

4.11.1 Participant Recruitment

For the focus groups, we recruited a total of ten peer mentors (eight male, two female) as participants through our partnership with the Dryhootch of America (202) dedicated to veteran welfare by giving an open call for peer mentors. Participating in this study was completely voluntary and participants could opt out at any time without penalty or loss of benefits to which they were otherwise entitled. The participants' age ranged between 25 and 69. They worked in different branches of the military including army, navy, marine corps, active duty, and reserves with highest in army. Four of them identify as African-American, five as White, and one chose not to disclose their preference. Their mentoring experiences varied from 0 years to 10 years. They all have gone through the peer mentorship training programs facilitated by the Dryhootch of America and other facilitators.

4.11.2 Data Collection and Analysis

We collected data through discussions and cognitive assessments involving drawings and sketches in the focus groups. Here, we describe the process of data collection,analysis and design followed in the three distinct phases. Figure 5 outlines these stages.

4.11.3 Phase 1: Identifying strategic application areas for persuasive reminders

In the initial phase, we conducted three focus group sessions, each spanning about 60 minutes. Following each focus group, two researchers conducted a debriefing session to document their initial impressions of the discussion, as well as any notable highlights, concerns, or immediate conclusions that had arisen. During these sessions, we inquired into the values and prerequisites necessary for crafting persuasive reminders within this particular context. The discussions centered around the firsthand experiences of individuals serving as peer-mentor veterans within a peer-mentor support system, focusing on their roles in supporting peer mentee veterans. We posed open-ended questions concerning engagement-related issues related to such interventions. The discussions were guided by an interview protocol but aimed to foster open dialogue. In one of these sessions, to illustrate the concept of persuasive reminders and stimulate creative input for future design, we also presented various low-fidelity prototypes of persuasive reminders that had already been employed in health-related contexts for different demographics, to get the participants acquainted with the concept of persuasive reminders. We sought feedback from our participants to discern which features were deemed valuable, required refinement, or should potentially be discarded. The initial focus group sessions primarily concentrated on the users' needs, perspectives, and concepts pertaining to reinforcing their ongoing commitment to the treatment process, with a particular emphasis on the role of persuasive reminders. These focus groups took place between September 2022 and December 2022.

4.11.4 Phase 2: Development of exemplary scenarios of persuasive reminders

Following the initial set of focus groups, a 12-week period from January 2023 to March 2023 was dedicated to the analysis of the data gathered during these sessions, with the aim of identifying strategic areas for the implementation of persuasive reminders. The audio recordings and accompanying notes were meticulously transcribed and anonymized by two of the authors. Subsequently, the collected information was organized according to specific questions and thematic categories, drawing upon established thematic analysis methodologies (200; 147). By discerning recurring key concepts that emerged throughout the focus group discussions, the principal themes and ideas relevant to our study's inquiries were derived.

Over the following 8 weeks, spanning from April 2023 to May 2023, we crafted exemplary scenarios for persuasive reminders based on the insights gleaned from our initial focus group sessions and informed by existing literature on persuasive reminders (210; 177). Paper prototypes of these scenarios

were designed. These designs underwent thorough careful examination and discussion with two subject matter experts, comprising an experienced peer mentor veteran and a professional psychologist. The professional psychologist, with nearly two decades of academic experience, has worked closely with the USMVs and served as a psychologist in a reputable medical college and veteran affairs. He is affiliated with different local organizations dedicated to veteran welfare and has published many peer-reviewed research articles on this topic. The experienced peer mentor veteran not only served in the Vietnam War but has also been an active mentor for over a decade, contributing to veteran welfare. He has participated as a researcher in multiple studies conducted by these collaborative effort. These experts played a crucial role in helping us understand the appropriate contextual language, predict potential responses, and determine the optimal way to present prototypes to the participants. Multiple rounds of improvements were carried out before presenting the final versions to the participants. We ended up with five distinct scenarios, each with two illustrative examples showcasing notable differences in language and underlying persuasion techniques. A detailed description of these scenarios is presented in section 5.2.

4.11.5 Phase 3: Exploring underlying values and needs for effective persuasive reminders

In the final phase, two sessions of participatory design workshops has been arranged in June 2023. It began with a brief presentation of some researchers' ideas about the proposed intervention, building upon the results from the first workshop. The second part involved presenting researchers' ideas on the proposed scenarios, followed by collecting participants' opinions and reflections on the proposed concepts and intervention features. The prototypes of the scenarios were printed in a large A3 paper with enough room on all sides, so that participants can write or draw sketches to present their own ideas and comments. (The study protocol for the participatory design workshops has been provided in the supplementary material section). These written expressions and drawings assisted us in grasping their thought processes regarding how we could tailor these reminders to better suit their individual needs. In certain instances, they composed entire texts that resonated with them, marked or circled words or sentences they found fitting, and utilized emojis to convey their approvals or disapproval. In some cases, we observed more drastic measures, such as crossing out entire sections of text or placing sticky notes over the content to indicate their disapproval. 6 illustrates a selection of photographs captured during our sessions. In one of the images, we observe a peer mentor wearing a marine hat while jotting down their own comments. Meanwhile, other peer

mentors, including members of the LGBTQ+ community, are actively engaged in the process. The image in the top right corner presents a scanned rendition of a sketch created by one of the peer mentors during one of the sessions. Our aim was to gain a deeper understanding of the intricacies of mentors' decision-making processes when responding to these notifications and to explore the context and rationale behind their responses. Two researchers from the team were present in them and took detailed notes of their involvements and comments. Subsequently, the audio recordings were transcribed and anonymized by two authors. The analysis comprised two phases. Initially, after becoming familiar with the data, we initiated the first phase by generating over 50 open codes to capture essential concepts from the participants. The relatively high number of open codes aimed to meticulously capture the nuanced insights within each discussion and was subsequently significantly reduced through the processes of memoing and clustering (200; 147). In the subsequent phase of analysis, we utilized a constant comparison method, specifically operationalized as affinity mapping, to memo and cluster the codes. This involved comparing each open code to others, positioning them based on their affinity to emerging themes and clusters, and exploring relationships between these themes. The ongoing process of constant comparison and collaborative discussions among the authors, who were domain experts, ensured the refinement of themes, resulting in a comprehensive and coherent representation of the data. Ultimately, a clear narrative was crafted to convey the identified themes, substantiating the findings with illustrative quotations and examples.

4.11.6 Data Integrity and Ethical Considerations

Considering that the discussion revolved around sensitive issues, we made sure to carefully consider the ethical implications of our work and to protect the rights of the participants. All sessions were recorded with participants' consent. The data was anonymized, with each dataset assigned a random code number in place of any personal identifiers such as participants' names or other identifying information. The study was approved by the Institutional Review Board (IRB) of the authors' institution.

4.11.7 Positionality

The research team was composed of individuals from three distinct fields of expertise, including three domain knowledge experts, one seasoned psychologist with a background in working with USMVs, two proficient peer mentors, and two researchers. One of these researchers specializes in intersectional stigma, online social support, and mental healthcare, while the other is an expert in working with marginalized and vulnerable populations. The entire group actively participated in

all aspects of the study, including designing the research, recruiting participants, and conducting both data collection and analysis. During the data collection stage, one of the researchers and two domain experts (the psychologist and one of the peer mentors) were present in all focus groups and participatory design workshops. These experts played a crucial role in facilitating the data collection process, ensuring the discussions ran smoothly, and offering guidance to the researchers in managing specific situations, such as maintaining coherent discussions and helping participants understand the topics in their terminologies. Furthermore, the presence of psychologists served as a precautionary measure to address any unforeseen situations, given that these veterans are highly susceptible to crises and can be easily triggered by sensitive discussion topics. Collectively, this team brings extensive experience in conducting collaborative, community-engaged research with USMVs, with some members having dedicated over a decade to this endeavor and the rest having more than two years of involvement in such research initiatives.

Furthermore, in pursuit to gain deeper insights into the dynamics of peer mentor-mentee relationships and to comprehend how peer mentors establish trust, provide support, and define the nature of their assistance, two members of the research team actively participated in weekly peer mentor meetings from January 2023 to August 2023. These observations assisted them in establishing a strong rapport with the peer mentors and gaining a deeper insight of the military context within the peer support intervention.

4.11.8 Limitations

Regarding data collection, we chose to engage only peer mentors in the initial phase, as we decided not to involve peer mentees due to their vulnerable state. However, it's worth noting that all the peer mentors themselves have either experienced being a peer mentee or have dealt with mental health-related issues. Furthermore, their close connection with their mentees and formal training of peer mentorship allowed them to effectively explore the needs and values of their mentees during the design process.

4.12 Results

The study was structured into three distinct phases. Phase 1 findings were instrumental in formulating strategies for the application of persuasive reminders in this context. In Phase 2, we developed these persuasive reminders based on the strategies. Finally, in Phase 3, our findings shed light on the essential values and requirements that should be integrated to enhance the effectiveness of these reminders.

4.12.1 Findings from Phase 1: Strategic Application Areas for Persuasive Reminders

To determine the optimal areas for deploying persuasive notifications, we initially conducted multiple focus groups to identify the domain-specific expertise needed. Our analysis of the feedback from peer mentors revealed a range of contexts where these persuasive reminders could be highly influential.

Reminding mentors to maintain communications with their mentees Peer mentors are the individuals who willingly offer support to fellow veterans facing mental health issues, leveraging their shared experiences. However, these mentors have their own lives, family commitments, and job responsibilities, making it challenging to strike a balance between their personal obligations and their role as mentors. The need to minimize their cognitive load is evident since constantly thinking about their peer mentees and support groups is not a viable option. One of the peer mentors said -

> *"I used to but I can't afford to do it for my own mental health, to take it home all their stuff. This is not an easy pace to me."*

On the other hand, peer mentees are in a vulnerable state and may experience crises at any time, requiring immediate attention and support. Given that these peer support groups operate through various mobile applications, where peer mentees are responsible for daily or weekly check-ins, these check-in results can serve as a valuable resource. Figure 7 shows such a snapshot from a mobile application displaying concerning check-in results from peer mentees, possibly warranting immediate attention. They can be leveraged to trigger more effective notifications when any of the mentees encounter issues related to their daily activities or mental health. This approach will help strike a balance between providing timely support and alleviating the cognitive burden on peer mentors

Reminding mentees to complete check-ins through the app or phone In a peer mentor-mentee support group, maintaining regular communication between mentors and mentees is of utmost importance. This ongoing interaction allows mentors to assess the current situations of their mentees, enabling them to provide the necessary support and attention when needed. However, there are instances where mentors encounter difficulties in reaching their peer mentees. One of the peer mentor said

> *"you know, whether he went two months without talking to me and ignoring my calls to the last trial for help, and what should I do?"*

In such cases, deploying persuasive reminders to motivate mentees to reengage in communication

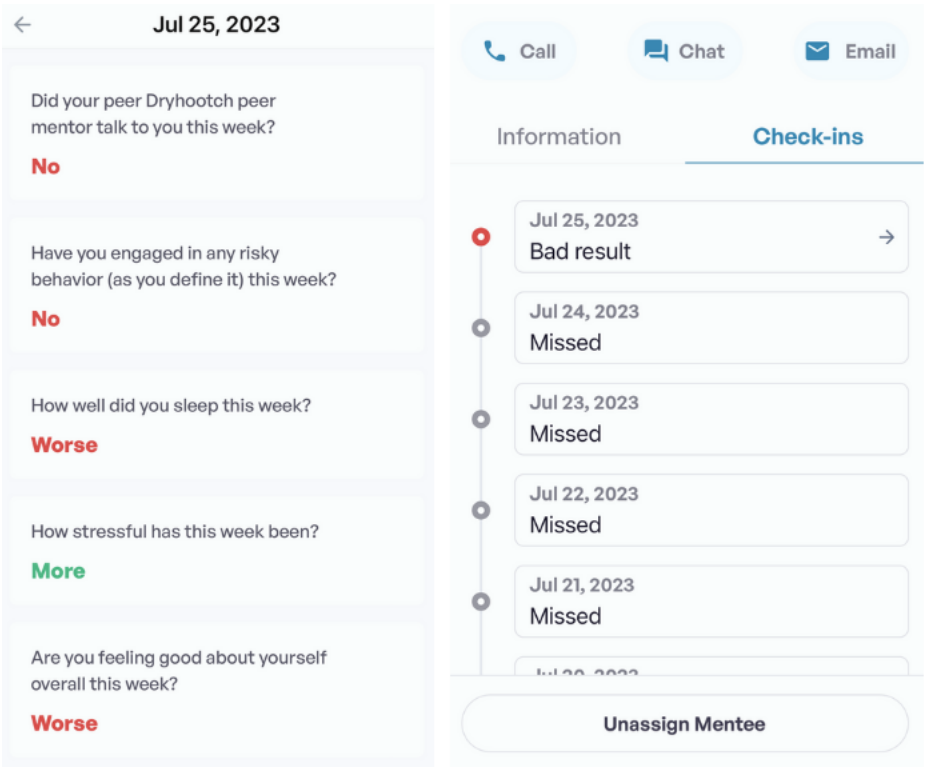


Figure 7: A snapshot from the "BattlePeer" app - a peer support smartphone application displaying concerning check-in results from peer mentees.

with their peer mentors can prove to be highly effective. These reminders can be served as a gentle nudge to encourage mentees to reconnect with their mentors, ensuring that the crucial lines of communication remain open for support and assistance.

Alerting peer mentees to immediate support during early crisis situations As per insights from peer mentors, it's important to recognize that peer mentees may find themselves in an early crisis situation or experience escalating negative thoughts at any given time. Leveraging their experience and training, peer mentors possess the ability to identify these critical situations and take proactive steps to provide assistance.

However, according to mentors, a common challenge arises when mentees tend to isolate themselves and discontinue communication during such difficult periods. Prior literature further highlights the challenge that USMVs encounter when it comes to openly sharing their trauma and related events with their family members (111). In such instances, it becomes essential to proactively remind these mentees of the available support networks and resources they can access. The objective is to motivate and encourage them to reestablish communication, ensuring they are aware of the assistance and guidance that remains at their disposal during challenging times.

Prompting peer mentees to access available wellbeing resources During our conversations with peer mentors, they have consistently emphasized a fundamental aspect of their role, which is to provide valuable resources related to self-care, medical facilities, housing, and employment to their peer mentees. Whenever mentors come across resources that could significantly benefit their mentees, they are eager to share this information. However, a challenge arises when mentors are unable to establish contact with their mentees.

In such situations, it becomes crucial for mentors to regularly remind their peer mentees about the existence of these valuable resources that can significantly contribute to their well-being. These reminders serve as prompts to ensure that mentees are aware of and can take advantage of the available resources that can enhance their overall quality of life. By maintaining open lines of communication and providing these timely reminders, mentors can better fulfill their role in supporting the well-being of their mentees.

4.12.2 Exemplary Scenarios of Persuasive Reminders

Based on the values and criteria identified in previous section, we have developed five distinct persuasive reminders, each meticulously crafted with unique language and motivational elements. We considered O'Leary et al.'s taxonomy of persuasive reminders for health self-management that discusses four different types: introspective, socially supportive, adaptive, and symbolic (210). **Introspective reminders** serve as prompts for individuals to contemplate their health goals, helping them plan and prioritize their daily health-related tasks. These reminders are timely and trigger self-reflection on daily activities, encouraging individuals to prioritize their mental well-being. Initiating these reminders ahead of tasks is crucial for fostering participants' awareness of their health journey and placing their health-related tasks in context. In contrast, **socially supportive reminders** are designed to cultivate motivational and mentoring relationships, enhancing users' confidence in performing health-related tasks. **Adaptive reminders** are tailored to persuade users by adjusting to their unique preferences and characteristics. Lastly, **symbolic reminders** motivate individuals to engage in health tasks by evoking personally significant reasons for adopting healthy behaviors (210). Below we describe the five scenarios of sending persuasive reminders under this taxonomy, which are further summarized in table 6. Throughout the rest of the chapter, we use the pseudonyms Josh and Kevin to represent a peer mentor and a mentee veteran respectively.

Scenario one

The reminder is designed to be sent to the peer mentors automatically to ensure that they maintain

	Example A	Example A illustration	Example B	Example B illustration
Scenario 1 (This text is meant for the peer mentors)	Hey Josh, Kevin's survey response is alarming. Click it here to check it now >>	1. More direct approach 2. Quick and actionable 3. Only shows alarming cases	Hey Josh, some alarming cases have been identified based on the peer vets' survey responses. Click it here to check it now >>	1. Allows them to identify less alarming cases too. 2. By going through the app, peer mentors will be able to judge themselves whether the case is alarming or not
Scenario 2 (This text is meant for the peer mentees)	Hey Kevin, last time we talked you felt much better. Give me a call if you want to talk again.	This text has clear value proposition. By giving them an example of previous time, the goal is to motivate them more towards check-in with their mentors	Hey Kevin, please know, I am here for you in your difficult times and not so difficult times. Give me a call if you want to talk	This text is more general and have a gentle nudge towards the value (without pushing too much)
Scenario 3 (This text is meant for the peer mentees)	Hey Kevin, you are on a roll. Complete this week's survey to maintain all the good work you have done so far. Click here to complete that >>	This text motivates the peer veterans by reminding them about their previous accomplishments	Hey kevin, you are on a roll. Only two days remaining to complete this week's survey. Click here to complete that >>	This text motivates the peer veterans by reminding them about deadlines
Scenario 4 (This text is meant for the peer mentees)	Hey, how are you? Are you having any negative thoughts lately? It's better to keep them in check while you can. Let me know what I can do for you.	1. This text creates urgency. Puts importance on completing the task. 2. NO NAME like "Hey Kevin" 3. More direct	Het Kevin, How are you? Just reminding you that there are great resources available that can help anyone deal with negative thoughts. Do you want me to share those resources with you?	1. This text is more general and have a gentle nudge towards the value (without pushing too much) 2. Asking them questions instead of request at the end
Scenario 5 (This text is meant for the peer mentees)	Hey Kevin, I recently learned some excellent information about steps you can take if you don't feel good. the next time we speak, please remind me to bring that up	1. This text alerts the peer mentees about the new resources. 2. Is it better to give the peer mentees an action by "reminding you"	Hey Kevin, _____ _____ _____ The next time we talk, I'll tell you about it.	1. Gave the participants an opportunity to share what they will send while sharing these resources with someone 2. Keep the control of the action by "I'll tell you"

Table 6: In-depth exploration of five persuasive reminder scenarios with actual texts and illustrations.

contact with their mentees. **Introspective reminders** hold significance here, primarily concerning the task of reminding mentors to maintain continuous communication, especially when mentees require urgent attention due to their specific mental health conditions. We presented two options to the participants, each with distinct characteristics. Example A takes a more direct approach by clearly identifying which mentees require extra attention. It doesn't require mentors to log in to the application, as they can connect with their mentees through alternative means like personal phone, email, or social media, enabling them to take immediate action. On the other hand, Example B empowers mentors to identify critical cases independently within the application. When they receive notifications, mentors are expected to access the application and assess the situations on their own. This approach also grants mentors the flexibility to reevaluate which mentees they want to follow up with.

Scenario two

Scenarios	Summary of reviews	Uncovered strategies
Scenario 1	- Mentioning peer mentees name directly (Kevin) is appreciated than generic alarming texts	* Strategies for personalization - Address with names
Scenario 2	- The idea of starting with "Hey Kevin" gets rejected - Use of emojis is recommended - Clear value proposition needs to be more tailored - Felt like too much use of filler languages - The idea of imposing full accountability on mentees (e.g. "Give me a call") gets rejected	* Strategies for personalization - Address with names - Use emojis * Strategies for providing encouragement - Encourage through past success stories Strategies for creating actionable reminders - Show intent to initiate communication rather than using filler sentences * Strategies for increasing mutual accountability
Scenario 3	- Motivating everyone with same type of text is not appreciated. Some needs more motivation for check-ins than others - For some demotivated mentees, just getting a short answer about their conditions needs to be prioritized than getting a formal check-in through the app - Term like 'survey' need to be avoided	* Strategies for providing encouragement - Understand the level of motivation needed - Select terminology carefully * Strategies for creating actionable reminders - Strike a balance between open-ended and yes/no questions
Scenario 4	- Starting with simple 'Hey' is appreciated - Term like "negative thoughts" needs to be avoided - Felt like too much use of filler languages - In some cases, asking about triggers is recommended - The idea of imposing full accountability on mentees (e.g. let me know, do you want me to) gets rejected	* Strategies for personalization - Address with names * Strategies for providing encouragement - Select terminology carefully - Address triggers in addition to negative thoughts * Strategies for creating actionable reminders - Show intent to initiate communication rather than using filler sentences * Strategies for increasing mutual accountability
Scenario 5	- Felt like too much use of filler languages - Providing direct guidance is recommended instead of giving options - Sharing accountability mutually between mentors and mentees is recommended (e.g. "please remind me") instead imposing the whole responsibility to mentors	* Strategies for creating actionable reminders - Show intent to initiate communication rather than using filler sentences - Provide direct guidance through constraints in choices * Strategies for increasing mutual accountability

Table 7: A brief overview of participants' feedback on five scenarios and uncovered strategies

This reminder is intended for peer mentees. We've introduced two options for consideration. Example A features a text with a distinct value proposition. It leverages a successful past instance to inspire peer mentees to establish a connection with their peer mentors. Conversely, Example B adopts a more generic and non-invasive approach. It aims to gently encourage mentees to find motivation in engaging with their mentors. This is an example of a **socially supportive reminders** where motivational factors and mentorship values have been incorporated.

Scenario three

This scenario incorporates elements from both **introspective and socially supportive reminders**. As part of the peer support group, it's the responsibility of mentees to fill out an EMA survey during daily/weekly check-ins. This reminder aims to encourage peer mentees to stay motivated and complete their weekly check-ins. We've provided two different examples, each with distinct approaches. Example A motivates peer mentees by reminding them of their past accomplishments in completing these check-ins, using positive reinforcement as a motivational tool. In contrast, Example B focuses on motivating mentees by emphasizing the deadlines for completing the surveys. Our intention here

is to explore what works best for the mentees—whether positive reinforcement or deadlines prove to be a more effective motivator.

Scenario four

This scenario addresses unique perspectives and characteristics found in peer mentees' individual situations. Consequently, **adaptive reminders** are the most suitable choice in this context. This text is designed to remind peer mentees that support is available to them during difficult times. There are distinct characteristics that set Example A and Example B apart in Scenario Four. First, Example A takes a more direct approach, creating a sense of urgency and encouraging mentees to actively seek the support they may need, using advisory language to gently nudge them in the right direction. On the other hand, Example B adopts a gentler tone, merely informing mentees about the various supports available to them in case they are going through challenging times. Additionally, we experimented with different language components. Example B begins by mentioning the peer mentee's name, making it more personalized, while Example A remains more generic. Towards the end, Example B concludes with a question to avoid establishing any authority over the mentees, whereas Example A sets a somewhat authoritative tone, gently encouraging them to seek support if necessary.

Scenario five

This final reminder serves as a conversation starter between mentors and mentees, helping to bridge any communication gaps that may arise due to conflicting schedules. Similar to scenario 4, **adaptive reminders** are the the most suitable choice for this scenario. In Example A, the reminder alerts peer mentees about new resources, providing a specific conversation point to rekindle their interaction. Additionally, by including the phrase "please remind me," it encourages mutual accountability in sharing important resources or information. Example B, on the other hand, is left open-ended and blank, allowing mentees to suggest topics or ideas for reconnecting in case of a time gap. However, it's worth noting that in this case, the mentor takes on the sole responsibility by saying "I'll tell you," which could imply greater control over the conversation and information sharing. The objective is to explore the level of control or accountability that both parties are comfortable with when it comes to reestablishing communication.

4.12.3 Findings from Phase 3: Underlying Values and Needs for Effective Persuasive Reminders

We organized two sessions of participatory design workshops to gather feedback on the design of our five unique reminders. The participants provided invaluable feedback and suggestions concerning various aspects, including strategies for personilization, providing encouragement, creating actionable reminders, increasing mutual accountability. Table 7 presents a brief overview of participants' feedback on five scenarios and uncovered strategies.

Strategies for personalization Past research has shown effectiveness of personalization in persuasive technologies (257). Our participants pointed out a few techniques for personalizing the persuasive reminders including the use of receiver's name, emoticons, and context specificity.

Address with names: In an effort to make text notifications more personalized, we incorporated the names of peer mentors and mentees into various parts of the messages. Initially, we used greetings like "Hey Kevin" (a pseudonym for a peer mentee) and "Hey Josh" (a pseudonym for a peer mentor). Our participants pointed out that, the effectiveness of personalization through names depends on the context and purpose of the communication. For example, in this case, despite addressing with names, the inclusion of generic greetings could lead to a loss of interest among peer mentees. They were concerned that this might give the impression that the messages were automated, ultimately undermining their motivational impact. However, they express a preference for the design presented in scenario 4, specifically example B, where the message initiates with a simple "Hey." Additionally, they suggest making the greetings less formal by using phrases like "Hey buddy" or "Hey man."

However, in the scenario of alerting peer mentors about a specific mentee needing immediate attention (example 1A), participants found it powerful and motivating. This personalized approach inspired them to take quick action through their preferred communication medium. They preferred this version over example 1B, where the focus is on encouraging mentors to use the mobile application to check on mentees that require attention. Moreover, receiving multiple reminders, one with each alarming participant name would be more effective, according to the mentors, as it would encourage accountability, too. They argued that if many of their mentees are struggling, since they usually follow similar mentoring techniques for all the mentees, the rest might be struggling, too. They would feel the urge to check on other mentees as well, to make sure that they are okay.

"If I recieve four different texts with four names and I have six mentees in my support

group, which will mean, Man, I am not doing a good job and this will in turn motivate me to contact with all the mentees indead of that four."

In summary, the use of names should be tailored to the particular context and the preferences of both mentors and mentees to achieve the desired level of engagement, motivation, and accountability.

Use emojis: Emojis can be helpful to communicate sentiment, and to improve communication and understanding. As we did not include any emoticons in our example scenarios, the participants placed a strong emphasis on the need for using humor and the inclusion of emoticons in order to humanize the communication. They believed that incorporating these elements would not only make the messages feel less automated but also infuse a sense of sentimentality into the interactions. By tailoring the language for particular mentees based on their relationships and previous communication styles with the mentor, and by injecting humor and emoticons into the messages, the participants can foster a more genuine connection. These small touches of personality and emotion were seen as essential in transforming routine reminders into meaningful and engaging exchanges, ultimately enhancing the overall quality of communication. The following example is from a participant's ideation showing the use of emoticons infusing a touch of humor to make the situation feel more relaxed and normal.

"Hey, I'm always a text away. butt call me once you get a chance [1.2][yellow!60!red]"

Strategies for providing encouragement It is essential for the mentors to make sure that mentees respond to check-in surveys through the app at least once a week, so that they are able to assess their mental health symptoms. The participant mentors discussed different approaches to provide constant encouragement, including adding success stories, identifying the level and type of motivation needed for each individual, addressing the triggers behind negative thoughts, and by careful consideration of words in the text content.

Encourage through past success stories: In Scenario 2, Participants emphasized the importance of a clear proposition tailored to the peer mentee's specific situation. They suggested focusing on successful events or strategies that both mentors and mentees have used in the past. Sometimes peer mentors assist mentees with various challenges such as substance abuse, housing, transportation, connecting them with VA facilities, or helping them find employment. They suggested adjusting the clear proposition to revolve around successful events or strategies that both mentors and mentees have experienced. Participants also recommended providing options with a few examplary tones to enhance engagement and capture the mentees' attention. Participants also suggested that sharing

anecdotes about their past successes in developing positive habits or accomplishing daily tasks could serve as a valuable strategy to motivate themselves and take initial steps toward enhancing their mental well-being. By recounting their own triumphs, they believed it would be easier to gain a foothold and begin the journey of improving their mental health. These personal stories would not only provide inspiration but also instill a sense of self-efficacy, empowering them to make gradual progress in their pursuit of better mental health. In essence, leveraging past achievements as a source of motivation and encouragement emerged can be a promising approach. One participant said if their mentees go to gym daily, a text should be something like -

"Hey, you haven't been to the gym. Are you ok?"

Understand the level of motivation needed: In Scenario 3, The participants emphasized the need for tailored and custom motivation. They acknowledged that some mentees consistently check in and respond well to a direct approach. For instance, they mentioned that for these mentees, a notification with a direct message such as, "Could you please check in on the app?" would suffice. However, others, who are less inclined to check in regularly, require a gentler approach. Instead of directly requesting check-ins, participants proposed an alternative approach that they found to be effective. They mentioned that while check-ins through the app are valuable, not everyone is enthusiastic about routine tasks like these. In such instances, initiating informal check-ins through text or phone calls should be prioritized. One participant gave such an example -

"Hey, I have not heard from you, what's been going on?"

Once they start feeling more comfortable, gently encouraging them to participate in app check-ins can be a successful strategy. Recognizing these differences in motivation levels is crucial for effective communication.

Address triggers in addition to negative thoughts: In Scenario 4, where the text notification addresses sensitive topics like negative thoughts, participants stressed the importance of asking about the triggers behind these thoughts rather than just addressing the concern. Understanding the triggers assists them in determining the appropriate next course of action. However, they pointed out that obtaining answers to such sensitive topics requires highly customized options tailored to their needs.

Select terminologies carefully: While reviewing the scenarios, participants emphasized the importance of choosing words carefully when sending text notifications to USMVs facing mental health challenges. They pointed out that certain words, such as "negative thoughts" and "crisis," should

be avoided for two significant reasons. First, using such terms might wrongly imply or assume that the peer mentee is in a crisis situation, potentially causing unnecessary discomfort or even triggering distress. Therefore, word choice in notifications needs to be crafted with caution, ensuring sensitivity to the mentee's mental state.

Additionally, during the examination of Scenario 3, participants noted that words like "survey" should be avoided as they can make the peer mentees feel like they are part of an experiment rather than receiving genuine support.

Furthermore, participants highlighted the importance of recognizing words with different meanings in a military context. For instance, the word "tools" may have a distinct connotation in military terminology. To avoid potential confusion or misunderstanding, it was suggested that instead of using "self-care tools," the term "beneficiary resources" should be employed to ensure clarity and relevance in communication with USMVs. This adjustment ensures that the language used is both respectful of their experiences and conducive to effective communication and support.

Strategies for creating actionable reminders

Our study participants placed significant emphasis on the importance of crafting reminders that are actionable with a clear intent to initiate communication. Instead of resorting to filler sentences, they advocated for concise and purposeful reminders. Additionally, they highlighted the value of offering direct guidance by implementing constraints in the available choices, making it easier for users to make decisions. Moreover, our participants stressed the need to strike a balance when designing questions. They recommended finding equilibrium between open-ended inquiries that allow for thoughtful responses and yes/no questions that provide straightforward answers. This approach will ensure that the interaction remains engaging and informative while also facilitating efficient communication.

Show intent to initiate communication rather than using filler sentences: Our initial scenario involving notifications sparked a discussion about the inclusion of unnecessary or filler language. None of the Scenario 2 examples was well-received as they contained excessive filler content. Mentors suggested that the messages should be actionable instead of false or unnecessary pretenses.

Rejecting both examples, participants took it upon themselves to propose their own text messages, and appeared to converge on a similar format. This format mainly involved prompting peer mentees to initiate communication when there had been a significant gap in their previous interactions. These

messages typically consisted of two sentences: one mentioning the lack of recent communication and the other requesting a call or text to kickstart the conversation. For example, one of the participant wrote -

> *"Been a few days since we talked. Hope all is well. Can you do a quick check-in on the app thanks"*

During the discussion, participants highlighted the issue of filler language when presented with Scenario 4 and Scenario 5. Their immediate reaction was that the text notifications felt artificial and insincere, describing it as "some AI nonsense." In multiple instances, peer mentors have put emphasis on the importance of carefully crafting the language to ensure it feels personalized and tailored. The use of generic and repetitive pushy language could lead to a loss of credibility toward the system. Peer interventions for USMVs prove effective because they are personalized and create a comfortable environment for exchanging support with those who share similar experiences. It's crucial for systems to emulate this personalized approach, providing users with a sense of familiarity and confidence akin to what they experience when interacting with their peers who have undergone similar situations.

Provide direct guidance through constraints in choices: In their attempts to explain how to make the notifications action-oriented, our participants emphasized on motivating mentees break out of their isolation, particularly during the early crisis stages. Mentees often tend to isolate themselves from their daily routines and tasks during such times, which can lead to the accumulation of negative thoughts. Therefore, one of the top priorities should be to encourage mentees to get off the couch and engage in their daily routines. In scenario 5, participants mentioned instead of mentioning about the choices, providing direct guidance would be effective.

Participants suggested a more focused approach with a single coping strategy instead of offering a choice among four different ones. This streamlined approach aims to provide clarity and direct guidance to the recipients to understand and implement the suggested coping strategy.

Strike a balance between open-ended and yes/no questions: Our participants mentioned instances where receiving a simple one-word response from their mentees become critically important, particularly when a mentee isolates himself, and the mentors have to repeatedly call or send texts for several days without getting any response. In such cases, they found that sending a text message that can be answered with a simple 'yes' or 'no' can be highly effective.

However, the text needs to be carefully crafted to strike a balance. It should convey a sense of

urgency and concern, yet also reassure the peer mentees that they will be okay. It acknowledges that the veterans might need some time alone to gather themselves and regain their strength while letting them know that support is available when they're ready to reach out. For example, *"Are you alive? Just let me know are you alive!"*

Strategies for increasing mutual accountability

To explore the shared responsibility of maintaining communication between mentors and mentees, each scenario in our notification system concluded with an actionable directive, such as "click it," "call me," or "remind me." In both Scenario 1 and Scenario 3, we implemented a convenient shortcut option for the respective user groups, i.e., peer mentors and peer mentees, allowing them to easily access the mobile application with just "one click." In Scenario 1, this feature primarily serves the needs of peer mentors, as they hold the responsibility of overseeing their peer support groups to identify those who may require additional attention. In contrast, in Scenario 3, this feature is tailored for peer mentees since they are solely responsible for completing the weekly check-ins. Although peer mentors might opt for direct communication with their mentees through alternative channels to remind them about completing check-ins, our participants generally expressed appreciation for having the option of automated reminder available. Now, in scenario 2 and 4, we framed the sentences ("Give me a call if you want to talk again", "Let me know what I can do for you") in a manner that placed full accountability on the mentee for sustaining communication between mentors and mentees. This approach was largely rejected by the participants.

"This "if" is a daydream"

Rather than offering excessive flexibility in the phrasing the message, participants suggested adopting a slightly more assertive tone. They recommended using expressions like "Could you perform a quick check-in within the app?" or simply "Please give me a call." This approach aims to gently encourage action and prompt engagement while maintaining a clear and direct communication style. In Scenario 5, the two examples presented significant differences regarding the distribution of accountability. In Example A, there's a shared sense of responsibility for maintaining communication. By phrasing it as "please remind me," both mentors and mentees are encouraged to play their part in the process. Conversely, in Example B, mentors assume complete responsibility by stating, "I'll tell you," indicating that they retain control over to ensure that the information reaches the mentees. Participants favored the approach of Example A. It allowed for a safety net in case someone forgot due to their other commitments or responsibilities. With both sides sharing the responsibility, the

likelihood of "information getting lost in the communication" is reduced. This mutual accountability contributes to a more robust and reliable system of communication, ensuring that important information is consistently conveyed and received.

4.13 Discussion

Our findings emphasize the importance of context-specific design and the potential of semi-guided reminders to alleviate the cognitive loads on peer mentors. Furthermore, it highlights the need for a more comprehensive user-centered persuasion design approach when addressing the distinct requirements of USMVs

4.13.1 Thinking Beyond Traditional Steps of Design

We initially planned to follow the traditional steps of user-centered design (formative research → prototype → evaluate) (8): to conduct a formative research to understand user needs of persuasive reminders to facilitate mentor-mentee peer support in the veteran's context, to utilize this understanding and data in designing persuasive reminders, and eventually evaluate these reminders in future. However, in order to conduct an ethical community based participatory research (CBPR) (270) we made sure to protect not only individual research participants but also the communities and the population. Contrasting the traditional user-centered design process, our approach includes an additional layer of validation before the actual development. We begin by comprehensively understanding the context and specifying user requirements, but before proceeding to evaluation, we prioritize validating aspects like language, motivational factors, and customization options in conjunction with the actual design. This iterative approach ensures that the design aligns closely with the needs and preferences of the community, enhancing its effectiveness in this specific context.

Our findings have challenged traditional concepts related to the implementation of persuasive reminders, with a few notable observations. First, contrary to the established belief in the efficacy of personalized reminders that use the recipient's name as a motivational tool, our study suggests this strategy might sometimes undermine trust in sensitive contexts such as mentor-mentee relationships involving USMVs with mental health (MH) issues. Second, while traditional persuasion principles favor a softer, non-aggressive tone utilizing gentle nudges to spur action (267), the approach can lose its potency, particularly with high-risk groups. Reminders containing filler sentences tend to lose their impact and are easily ignored by this demographic. We advocate for a more straightforward communication strategy that invites short, one- or two-word responses, accommodating the potential lack of mentees' energy for extended replies, but still facilitates ongoing communication

by encouraging brief feedback.

4.13.2 Implications for Future Research and Design

Context-specific design To create reminders that truly resonate with this demographic, it is imperative to integrate veteran's perspectives into the design process. Prior study has shown that even after leaving the military, many veterans continue to live within the ideological framework of their military experience (165). Incorporating this context into veterans' daily lives is crucial for the effectiveness of interventions targeted towards them. Our research uncovered specific nuances in communication with veterans that have important implications for designing persuasive reminders. It became evident that certain words should be avoided when crafting reminders. Terms like "survey" can trigger a sense of being involved in an experiment, which they may not receive well. Additionally, the use of phrases related to "negative thoughts" may not be effective, as traumatized veterans often respond poorly to negative reinforcement. Even seemingly innocuous words like "tools" can carry a different connotation within the military context. These insights highlight that not all elements of language from prior research on persuasive technology are directly applicable for in this setting.

Furthermore, implementation of persuasive reminders often encourage guiding users with a wide variety of choices to facilitate decision-making as a part of empowering the users of making their own choice (253). Contrary to this, our findings suggest that a more constrained approach to providing choices may be beneficial. We identified that during complex situations, tasking mentors with simple, daily responsibilities can be challenging, but engaging in these tasks can significantly improve their well-being. In such cases, simple and single goal-oriented reminders might be more effective. For military context, this approach can be likened to assigning a mission, motivating individuals to work diligently towards completion. These findings highlight the importance of contextual design, particularly when supporting vulnerable populations like USMVs through peer intervention. It highlights that different groups have distinct approaches to recovery, emphasizing the need for tailored design strategies that consider the unique context and needs of the target population.

Implementation of semi-guided reminders

Using past success stories as a persuasive technique is a well-established method to motivate individuals to complete future tasks (13). In contrast, our research highlighted that employing generic success stories may not be as motivational as tailoring them to individuals' specific situations. Furthermore, mentors emphasized that different mentees require varying approaches to motivation. Some respond well to a direct approach, while others benefit from a more gentle tone.

In this context, incorporating human involvement, such as peer mentors, can be advantageous. Peer mentors possess a deep understanding of their mentees' current situations and can effectively gauge the most suitable motivational approaches. This insight can be leveraged to implement semi-guided persuasive reminders, where a range of exemplary tones for reminders is presented, but peer mentors retain control in setting up these reminders. Peer mentors can be allowed to customize the communication style and type of the digital reminders in a few ways : (a) they can be shown option to choose between a more direct approach or a gentler tone, (b) they may have the authority to prioritize the types of reminders sent to specific peer mentees. For example, some mentees who are proactive and do not require to be frequently reminded to complete check-ins, might benefit from reminders about available resources instead, (c) interactive reminders with quick and easy responses can be sent, facilitating the flow of information that mentors need to better understand the situation.. Prior literature suggests that semi-guided interactions can be particularly beneficial during vulnerable stages when individuals are dealing with significant concerns (141). This approach helps the system extend its reach and prompt early intervention when needed. These findings underscore the importance of not relying entirely on automated texts. It is essential that peer mentors are provided with sufficient flexibility within the system to customize various aspects of the communication, ensuring a more personalized and effective mentor-mentee interaction..

Another design recommendation involves implementing a "Battle Buddy" system for the mentees, drawing inspiration from the United States Army's practice of assigning partners to soldiers. In the military, a battle buddy is tasked with providing support to their partner, both in combat and in other aspects of military life. This concept can be adapted to peer support interventions, where the role of supporting each other is paramount (225). While the battle buddy concept has been explored in prior literature (187), with buddies responsible for preventing negative thoughts and, in extreme cases, suicide, it has its drawbacks, particularly when pairings are not well-matched (18). However, in the context of peer mentor-mentee support interventions, where mentors often shoulder these responsibilities, the battle buddy system can be employed for minimal interactions. A battle buddy's role would primarily involve monitoring their partner's activities, ensuring they complete daily tasks and check-ins with their mentors. This military concept can be particularly useful in designing "Socially supportive reminders" (210) that involve motivational and mentoring relationships. This type of persuasive health reminders have been used to leverage social support to not only remember tasks, but to increase confidence in performing them (210). In our context, two peers can be teamed up by mentors to achieve some goal (e.g., complete weekly check-ins) and one

peer can send reminders to the other peer about going through the activities. Then at the end of some achievements (e.g., if both complete check-ins in a timely manner), they both get rewarded. Veterans share similar values and experience, and thus can relate better to each other, making this type of gamification work well. Moreover, buddies can exchange information, allowing them to be aware of each others' situations. If a buddy is equipped with similar or adjusted access to generate semi-guided reminders, that can help relieve cognitive loads from the mentors to some extent as well. The customization of notification settings would be determined by the buddies themselves, and they would serve as an additional layer of intervention to assess the current situation of a peer mentee. Subsequently, they can report back to peer mentors, who can then take appropriate measures as needed.

4.13.3 Implications for Reducing Cognitive Loads of Peer Mentors

Our findings related to strategic areas where persuasive messages can be employed have implications for alleviating the cognitive loads on peer mentors, thus enhancing the effectiveness of peer support interventions. Therefore, our initial focus was to pinpoint strategic areas where these persuasive reminders could be most beneficial. Peer mentors were enthusiastic about the concept from the the very beginning and acknowledged its potential to reduce the cognitive load on peer mentors. Peer mentorship is a demanding role, often undertaken voluntarily out of goodwill. However, our findings illuminated instances where peer mentors experienced a significant mental workload. Each mentor is potentially assigned several mentees, each with unique expectations and support requirements from their mentor. Furthermore, mentors themselves frequently deal with their personal MH challenges. In this scenario, managing effective communication with all mentees can be a time-intensive task and may contribute to a significant cognitive load for the veteran mentors. Periodic reminders can play a crucial role in reducing mentees' cognitive loads, especially for those facing MH challenges. Managing multiple medications and therapies can complicate memory issues, hindering their ability to remember important tasks or self-care routines. Integrating persuasive reminders through a digital platform helps alleviate memory-related loads. Timely and tailored reminders for actions like check-ins or self-care activities reduce cognitive burden, aid in memory recall, and encourage adherence to essential health routines. The reminders may also offer a layer of consistency and reliability that can be comforting for individuals dealing with MH issues. In a landscape where MH conditions can introduce unpredictability and disarray into daily life, the steadiness provided by these persuasive reminders can be a source of stability and reassurance. Leveraging mobile technology not as a replacement but as an amplification tool can address this issue. Kentaro et al. proposed

that technology tends to amplify existing human forces, enhancing effectiveness where intentions and capabilities are positive, but yielding little or negative impact in areas where human forces are indifferent, dysfunctional, or corrupt (264). In our case, identifying situations where notifications can prompt actions, such as motivating check-ins, alerting peer mentees during early crises, or encouraging access to available resources, can amplify existing peer mentor-mentee intervention strategies by reducing the cognitive load on peer mentors. This, in turn, enables them to provide support more effectively and efficiently.

Furthermore, it's crucial for mentors in peer-mentor support groups to be informed about their mentees' current circumstances, especially considering the vulnerability of individuals with past trauma to crises at any time. Designing warning notifications for mentors regarding a potential crisis in one of their mentees is based on the fundamental expectation that those who receive such alerts will understand them and take appropriate action, whether by intervening themselves or notifying the relevant authorities. However, decades of social science research suggest that human responses to warnings are intricate (73). Even when an alert is issued and understood, it may not necessarily lead to the intended response, often due to a normalcy bias, where individuals downplay potential danger even when warned (73). Identifying situations where peer mentees can take independent actions through shared well-being resources and prompting them to seek immediate support adds an additional layer of safety, complementing the role of mentors and making peer support interventions more efficient.

4.14 Conclusion

In this study, we investigated persuasive reminders in peer support with ten US military veterans, emphasizing the significance of tailored strategies for effective communication, crisis symptom alerts, and well-being resources. Our findings highlighted the need for personalized approaches, including success story delivery, motivation understanding, language choice, actionable reminders, and mutual accountability. Designing mental health technologies for and with peer supporters is crucial, especially for veterans who face unique challenges. Trauma experiences impact their seeking and receiving of peer support differently from civilians. To address the unmet mental health needs of veterans and similar communities, understanding their perceptions of and coping with mental health crises is essential. We provided implications regarding the importance of context-specific design, elements to reduce cognitive load for peer mentors, and custom user-centered persuasion design for the distinctive needs of military veterans.

CHAPTER 5: BEYOND THE APP STORES: TOWARD DESIGNING A WEB-BASED MOBILE MENTAL HEALTH APPLICATIONS RECOMMENDER SYSTEM

In recent years, Mobile Mental Health Applications (MMHAs) have become increasingly popular. However, choosing the right app can be difficult due to the large number of options available and limited information about their features on online marketplaces. Online ratings and reviews offer a broad perspective different features and shortcomings but often fall short of providing comprehensive insights. While app guide platforms like Psyberguide, Mindapps, and MARS present expert evaluations, their integration of user feedback and preferences is limited. To address this issue, there is a need for an inclusive app recommender framework that systematically combines insights from both experts and users, presenting them to potential app users in an organized manner to aid in making informed decisions about app downloads. This study conducts a systematic review of existing platforms like Psyberguide and Mindapps and analyzes recent literature on user perspectives of MMHAs. We identify prevalent themes (such as cost, privacy, and user-friendliness) and discrepancies between what experts recommend and what users actually need (such as the integration of AI functionalities like chatbots, cultural inclusivity, and accessibility features). These findings can guide the development of a comprehensive web based MMHA recommendation tool. The proposed framework aims to prioritize users' needs and values when selecting an MMHA, thereby generating more engaging and effective app recommendations. By bridging the gap between popularity metrics and user experience, this framework will help users make informed decisions about which MMHA to use.

5.1 Introduction

In 2016, mental health conditions affected over a billion people globally, with depression in 2020 recognized by the World Health Organization as a leading cause of disability worldwide (100). Despite ongoing efforts to enhance accessibility, significant disparities in access to mental healthcare persist across all countries (32). In recent years, Mobile Mental Health Applications (MMHAs) have emerged as potential solutions to address these challenges (260). MMHAs, also known as mental health apps, are software applications designed for mobile devices like smartphones and tablets. They aim to support mental well-being, provide mental health information, and offer tools or interventions to manage and improve mental health (23). These apps cover a wide range of mental health aspects, including stress management, mindfulness, meditation, mood tracking, and access to mental health

resources (112). They can be utilized for self-help, as supplements to therapy, or alongside traditional mental health services (132). MMHAs are cost-effective, accessible, and scalable tools that have the potential to expand access to mental healthcare worldwide (132).

The rapid expansion of MMHAs has been fueled by the widespread ownership of smartphones and internet connectivity. Presently, there exist approximately 10,000 MMHAs geared towards mental health (112). However, despite their promising potential, many of these applications lack substantial evidence validating their effectiveness (27). The abundance of MMHAs often leaves users and clinicians grappling with the challenge of selecting the most suitable one. Users frequently rely on star ratings and reviews as guiding factors in their decision-making process (199). However, these metrics typically fail to offer a structured approach based on user priorities (48; 199). While star ratings and reviews can provide a useful starting point, they should not be the sole determinant.

To address this issue and enhance transparency regarding the clinical efficacy of these apps, researchers and practitioners have developed various evaluation frameworks (48). Examples include the American Psychiatric Association's APA App Evaluation Model (17), ORCHA Guide (245), MARS Evaluation Tools (251), PsyberGuide (One Mind PsyberGuide), and MIND Tools (65). These frameworks function as criteria-based tools or methods for assessing app quality. Although the concept of quality may differ across these frameworks, they commonly evaluate factors such as accessibility, empirical evidence base, acceptability, usability, safety, privacy, security, and overall helpfulness (48).

Despite recent efforts to streamline guidelines for mental health app selection, these guidelines still require careful evaluation of apps for security, credibility, and clinical efficacy (129). However, they may prove challenging for regular users seeking simpler information to make choices. Prior literature found that these frameworks, developed by mental health professionals, may overlook what day-to-day users truly desire (129). Prior studies also observed low agreement between users' and professional ratings, with users valuing features often overlooked by professional reviewers (48; 149; 146). As technology evolves, AI tools like chatbots are increasingly integrated into MMHAs, but user perspectives point out limitations in these new tools such as emergency support capabilities, considerations factors regarding diversity, equity, and inclusion (DEI) crucial for reducing mental health disparities among marginalized communities (224). This highlights the need for a universal framework that acknowledges both user and expert opinions.

5.2 objective of the Study

While frameworks developed by mental health professionals offer valuable knowledge, incorporating nuanced user perspectives can provide a holistic understanding, aiding both clinicians and users to make informed decisions when choosing a mental health app. This inspired us to conduct an empirical analysis of two prominent MMHA rating framework in relation with user priorities in MMHA requirement pointed out in relevant prior research articles. We organized our study based on the following two research questions:

RQ1: What are the similarities and disparities exist between established evidence-backed evaluation and assessment tools and users' preferences when selecting MMHAs?

RQ2: What are the essential factors or categories that should be integrated into the development of a comprehensive framework for recommending mobile mental health applications (MMHAs)?

RQ3: What are the mental models that users and experts utilize when deciding on mobile mental health applications?

5.3 Background

Users seeking MMHAs need clear and comprehensive information to make informed choices. Moreover, the rapid proliferation of MMHAs has highlighted the need for reliable frameworks to evaluate their safety and effectiveness (162). Prior literature acknowledges the needs and provides implications around how experts, policymakers, and researchers can provide guidance in this area (166). Following up these concerns, different organizations have developed such frameworks, each with its own focus and methodology. Numerous Assessment and Evaluation Tools such as frameworks, guidelines, rating systems, or app libraries, have been established worldwide (48). Nevertheless, these efforts encounter challenges. For instance, the NHS Apps Library (128), which evaluated apps based on specific criteria, was introduced but promptly withdrawn due to public backlash after reports surfaced regarding privacy and security vulnerabilities in several of the apps (128). Mobile App Rating Scale (MARS) (251) and Organization for the Review of Care and Health Applications (ORCHA) developed their own guideline tools focusing on aspects such as engagement and usability (170). MARS evaluates apps based on criteria including engagement, functionality, aesthetics, information accuracy, and subjective quality (251). Another prominent guidance framework is the American Psychiatric Association's APA App Evaluation Model, which prioritizes factors such as privacy, clinical foundation, and therapeutic goal. This model underwent refinement by a diverse

committee, aiming for objectivity and evidence-based assessment (17). However, evaluation frameworks such as MARS and the APA App Evaluation Model aim to guide users through app selection questions, but may lack clear metrics for decision-making (161; 194). They also lack the sufficient detail for users to determine an app's suitability and do not provide clear metrics to guide app choices (194).

To facilitate app evaluation, organizations have created databases or evaluation hubs, such as the MIND database (65), which streamline the process by categorizing apps based on evaluation criteria. It has been developed in collaboration with the American Psychiatric Association (APA) (17), focusing on five priority levels: background info, privacy and safety, evidence, ease of use, and data integration (161). This framework consists of 105 objective questions, aiming to provide a data-driven approach to evaluating health apps and enabling customization to meet individual needs. Crowdsourcing is also incorporated to ensure ongoing evaluation and identification of unsafe apps (161). While the framework may be comprehensive, efforts have been made to make it practical for clinicians and users, with a focus on usability and objectivity. Another such effort led to the creation of One Mind PsyberGuide (One Mind PsyberGuide), a nonprofit project established in 2013, aims to assist individuals in using technology for better mental health. Its evaluations focus on three key metrics: credibility, user experience, and transparency regarding data security and privacy practices (194). Instead of assigning an overall score, PsyberGuide presents these metrics separately to allow consumers to prioritize factors based on their individual needs and preferences (195).

However, despite recent efforts by a few evidence-based works such as PsyberGuide and MIND to streamline these guidelines or provide additional materials to support their use, these guidelines require careful consideration and evaluation of apps for security, credibility, and clinical efficacy. Consequently, they may prove even more challenging for general users who seek simpler information to make choices. Carlo et al. discovered that the most common criteria of these frameworks, such as Psyberguide, Mindtools, and ORCHA, often have notably discordant reviews. Moreover, these frameworks and guidelines are developed by mental health professionals and experts who possess clinical knowledge but may lack insight into what day-to-day users truly desire (48). Prior literature has shown significant differences in opinions between experts and users on several criteria. Hudson et al. found low agreement between users' and professional ratings, and they identified several features missing from professional ratings, such as customization and integration with daily life. They also highlighted that participants particularly valued certain aspects of mental health apps that appear to be overlooked by professional reviewers (129). This suggests that academic researchers and

clinicians designing frameworks and evaluating apps may be missing what consumers are actually seeking. With the rapid growth of technology and advancements in different AI tools such as chatbots, these technologies have made their way into intervention techniques supported by MMHAs. Prior literature has pointed out the limitations of these new tools through user perspectives (112). Additionally, Ramos et al. demonstrated that it remains unclear whether existing app assessment tools adequately address considerations of diversity, equity, and inclusion (DEI), which are particularly crucial for reducing mental health disparities among marginalized communities (224). This underscores the need for a universal framework that acknowledges both user and expert opinions. Frameworks developed by mental health professionals undoubtedly carry significant value and knowledge that should be considered when choosing a MMHA. However, incorporating user perspectives in accordance with that has the potential to reveal a complete picture that will help both clinicians and users make informed decisions in choosing a MMHA.

5.4 Methods

Our methods has been divided into two parts, First, to identify similarities and gaps between expert opinions and user preferences, we first analyzed the most recent developed app evaluation frameworks with an aim to guide users about different factors that should be considered while choosing a mobile mental health applications (MMHAs). then on the second part, we have done a semi-structured interview with 10 participants to understand how they choose MMHA so that the elements from their mental model can be translated to our final design of the web-based MMHA recommender tool.

5.4.1 Phase 1: Analyzing the most recent developed app evaluation frameworks and user perspectives from prior literature

Selection of Existing App Assessment and Evaluation Tools: To extract the criteria from expert opinions we have used most recent and user-friendly MMHA Assessment and Evaluation Tools - PsyberGuide (One Mind PsyberGuide) and MIND (65). PsyberGuide is based on Mobile App Rating Scale (MARS) (251) and MIND is based on American Psychiatric Association's APA App Evaluation Model (17) - two most popular and effective assessment tools according to prior literature. However both MARS and APA might not be very user-friendly. So MIND and PsyberGuide both took an attempt to streamlined the process. PsyberGuide is a non-profit organization dedicated to evaluating digital mental health tools. It provides impartial reviews and ratings of various mental health apps and online programs to help users make informed decisions about their mental health

care (One Mind PsyberGuide). PsyberGuide's evaluations are based on scientific research and expert input, offering users valuable insights into the effectiveness, usability, and privacy features of different digital mental health resources. PsyberGuide assesses digital mental health tools using various criteria, including effectiveness, credibility, usability, privacy and security, and overall user experience (One Mind PsyberGuide; 194). Additionally, expert opinions on themes such as Recommendations for Use/ Target Audience, Content/ Appropriateness of Content, Ease of Use and User Experience, Visual Design and User Interface, Appropriateness of Feedback, Ease of Account Management/ Cognitive Challenge, Scientific Basis, Qualitative Review of Program Efficacy, Estimate of Efficacy Relative to Similar Products, Cost, and Security/Privacy/Data Sharing are provided for certain apps (One Mind PsyberGuide; 194). A screenshot of PsyberGuide is depicted in Figure 8 (left). Similary, the MIND (Measurement Inventory for Digital Health) framework is a comprehensive tool designed to evaluate digital mental health interventions (65). It takes a multi-dimensional approach, covering domains such as clinical effectiveness, user engagement, safety and privacy, user experience, interoperability, and scalability (65; 161). In an effort to convey accurate information to users, MIND offers a qualitative assessment (descriptive) and quantitative ratings based on the criteria such as Conditions Supported, Treatment Approach, Inputs/Outputs, Accessibility, Clinical Foundation, Engagement, Cost, Privacy, Features, Inputs, Uses, and Developer Types (65; 161). This multifaceted evaluation framework aims to offer comprehensive insights into the effectiveness, usability, and credibility of digital mental health interventions, facilitating informed decision-making among users (65; 161). A screenshot of the evaluation page of an MMHA has been depicted in Figure 8 (right). PsyberGuide and MIND both offer app descriptions and screenshots sourced from mobile app marketplaces.

Selection of Prior Literature on User Perspectives Regarding MMHAs

To extract user opinions and requirements regarding MMHAs, we utilized existing literature where researchers analyzed user needs and viewpoints from various sources concerning MMHAs. Papers were acquired from five databases: the ACM Digital Library (1), IEEE Xplore (3), EBSCOHost Academic Search Premier (2), ScienceDirect (5), and PubMed (4). These databases provide insight on the subject from a variety of relevant subfields in this area, such as human-computer interaction (1; 3), medicine (5; 4), and general academic literature (2). We used specific databases rather than Google Scholar because Scholar uses a non-transparent algorithm (that is frequently changed/updated) to rank results, which limits the possibility to replicate this analysis.The lead author executed the review and analyzed the literature. All databases were searched between December

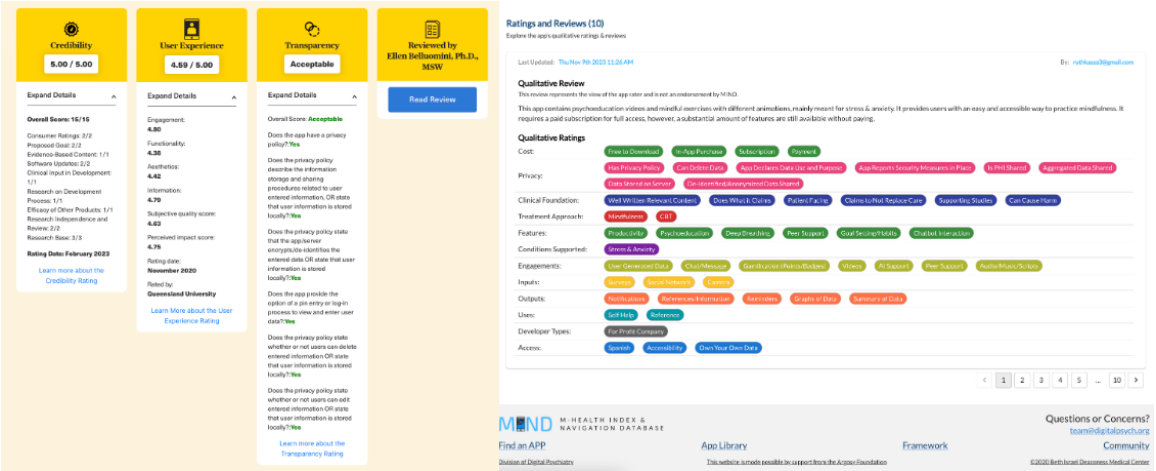


Figure 8: Screenshots of two leading and contemporary app evaluation tools: PsyberGuide (left) and MIND (right)

4 and December 17, 2024.

Query strings

We conducted a comprehensive search for literature containing insights into the selection process of mobile mental health applications (MMHAs), user perspectives on MMHAs, qualitative assessments of user requirements for MMHAs, and insights derived from user reviews of MMHAs. This involved exploring various scholarly papers and studies to gain a thorough understanding of how individuals engage with and evaluate MMHAs, aiming to gather valuable insights into user preferences, needs, and experiences with these applications. The search query string involved a combination of words of three different types: (1) words and phrases related to the mobile app and mental health, such as "mobile mental health applications" (2) words and phrases related to technology, such as "internet" or "technology;" and (3) words and phrases related to support, such as "users" or "intervention." At least one word or phrase from each type of the search query needed to be present. We brainstormed relevant keywords based on our experience with the previous literature and exploratory searches in each of the four databases. Some of the exemplary keywords we have used are "mobile mental health applications," "mental health apps," "user preferences," "user experiences," "user reviews," "qualitative assessment," "user requirements," "user perspectives," "mobile mental health application evaluation," "mobile mental health selection process," "user satisfaction," "app usability,"app effectiveness," "app features," and "user engagement." In order to enhance the relevance of our search results and access more comprehensive literature, we integrated the specific term "mobile mental health app" into our search queries. This strategic addition helped us refine our searches and target literature more directly related to our topic of interest. By incorporating this keyword

consistently, we aimed to ensure that the retrieved literature would be more closely aligned with our research objectives and provide valuable insights into mobile mental health applications.

Title	Authors(s)	Publication platform	Publication year
"For an App Supposed to Make Its Users Feel Better, It Sure is a Joke" - An Analysis of User Reviews of Mobile Mental Health Applications	Haque et al. (112)	Proc. ACM Hum.-Comput. Interact. 6, CSCW2	2022
A qualitative analysis of user reviews on mental health apps: who used it? for what? and why?	Thach et al. (259)	IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)	2019
Comparing Professional and Consumer Ratings of Mental Health Apps: Mixed Methods Study	Hudson et al (129)	JMIR Formative Research	2022
Potential and Pitfalls of Mobile Mental Health Apps in Traditional Treatment: An Umbrella Review	Koh et al. (152)	Journal of Personalized Medicine	2022
Stakeholders' views and opinions on existing guidelines on "How to Choose Mental Health Apps"	Khan et al. (146)	Frontiers in public Health	2023
What criteria are young people using to select mobile mental health applications? A nominal group study	Kabacińska et al. (136)	Digital Health	2022
Considerations of diversity, equity, and inclusion in mental health apps: A scoping review of evaluation frameworks	Ramos et al. (224)	Behavioral Research and Therapy	2021
Exploring how informed mental health app selection may impact user engagement and satisfaction	Kopka et al. (153)	PLOS Digital Health	2023
Methods for Navigating the Mobile Mental Health App Landscape for Clinical Use	King et al. (149)	Curr Treat Options Psychiatry	2023
Users' Adoption of Mental Health Apps: Examining the Impact of Information Cues	Huang et al. (127)	JMIR Mhealth Uhealth	2017
Insights from user reviews to improve mental health apps	Alqahtani and Orji et al. (16)	Health Informatics Journal	2020

Table 8: An overview of the peer-reviewed articles considered in this study

Inclusion/exclusion criteria

The search process was guided by a set of inclusion and exclusion criteria to ensure the relevance and quality of the retrieved literature. Firstly, papers considered for inclusion had to be published in or after the year 2000, thereby focusing on more recent research within the field. Additionally, papers were required to be written in English and undergo peer review to uphold academic standards. Furthermore, to align with the research focus on general mobile mental health applications, papers had to specifically address this topic. Importantly, user perspectives were prioritized, and therefore only papers that extracted insights from user studies or user reviews were included. Systematic literature review papers were deliberately excluded to focus on obtaining information directly from actual users, ensuring a more authentic representation of user opinions and experiences.

Selection process

By incorporating the keyword "Mobile mental health app" into each search query, we retrieved approximately 96 papers. Subsequently, we established inclusion and exclusion criteria, resulting in the exclusion of 39 papers. Titles and associated keywords for each remaining paper were reviewed, leading to the exclusion of an additional 19 papers. The remaining 19 papers underwent further screening through abstract review, resulting in the exclusion of 11 papers. Finally, these 11 papers were selected to extract user perspectives. The list of these papers have been presented in Table 8. It's important to note that our approach did not involve a full systematic literature search; rather, we aimed to reach theoretical saturation by selecting 11 papers that provided sufficient information. However, if necessary, we were prepared to repeat the process to identify any potentially overlooked papers.

Data Analysis

The first author diligently collected data throughout the review process, meticulously compiling it using an Excel spreadsheet. Initially, the researcher categorized all the criteria outlined in both the PsyberGuide and MIND frameworks (One Mind PsyberGuide; 65; 194; 161), providing succinct descriptions of each criterion to enhance clarity. Subsequently, user requirements and opinions were meticulously extracted and categorized from all ten papers, each category accompanied by a brief description clarifying its significance. The analysis proceeded in two stages. Initially, we performed open coding, generating over 30 codes to capture diverse viewpoints from the reviews, aiming to encompass nuanced insights. This number was then reduced through memoing and clustering (200). In the second phase, we utilized a constant comparison method, specifically affinity mapping, to cluster the codes based on their affinity within emerging themes. Following this categorization process, two authors collaborated to cross-reference all the categories. Through a thorough thematic analysis (200), prominent themes emerged, offering insights into the data extracted from the literature and existing frameworks.

Limitations of phase 1

Some biases were present during the selection process of the review. Papers were only included if written in English, which may exclude any number of relevant papers in various languages. Papers were also only included based on a subjective set of keywords set by the authors of this paper, allowing for bias to exist in either the selection of papers to be shown or included. Both authors have a background in computer science, which could introduce bias in both the selection of papers and the understanding of results within papers and overall during the analysis. Finally, only the

leading author read all papers and conducted the review; this could lead to subjective bias based on their understanding of both the review process and the literature.

5.4.2 Data Collection from Mobile Mental Health App (MMHA) Users

To understand the mental models guiding the selection of mobile mental health apps (MMHAs) by users, we conducted a semi-structured interview study with 10 participants (5 male and 5 female) in the USA. This study, approved by the Institutional Review Board (IRB), sought to identify the values and needs essential for designing a web-based recommender system for MMHAs.

Study procedure

We recruited a total of 10 participants (5 male and 5 female) with diverse backgrounds through a combination of personal contacts and snowball sampling (**?**). People who agreed to participate were interviewed based on in-depth semi-structured questions virtually using Microsoft teams platform.

The interview process comprised two distinct phases. Initially, participants were introduced to the study's objectives through a brief presentation. Following this, one of the researchers led an open discussion exploring participants' approaches to selecting mobile mental health applications. This discussion delved into their personal experiences, examining the factors influencing their decisions, elucidating their selection criteria, and uncovering how they became aware of these criteria.

In the second phase of the session, participants were presented with genuine user reviews sourced from mobile app stores such as the Apple App Store or Google Play Store, categorized using the ChatGPT API. Three overarching themes - privacy, cost, and ease of use - were selected based on a synthesis of expert opinions and user perspectives from prior literature. These reviews, provided in Table 10, aimed to solicit initial feedback on their usefulness.

The entire study duration lasted 60 minutes. Following the completion of the study, participants received a 30 USD Amazon gift card as compensation for their valuable time and participation.

Participants

Before participating in the interview, interested individuals were provided with a pre-screening questionnaire in which we inquired about their demographic information and their experience with using MMHAs. They were asked about the number of apps they had tried before, the duration of their usage, the purposes for which they had used them, and the factors motivating them to either continue using or discontinue using a particular app. Through our pre-screening process, we identified 10 participants with diverse backgrounds who had each used 1-3 MMHAs before, for a

Participant	Age range (in years)	Ethnicity	Gender	Education qualification	Annual household income range (in USD)	Mental health diagnosis
P001	25-34	Asian	Female	Master's	25,000-50,000	Not diagnosed
P002	25-34	Hispanic	Female	No degree	25,000-50,000	Prefer not to disclose
P003	18-24	African American	Male	High school diploma	50,000-100,00	Clinically diagnosed
P004	25-34	Asian	Female	Bachelor's	Less than 25,000	Clinically diagnosed
P005	35-44	African American	Male	Master's	More than 150,000	Clinically diagnosed
P006	35-44	Asian	Male	Master's	Less than 25,000	Self diagnosed
P007	25-34	Asian	Male	Doctorate	50,000-100,00	Clinically diagnosed
P008	25-34	African American	Female	Bachelor's	More than 150,000	Clinically diagnosed
P009	18-24	African American	Female	Bachelor's	25,000-50,000	Self diagnosed
P010	25-34	Hispanic	Male	Bachelor's	50,000-100,00	Clinically diagnosed

Table 9: List of particapnts and their demographic details

minimum duration of at least one month. Among the 10 participants, 5 were male and 5 were female. Six participants fell within the age range of 25-34 years, 2 were aged 18-24 years, and 2 were aged 35-44 years. In terms of ethnicity, 4 identified as Asian, 4 as African American, and 2 as Hispanic. Regarding educational attainment, 4 held bachelor's degrees, 3 held master's degrees, 1 held a doctorate degree, 1 had a high school diploma, and 1 did not possess a formal degree. In terms of household income, 3 participants had incomes ranging from 25,000 USD to 50,000 USD, 3 had incomes ranging from 50,000 USD to 100,000 USD, 2 had incomes below 25,000 USD, and 2 had incomes exceeding 150,000 USD. Among the participants, 6 were clinically diagnosed, 2 were self-diagnosed, 1 was not diagnosed, and 1 preferred not to disclose their diagnostic status.

Data analysis

The interviews were audio-recorded. Two researchers from the team were present during the interview and took detailed notes of their involvements and comments. Subsequently, the audio recordings were transcribed and anonymized by two authors. The analysis comprised two phases. Initially, after becoming familiar with the data, we initiated the first phase by generating over 30 open codes to capture essential concepts from the participants. The relatively high number of open codes aimed to meticulously capture the nuanced insights within each discussion and was subsequently significantly

reduced through the processes of memoing and clustering (200; 147). In the subsequent phase of analysis, we utilized a constant comparison method, specifically operationalized as affinity mapping, to memo and cluster the codes. This involved comparing each open code to others, positioning them based on their affinity to emerging themes and clusters, and exploring relationships between these themes. The ongoing process of constant comparison and collaborative discussions among the authors, who were domain experts, ensured the refinement of themes, resulting in a comprehensive and coherent representation of the data. Ultimately, a clear narrative was crafted to convey the identified themes, substantiating the findings with illustrative quotations and examples.

Data Integrity and Ethical Considerations

Considering that the discussion revolved around sensitive issues, we made sure to carefully consider the ethical implications of our work and to protect the rights of the participants. All sessions were recorded with participants' consent. The data was anonymized, with each dataset assigned a random code number in place of any personal identifiers such as participants' names or other identifying information. The study was approved by the Institutional Review Board (IRB) of the authors' institution.

5.5 Results

Our analysis unfolded in three main stages. Firstly, we compared user needs with expert criteria within the app evaluation framework to identify commonalities. Next, we presented the discrepancies on the same criteria between user requirements with established evaluation criteria. Finally, we pinpointed gaps in the app evaluation frameworks from the user's perspective.

5.5.1 Similarities Between User Needs and Expert Criteria in App Evaluation Framework

In this section, we will provide a brief overview of the categories that app evaluation frameworks can effectively incorporate to meet user requirements. Given that these metrics are well-established and readily available in existing frameworks, our analysis will be concise.

Recommendations for Use/ target audience

Target audience and supported intervention techniques are crucial considerations when choosing a MMHA. Both PsyberGuide and the MIND framework emphasize the importance of providing information regarding the target audience and recommended use. They offer detailed breakdowns of

the intervention techniques utilized by MMHAs, such as mindfulness, symptom tracking, goal setting, and reframing negative thoughts, as well as the mental health concerns they support, including depression, anxiety, stress, and sleep disorders. These frameworks feature user-friendly interfaces with filter criteria, allowing users to select MMHAs based on their preferences for intervention techniques and targeted mental health concerns (One Mind PsyberGuide; 65) Previous literature also underscores the significance of these factors. Haque et al. and Thach et al. have highlighted the importance of understanding the target audience and intervention techniques, recognizing their role in effective mental health care (112; 259). Additionally, research by Koh et al. found that supported conditions are among the most preferred criteria identified by users in their surveys (152). Therefore, as essential metrics in the design of app guidance tools, these categories are crucial considerations, with both experts and users in agreement on their importance.

Ease of use and user experience:

Ease of use is a critical factor in the effectiveness of MMHAs, as it significantly impacts users' experiences and engagement levels. PsyberGuide and MIND framework both recognize that MMHAs should prioritize ease of use by ensuring intuitive navigation, smooth performance, and minimal disruptions from advertisements (194; 161). Customizability features should also be incorporated to allow users to personalize their experience and maintain control over their privacy settings. Clear instructions and user-friendly design elements are essential to enhance usability and promote engagement with MMHAs for effective mental health support (194; 161).

Kabacinska et al. found that Users commonly report that MMHAs should be intuitive and straightforward to navigate, with features easily accessible without extensive effort (136). For instance, participants in various studies noted that apps described as "very easy and simple" to use were preferred, with navigation that did not require significant effort to locate features (136; 16). Additionally, the performance and efficiency of MMHAs are paramount, with users expressing reluctance to use apps with technical issues such as glitches or slow loading times. Participants universally agreed that they would not continue using an app with bugs or performance issues, highlighting the importance of smooth and reliable app performance. Moreover, the presence of advertisements within MMHAs can negatively impact usability, particularly if they disrupt the user experience with intrusive pop-ups or take over the entire screen. While participants understood the necessity of advertising for app developers, they emphasized the need for ads to be relevant and non-intrusive, suggesting that ads tailored to mental health needs or local services could be more acceptable (112; 129). Hudson et al. found that customizability emerged as another key aspect of ease of use, with participants valuing

the ability to personalize settings within the app to better suit their preferences and needs. This includes options such as customizing colors and avatars, as well as adapting the app to specific user requirements. Participants also emphasized the importance of clear instructions within the app to aid navigation, particularly for more complex features. Moreover, the ability to customize notifications and control access to personal information within the app were identified as important features, allowing users to tailor the app to their individual preferences and privacy concerns. The consensus among both experts and users is clear regarding the criteria for app evaluation and assessment tools.

Accessibility

MIND's framework integrates offline usability, functionality, and accessibility features within its initial level, recognizing their significance as fundamental components of access. These attributes are pivotal for ensuring that individuals with diverse needs, including those with limited internet connectivity or disabilities, can effectively utilize mental health apps (161). Similarly, PsyberGuide emphasizes comparable criteria, such as considering the cognitive challenges associated with accessing app features (194). Prior literature also underscores the multifaceted nature of accessibility, encompassing factors like app affordability, language options, and accommodations for users with sensory impairments (136). Furthermore, the ease of app registration and account creation is crucial for facilitating user engagement (112). Ensuring accessibility is not only a prerequisite for app usage but also aligns with principles of inclusively and equitable access to mental health resources. By assessing the presence of these features, both MIND and PsyberGuide contribute to fostering a more accessible and user-friendly landscape for mobile mental health applications.

Engagement challenges and effectiveness

Engagement and effectiveness are fundamental criteria integrated within both the PsyberGuide and MIND frameworks, reflecting their significance in evaluating mobile mental health applications (MMHAs). PsyberGuide categorizes this aspect as a "qualitative review of program efficacy," offering a comprehensive assessment of effectiveness through expert evaluations (One Mind PsyberGuide). Additionally, PsyberGuide provides a quantitative summary of user experience, capturing insights into engagement levels (194). Conversely, MIND evaluates engagement through criteria focusing on various intervention techniques aimed at maintaining user commitment, such as peer support, gamification, motivational quotes, and multimedia content (65). Prior literature underscores the importance of providing information on engagement and effectiveness to build trust among users regarding MMHA efficacy. By incorporating these criteria, both frameworks contribute to enhancing

transparency and aiding users in selecting MMHAs that align with their engagement preferences and effectiveness expectations (259; 152).

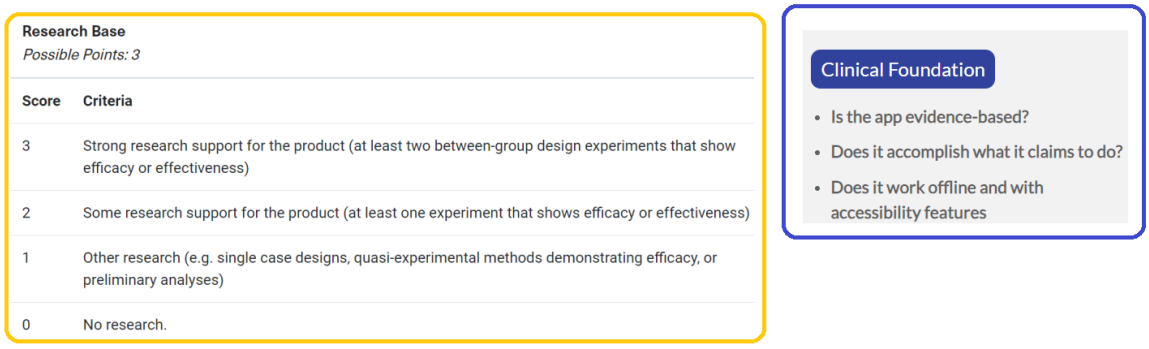
5.5.2 Discrepancies Between User Needs and Expert Criteria in App Evaluation Frameworks

In this subsection, we explore the elements highlighted by existing app evaluation frameworks, such as PsyberGuide (One Mind PsyberGuide) and MIND (65), which, despite their emphasis, do not always meet user needs or facilitate the app selection process effectively. We begin by outlining the evaluation criteria utilized by these frameworks, followed by identification of the gaps in these criteria from the standpoint of user experiences, as evidenced in previous studies. For each identified gap, we propose further considerations that should be included in a comprehensive app evaluation framework to better align with user expectations and requirements.

Inclusiveness of content across diverse demographics

As smartphones become ubiquitous, Mobile Mental Health Applications (MMHAs) are proliferating across diverse cultural and demographic groups. However, the rapid expansion prompts an important inquiry: are MMHAs suitable for users from varied backgrounds? Ramos et al. identified a deficiency in existing app evaluation frameworks, notably the absence of diversity, equity, and inclusion variables (224). This gap may hinder app-based interventions' ability to reach marginalized communities and could potentially exacerbate disparities. Haque et al. highlighted users' confusion over MMHAs' failure to tailor interventions for different age groups, revealing a generational gap that undermines effectiveness (112). Additionally, Hudson et al. demonstrated users' desire for personalized insights tailored to their mental health needs (129). In our research, we found that PsyberGuide introduced the concept of "appropriateness of content" to assess an app content's suitability through expert opinions (194). However, it overlooks nuances regarding whether MMHAs incorporate diversity, equity, and inclusion variables. Emphasizing cultural competence and gender inclusivity within app evaluation frameworks is essential to assess whether individuals from diverse cultural and demographic backgrounds feel supported and included when using these applications. This necessitates assessing whether content:

- Respects and accurately reflects diverse cultural values and beliefs.
- Offers support and resources tailored to specific cultural needs.
- Employs inclusive language that avoids bias and stereotypes.



Research Base
Possible Points: 3

Score	Criteria
3	Strong research support for the product (at least two between-group design experiments that show efficacy or effectiveness)
2	Some research support for the product (at least one experiment that shows efficacy or effectiveness)
1	Other research (e.g. single case designs, quasi-experimental methods demonstrating efficacy, or preliminary analyses)
0	No research.

Figure 9: Research evidence considered in the frameworks: PsyberGuide (left) and MIND (right)

- Provides gender-neutral or gender-affirming options and interventions.

Scientific evidence for effectiveness

Both the PsyberGuide and MIND evaluation frameworks emphasize the importance of incorporating a scientific foundation as one of the criteria for assessing MMHAs (194; 161). PsyberGuide rates an app on a scale from 3 to 0 based on strong to no research studies on app evaluation. MIND incorporates answers to three yes/no questions related to clinical foundation (see Fig. 9) (194; 161). However, from the users' standpoint, many users are uncertain whether these criteria are adequate for determining effectiveness. Koh et al. observed a lack of empirical support for the long-term efficacy of these interventions, as most studies concentrate on short-term results (152). Alqahtani et al. discovered that users find reassurance in hearing about positive real-world experiences from other users, suggesting that in addition to quantitative rating, evaluation frameworks should also include qualitative findings (e.g., positive comments from study participants) (16).

Merely considering short-term scientific experiments conducted behind the scenes may not provide a comprehensive understanding; thus, evaluation frameworks can instead concentrate on aligning user expectations with clinical evidence. This shift in focus requires:

- incorporating user-reported outcomes and real-world data alongside traditional scientific measures,
- including user-friendly summaries of clinical evidence within the app or evaluation framework,
- prioritizing studies that assess long-term efficacy and generalizability to real-world contexts.

Cost transparency

Existing frameworks typically address cost, indicating whether an app is free or offers paid versions.

While frameworks like MIND explore into subscription and purchase options (161), crucial details remain confusing. This limited transparency significantly impacts user adoption, particularly for individuals with lower socioeconomic backgrounds who seek affordable alternatives to traditional therapy. Research by Haque et al. highlights user dissatisfaction when they are required to enter personal information only to discover that most app functionalities are locked behind a subscription model. Users require clarity on what features are freely accessible, what necessitates payment, and if the cost aligns with the service's value proposition (112).

To empower informed decision-making and ensure accessibility, MMHA evaluation frameworks must move beyond a simplistic free/paid label by:

- **Justifying detailed app costs:** Provide a rationale for the app's pricing, detailing development costs, ongoing maintenance expenses, and the value proposition offered to users.
- **Unveiling the "Freemium" landscape:** Many apps employ freemium models with limited free features. Frameworks should disclose these constraints and clearly outline features accessible in the free version, including update frequency and limitations.
- **Expanding the cost conversation:** Users consider factors beyond the initial price tag, such as health insurance coverage, post-subscription charges, in-app purchases, unexpected fees, etc. Transparency regarding any potential fees should be explicitly mentioned upfront.

App integrity beyond organizational credibility

The MIND app evaluation framework introduced a criterion known as "developer types," aimed at informing users about the entities behind the app's development, such as government agencies, for-profit, or non-profit organizations (161). The intention behind this introduction is to build trust among users. However, users have expressed that providing additional information about organizational credibility would be more beneficial. Alqahtani and Orji et al. identified two crucial factors for earning user trust (16). Firstly, ensuring that the application undergoes regular updates and offers fresh content can reduce attrition rates. Users appreciate seeing new features like meditations, activities, backgrounds, pictures, quotes, and dashboard updates regularly, as it suggests the app remains interesting and engaging. Secondly, having responsive customer support is essential for users, as they expect quick and communicative assistance with their concerns (16). The integrity of the app and the credibility of the organization are closely intertwined. Users with mental health concerns rely on certain intervention techniques, and if the app breaks or fails to maintain their established lifestyle, it can be detrimental. To effectively address these considerations, MHA

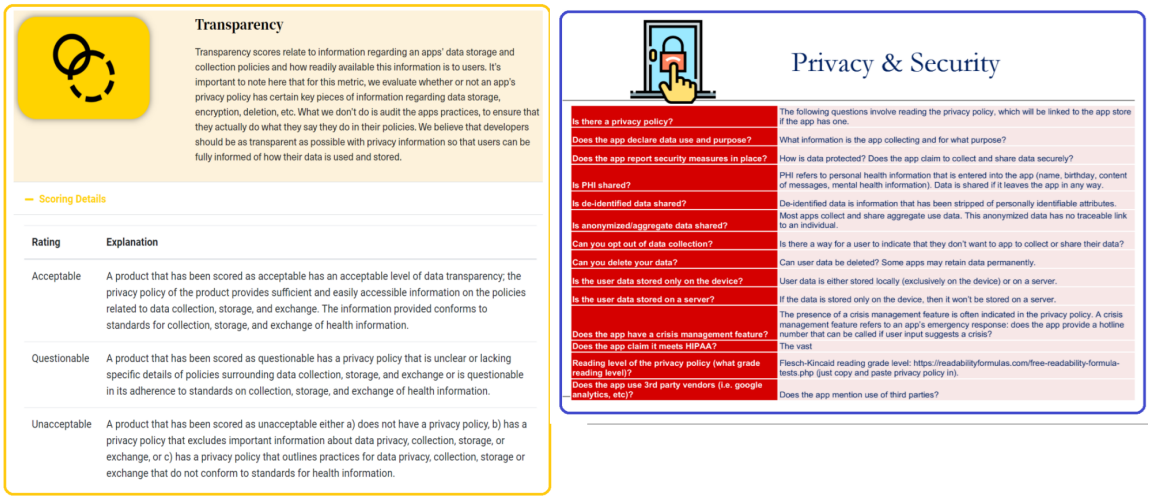


Figure 10: Privacy policy considerations in the frameworks: PsyberGuide (left) and MIND (right)

evaluation frameworks should be expanded to incorporate the following:

- **App update frequency and content freshness:** This may include the regularity of new feature introductions, focusing on examples relevant to user needs, such as meditation practices, activity types, and interface enhancements.
- **Customer support responsiveness:** Response times, communication effectiveness, and problem resolution rates would inform users about the level and quality of support available.
- **User-reported experiences of app integrity:** Gathering data on app functionality, stability, and potential disruptions to established routines could provide valuable insights into real-world app performance.

Beyond privacy policies

Privacy and security are essential criteria found in every app evaluation framework, including PsyberGuide (One Mind PsyberGuide) and the MIND framework (65). PsyberGuide goes further by not only providing expert opinions on policies but also offering quantitative ratings of various subcategories related to privacy. These include aspects such as whether the privacy policy describes information storage and sharing procedures, whether encryption or de-identification of data is implemented, whether pin entry or log-in processes are available for accessing user data, and whether users can delete entered information. Similarly, the MIND framework includes criteria related to the presence of a privacy policy, data sharing practices, and the option for users to opt out of data collection (Figure 10) (194).

Both frameworks offer comprehensive insights into the privacy policies of different mobile mental health applications (MMHAs). However, analyzing these policy recommendations requires a level of data literacy that some users may not possess. Additionally, given the global distribution of apps, it can be challenging to determine the applicability of these policies and whether they align with users' rights. Privacy policies established in one region, such as Europe's General Data Protection Regulation (GDPR) (**?**), may not necessarily meet the standards of users in other regions like the United States.

In light of these challenges, it is crucial to consider not only the existence and strength of privacy policies but also how users perceive their safety and security while using MMHAs. For instance, Haque et al. found that regardless of the presence of privacy policies, some users may still feel unsafe or violated while using certain MMHAs. Users may also express concerns about providing unnecessary personal information or granting permissions to certain features of their phones, which can make them feel vulnerable (112).

Therefore, in app evaluation criteria, a more holistic approach to incorporate insights into how users feel while using these apps can be adopted by:

- **In-depth analysis of negative reviews:** By qualitatively analyzing and presenting summaries of negative privacy-related reviews, MMHA evaluation frameworks can provide a nuanced understanding of whether users feel vulnerable or insecure while using a particular application.
- **Contextual considerations:** Recognizing diverse user experiences demands considering location, cultural background, and legal environment when interpreting reported experiences.

5.5.3 Missing Criteria in Existing Frameworks: User Perspectives

Our analysis of literature on user requirements and consumer reviews revealed crucial considerations missing from current mental health app evaluation frameworks. This subsection highlights these unmet user needs, advocating for their inclusion to bridge the gap between existing frameworks and real-world user experiences.

The capability of providing emergency support

Incorporating an evaluation criterion focused on emergency support provision within mobile mental health app frameworks is imperative due to the critical role it plays in aiding individuals during mental health crises. These features serve as lifelines during moments of vulnerability, offering users

reassurance and assistance when they need it most. Koh et al. mentioned by assessing whether an app provides comprehensive crisis-related information, including guidance on managing acute symptoms and accessing emergency services, users can ascertain the app's preparedness to address urgent mental health needs. Additionally, haque et al. found that the capability of these apps to assist users in creating personalized crisis plans enhances individuals' sense of agency and control over their mental health care (112). Furthermore, Alqahtani and Orji et al. found that users appreciate the ability to store family contacts for crisis communication ensures that users have immediate access to their support networks during times of distress, fostering a sense of connection and security (16). Moreover, apps that facilitate seamless connections with professional help services enable timely intervention and access to specialized care when required, potentially averting crises and promoting recovery. Furthermore, the inclusion of self-care tools within these apps equips users with practical strategies to manage their mental health autonomously, promoting resilience and coping skills (152; 16). Finally, apps that incorporate crisis detection mechanisms can provide early intervention by identifying signs of distress and prompting users to seek support proactively (152; 16). In essence, by evaluating these emergency support-related criteria, users can make informed decisions about which apps align best with their individual needs and circumstances, ultimately enhancing their ability to manage their mental health effectively and promoting overall well-being.

Therefore, app assessment and evaluation tools must be adapted to incorporate whether they encompass the following factors regarding the ability to provide emergency support.

- **Emergency support**: Apps must offer comprehensive crisis-related guidance and facilitate access to emergency services by offering offer comprehensive crisis-related guidance and facilitate access to emergency services. Including self-care tools equips users with strategies for managing mental health independently as well.
- **Seamless Professional Connections**: Apps should enable seamless connections with family members and professional help for crisis communication fosters a sense of connection and security.

Community and social support

The consideration of community guidelines and social environment within the app evaluation framework is paramount as it significantly influences users' experiences and ultimately aids in their selection of a suitable mobile mental health application. Alqahtani and Orji showed that community guidelines serve as the backbone of user interactions within the app, shaping the overall social en-

vironment and setting the tone for engagement (16). By examining community guidelines, users can gain insights into the platform's approach to moderating user behavior and fostering a safe and supportive online community. Haque et al. found some key factors from user review analysis such as the level of moderation strictness and the presence of moderator bias, provide valuable indicators of the platform's commitment to maintaining a positive and inclusive social environment (112). Users seek reassurance that they can engage with the community without fear of discrimination or predatory behavior, emphasizing the importance of feeling safe and respected within the app's social space. Furthermore, assessing the attitudes and values prevalent within the community offers users valuable information about the overall tone and culture of the platform beyond the written guidelines, enabling them to gauge whether it aligns with their values and preferences. By considering community guidelines and social environment as integral components of the evaluation framework, users are empowered to make informed decisions about which apps are suitable to foster a supportive and inclusive online community conducive to their mental well-being.

To offer a overview of the community and social support, app guidance tools should incorporate the following metrics to provide users with a comprehensive understanding of a particular MMHA.

- **Community Guidelines:** Serve as the foundation of user interactions within the app, shaping the social environment and engagement tone.
- **Moderation Standards and Bias:** Indicate the platform's commitment to maintaining a positive, inclusive atmosphere, gauged by the level of moderation strictness and absence of bias.
- **Safety and Respect:** Assurance of engagement free from discrimination or predatory behavior, underscoring the need for a safe, respectful space.
- **Attitudes and Values:** Insight into prevalent community attitudes and values beyond written guidelines aids users in aligning with their preferences.

Sensitivity towards vulnerable population

Individuals experiencing various mental health concerns are often considered vulnerable, requiring precise care and intervention techniques to facilitate effective recovery. Mishaps or ineffective tools can have detrimental effects, potentially exacerbating their condition or leading to serious consequences. As Haque et al. highlight, "one-size-fits-all" interventions can backfire, causing distress for certain individuals (112). App evaluation frameworks must prioritize sensitivity towards diverse

populations,including military veterans (consider specific needs related to PTSD, combat stress, and reintegration challenges), parental populations (address unique anxieties and pressures specific to parenthood), minors (ensure age-appropriate language, content, and safety measures), and others (include diverse groups like LGBTQ+, individuals with disabilities, and specific cultural considerations) (224). Certain vulnerable populations may be hesitant to share personal information or may not respond well to specific intervention techniques. Consequently, the evaluation framework should comprehensively outline how sensitive it is to the needs of these groups, empowering users to make informed decisions about selecting a MMHA that best caters to their unique requirements. By enhancing awareness of the nuances surrounding vulnerable populations, the evaluation framework can better serve diverse user bases and contribute to more inclusive and effective mental health care provision. To summarize the needs for vulnerable population, App evaluation frameowrk needs to consider following two categories -

- **Sensitivity to Diverse Populations:** Evaluation frameworks must prioritize sensitivity towards various vulnerable groups, including military veterans, parental populations, minors, LGBTQ+ individuals, those with disabilities, and specific cultural considerations.
- **Tailored Intervention Techniques:** Certain vulnerable populations may require customized intervention techniques and privacy measures, necessitating an evaluation framework that outlines its sensitivity to these needs.

The efficacy of AI tools

The advent of AI tools has revolutionized intervention techniques in MMHAs, offering customized and personalized features such as journaling, mindfulness exercises, symptom tracking, and goal setting. A recent trend in AI implementation is the introduction of chatbots, which emulate human conversation and provide tailored intervention suggestions while serving as conversational companions. However, despite their potential benefits, Haque et al. have highlighted limitations in these chatbots' interaction capabilities, potentially fostering over-reliance and social conservatism among users, and hindering their ability to provide emergency support (112). Given the limited evidence supporting these interventions, it is crucial for users to understand both their potential benefits and limitations. Managing users' expectations is essential, as evidenced by Haque et al.'s findings (112), which suggest that by tempering expectations, these tools can offer more accessible and affordable mental health care. Consequently, the app evaluation framework should provide comprehensive information regarding the efficacy and limitations of AI-based intervention tools. By equipping users

with this knowledge, they can make informed decisions about whether these tools provide the level of support they seek in an MMHA, ultimately enhancing their mental health care experience.

In summary, the following metrics are crucial in the guidance framework to ensure users can make a consistent choice.

- **AI Intervention Techniques in MMHAs:** Transperancy is expected where AI intervention tools are being used in intervention techniques with name and details of the intervention techniques.
- **Understanding Benefits and Limitations:** Users must be provided with a total picture of the potential benefits and limitations of AI-based intervention tools to manage expectations effectively.

Clear guidance and education

Through user review analysis, Haque et al. found that users highly value MMHAs that offer clear guidelines on app usage. They appreciate educational content provided by MMHAs, which helps them understand the efficacy of intervention techniques and why they are used (112). Similarly, Hudson et al. found that users appreciate informative content, such as articles and blog posts, within MMHAs, particularly those addressing symptom management and coping strategies for anxiety and depression. (129) This educational material enhances users' understanding of their mental health, providing both insights into anxiety and stress and a more technical understanding of these conditions. Alqahtani and Orji et al. also noted that providing education on mental health concerns and intervention techniques increases user engagement (16). Therefore, including clear guidelines and educational content in MMHAs is crucial for attracting and retaining users. Hudson et al. revealed that users valued engaging exercises and courses that prompted introspection and thought-provoking insights. Specific information on mental health status and coping strategies for adverse experiences was well-received. Guided journeys aided in identifying and addressing negative thought patterns, ultimately reducing feelings of anxiety and low mood. Users appreciated both coping exercises and the knowledge provided, which helped them understand and manage their anxiety effectively (129).Considering these factors as criteria in the app evaluation framework would facilitate informed decision-making among users, ensuring they select MMHAs that meet their needs effectively. Hence, App evaluation framework needs to be designed to include these factors, facilitating informed decision-making among users in selecting effective MMHAs.

- Presence of clear guidelines on app usage to enhance user understanding of intervention tech-

niques and

- Availability of informative content such as educational materials, articles and blog posts addressing symptom management and coping strategies for different mental health concerns.

Integration With daily life and impact on well-being

Our analysis found that users appreciate the fact how easily the interventions through MMHAs can be integrated with daily life and how much impact it can be on well being. Hudson et al. found that Well-being apps offer support to users throughout the day and across various locations, allowing for convenient scheduling according to individual preferences. Users appreciate features like specific meditations tailored for different times of day and content addressing their specific life circumstances, such as coping with challenges like the COVID-19 pandemic (129). While some users value the brevity of exercises, making them suitable for busy lifestyles, others find the time commitment unsustainable for long-term use. They also found that users appreciate the information to know that these MMHAs has brought positive changes in other users' well-being, such as experiencing reduced anxiety levels. According to Thach et al., users highlighted several factors associated with their well-being that influenced their perceptions of mental health apps (MMHAs). These factors include whether the apps contribute to feelings of happiness, positivity, self-awareness, calmness, focus, relaxation, motivation, mindfulness, and emotional control among users (259). Users also prioritize whether the apps enhance mood regulation, provide a sense of control over their lives, promote organizational skills, and increase awareness of mental health status. Furthermore, users value apps that help alleviate feelings of loneliness by fostering a sense of connection and support. Overall, users' experiences with MMHAs are characterized by positive emotions and an enhanced sense of mental well-being, which can serve as motivation for others to use these applications. Based on the preceding discussion, the following factors should be incorporated into app guidance tools.

- **Time commitments**: The anticipated duration necessary to complete various intervention techniques and their compatibility with hectic schedules.
- **Promoting positive life-styles:** users prioritize whether apps contribute to mood regulation, provide a sense of control over their lives, promote organizational skills, and increase awareness of mental health status.

Healthcare coordination

In certain instances, individuals utilize MMHAs as a supplementary support system alongside con-

ventional and formal treatments they receive. For example, clinicians may prescribe exercises to reframe negative thinking, track symptoms or goals, encourage journaling, or promote meditation, and MMHAs have demonstrated efficacy in facilitating these practices. Kabacińska et al. observed that users value the integration of MMHAs with the professional help they are receiving, appreciating the coordinated approach. However, users also expressed concerns regarding the sharing and extent of data, emphasizing the importance of user control in this regard (136). Similarly, Haque et al. noted that users appreciate the seamless transfer of information between healthcare providers and the MMHAs they utilize (112). Ensuring healthcare coordination, transparency, and user control over data sharing are essential factors for users to make informed decisions and maintain consistency in their mental health management. The following information should be conveyed to users to help them comprehend the capabilities of a specific MMHA in facilitating seamless coordination with healthcare providers.

- The integration capabilities of MMHAs with professional help, appreciating a coordinated approach to mental health management.
- Concerns regarding data sharing and user control over personal information, emphasizing the need for transparency and user autonomy in this aspect.

subsectionFindings

5.5.4 Strategies for finding MMHAs

In our investigation into the mental model of Mobile Mental Health Application (MMHA) users, we've found numerous pathways through which individuals come to know about these apps.

Online search and app marketplace

A large number of users rely on Google searches to explore options tailored to their specific needs. Once they compile a shortlist of recommendations, they typically take a glance at the descriptions provided in app stores to gain insights into each application's functionalities. This involves understanding the mental health symptoms addressed, the types of interventions or therapies offered, and how these align with their individual requirements. It's evident that users' needs vary considerably depending on their circumstances. Some seek calming videos, others prioritize access to professional therapy, while some are focused on symptom tracking and desire tailored activities like mindfulness or meditation.

> *" I look at the Play Store comment section where the people answer after from after first*

and second comment after three or four comments. There were I think people did really actually give actual comments so I follow them. And based on the star rating, I think I first approach to my first step. - P002"

Furthermore, our research underscores the significant role of visual elements in users' decision-making processes regarding MMHAs. To evaluate an app's suitability, individuals often scrutinize the screenshots provided, assessing factors such as visual appeal and ease of use. This visual exploration serves as a crucial determinant in their ultimate selection process, influencing whether they opt for a particular MMHA or continue their search elsewhere.

Recommendations from friends and family

Another way users discover specific applications is through recommendations from friends and family. Our study participants shared that during discussions about their mental health concerns, loved ones often shared their own experiences with MMHAs and suggested suitable options. Notably, these recommendations are especially valuable when users seek an application tailored to their demographic characteristics, such as gender, age, ethnicity, or nationality. Such personalized endorsements play a crucial role in fostering trust and confidence in the recommended MMHAs.

*"For example, I have a *MH* app that I **. So I think I need that app. I'm dependent on that. app. So at first I tried to figure out which app is the best. And then some of my friends suggested that this app you you are using this but I'm using that one and this one is really good. so I go for it" - P001*

Recommendations from mental health professionals

Moreover, users highly regard recommendations from mental health professionals, such as psychiatrists or psychotherapists. Participants disclosed that these professionals, equipped with a wealth of knowledge about mental health issues and adept at assessing individual situations, offer practical and beneficial suggestions. Additionally, users place significant trust in the judgment of these professionals, which in turn encourages them to take the next step towards utilizing the recommended MMHA. However, it's important to note that both personal and professional recommendations are contingent upon access to these support systems. Accessibility to such networks plays a pivotal role in the effectiveness of these recommendations.

" If psychotherapist or psychiatrist or someone who can help you professionally also recommend you this app that this when I am not around or when you are feeling or

you're feeling confused, you can go to this app and take help from" - P006

Through social media

Finally, our research uncovered the significant impact of social media in shaping users' decisions to adopt specific MMHAs. Participants disclosed that extensive advertising and marketing efforts for particular MMHAs played a pivotal role in piquing their interest and prompting them to "give them a try".

> *"So I looked in the internet, to any app, which can help me to refocus my thinking and keep me calm in maybe some sort of music or, or any idea or guidance so I started them, and then some recommendation came and also after that, now this is as it is the social media age. So after that, so many ads were coming in my Facebook and Instagram. These are the apps you can use to help yourself. So I chose them. Some of the apps were really repetitive, repetitive. Very much so kind of they were pushed, they were pushing me to download and try one. So I downloaded at first maybe comm app and I found it a little bit helpful" - P004*

Furthermore, users also discovered various MMHAs through online communities and groups, such as Facebook groups, Reddit, and Quora. Several participants mentioned actively engaging with these communities aligned with their preferences to explore diverse recommendations and gain insights into real-life experiences and use cases of MMHAs. By leveraging these platforms, users sought not only recommendations but also sought genuine opinions and firsthand accounts to inform their decision-making process. This illustrates the multifaceted role of social media in both introducing users to MMHAs and facilitating informed choices through community engagement.

> *"I asked about the apps in Reddit and some comments were suggesting me that you can go through these apps people who use them or have already have experience on them. I have seen their comments and and also, before downloading the app" - P006*

5.5.5 Potential of user reviews

Our research uncovered the potential of user reviews within app marketplaces to foster trust and credibility among users. Moreover, they serve as critical tools for evaluating various aspects such as privacy, safety, cultural and gender appropriateness, and cost-effectiveness. Lastly, participants emphasized the significant role of user reviews in determining their decision to adopt a particular app, empowering them with greater control over selecting a suitable MMHA.

Establishing trust and credibility

Our study participants highlighted the pivotal role of user reviews in fostering trust and credibility regarding specific MMHAs. They emphasized that user reviews serve as credible sources where genuine user experiences are shared, offering valuable insights into the real-life utility of these applications. Participants perceive user reviews as trustworthy channels for accessing authentic use cases from other users.

> *"I see how other people love experience. Then I download the app. Okay. Then use it." - P008*

> *"I feel like in a lot of the apps that I do like if I see a review saying that like they have stayed consistent and stuff like that. That's usually like a plus for me. Sometimes like Oh, go through the reviews and see like, Oh, they've changed their entire model like everything is different and then I'll probably stay away from that." - P002*

Interestingly, our participants indicated that, in most instances, they do not actively seek scientific evidence supporting these applications, either by choice or due to the absence of such information upfront before downloading. Moreover, they often lack clarity on what constitutes scientifically evident measures for an MMHA. In such scenarios, participants noted that relying on user reviews helps them establish credibility and confidence in the efficacy of the application. One participant aptly remarked that if an MMHA proves effective for six other users, it's likely to work for them as well, underscoring the significance they attribute to collective user experiences. This highlights the reliance on peer feedback as a crucial determinant in decision-making processes surrounding MMHAs.

> *"Okay, oh, I do that on my own actual research ... Yeah. So for me I would say depends amounts of users. You know, check the reviews and check the futures of these apps." - P005*

Communicating crucial information about MMHA

User reviews can provide valuable information to users, making the decision-making process easier. Some specific cases illustrate this potential.

Our study participants disclosed that while they highly prioritize privacy and regard it with utmost seriousness, they often find it inconvenient to thoroughly review privacy policies. Instead, they opt to rely on user reviews to navigate the privacy and safety aspects of an application. In doing so, they

specifically look for feedback from other users regarding concerns such as the extent of information collected, the overall safety of usage, and whether the application is perceived as violating privacy rights.

> *"reading the terms and conditions and the privacy policy would be out of convenience, just don't do it, because it seems kind of long process to go through, like, sort of like a legal document to go through." - P007*

> *"typically I don't read policies ... I usually just read the reviews and then download it" - P002*

Participants noted that user reviews serve as a valuable resource for understanding the privacy implications of an MMHA. They pay close attention to insights shared by fellow users regarding any perceived overreach in data collection, potential risks to personal information, and overall adherence to privacy standards. By considering the collective experiences and observations of other users, participants aim to make informed decisions regarding the privacy implications of using a particular MMHA. This highlights the significance of peer feedback in supplementing traditional privacy policy assessments and shaping users' perceptions of privacy and safety within the digital landscape.

Our research also uncovers certain information that may not be readily apparent in app descriptions or expert opinions, such as whether a particular app caters to specific demographics. Our participants emphasized the importance of understanding whether an app is suitable for different genders, ages, or cultural backgrounds. They highlighted the significance of gender and ethnic appropriateness to ensure that these apps are better tailored to meet individual needs.

> *"Every culture is kind of different from another. So they're the people of that culture. They also vary their mental outbreaks are very different one another." - P004*

> *"different different ethnic groups, experience, you know, different treatments and so, that itself can contribute to more stress just from a socio cultural perspective. And as such, I think it's important to, you know, cater to, you know, different people of color as well." - P003*

When asked where they could find this information, participants indicated that the primary source is community feedback or user reviews. Learning about the experiences of users from similar demographics instills confidence in new users to proceed with downloading and utilizing the app for their recovery journey. This highlights the value of peer experiences in providing insights into demographic

suitability and enhancing user trust in the effectiveness of mental health apps.

Another significant aspect highlighted by participants is the cost versus usability tradeoff, which they glean from user reviews. They emphasized that while they recognize the inevitability of paying for certain applications, they are willing to do so if the app offers genuine value and can positively impact their mental health.

> *" If that is the monthly or yearly subscription, I think I will go for the for the next paying portion because it's about my mental health health and as it is helping me to gain my sense. So why wouldn't I happy to spend few money on it? Right." - P004*

Now, they stressed that the only way to ascertain whether an app meets their expectations in relation to the cost is through user reviews. Additionally, participants noted that user reviews can alert them to any hidden fees or unexpected costs associated with an application.

> *"I try to see (in user reviews) if there are any hidden fees, that's all." - P006*

In essence, user reviews have the potential to convey information to new users that delves deeper than the surface-level cost visible in the app marketplace. By providing insights into the actual usability and value proposition of an app in relation to its cost, user reviews empower users to make informed decisions regarding their investment in mental health resources.

Reinforcing decision-making process

In the second phase of our interview study, we presented our participants with actual user reviews categorized by the ChatGPT API into three main categories: cost, privacy, and ease of use. These reviews were documented in Table 10. Our objective was to understand our participants' perspectives upon encountering such reviews and how it might influence their decision-making process regarding whether to download the application in question.

Interestingly, we observed that these reviews played a pivotal role in aiding participants' decision-making processes regarding app download. In each case, participants were able to identify specific aspects of the reviews that significantly contributed to their decision. For instance, Participant P007 highlighted how initially, the subscription cost seemed reasonable, but upon reading user reviews, it became apparent that crucial features were locked behind a paywall, leading them to reject the app.

> *" it doesn't seem like that it's too expensive to have a subscription to the app itself. But if everything is beyond a paywall, then I don't think I questioned the ethics of the app and the developers themselves." - P007*

Categories	Review number	User reviews categorized by ChatGPT API
Privacy	UR001	*"Ok so as a teenager its like really good to have a daily journal especially bc u can keep track of how ur doing on some days and u can keep track of how ur health is on this app too. Only complaint is that I have to have a premium subscription to get a pin number lock on the app?? It shouldn't be like that. If I don't want my cousins or nephews and nieces or anyone looking at this I would have to pay? Its not fair. PLEASE whoever works for this apps company let the password feature be unlocked for people without a subscription. Privacy is a very important thing and as I said, we should not have to pay for our own privacy."*
	UR002	*"I don't like apps that ask personal information and then have to link them to an email account. It's too stressful- no privacy- when it asked for my email I deleted it- never found out if it was good or not."*
	UR003	*"This app collects too much information about me that is not necessary for me to use their app. Look at their new privacy info in the App Store. I value my privacy and wish they didn't collect so much info and use it."*
Cost	UR004	*"But once it's up, almost everything is behind a paywall. If you guys want to sell a paid app, just do it up front. I know it sounds counter-intuitive to meditation, but I'd even watch ads before the sessions if that would help them offset production costs. Super expensive otherwise."*
	UR005	*"I like it a lot, but I wish I didn't have to pay for sessions. I have a problem with sleep, I can't get enough sleep so, that makes a big impact on school. I need and want to have a good education. I would buy it but I need money for other things. In conclusion, it's good but costs money."*
	UR006	*"Watch out for the auto renew option, it's not clear and canceling subscription was not easy, had to research a little since other services like Spotify, HBO Now, etc... make it very easy to unsubscribe.DO NOT TRUST developers that want to keep you engaged at any cost."*
	UR007	*"I love this app. The daily activities and games help me stay focused and positive. I went ahead and paid for the subscription after testing out the free version. For me, it's worth the cost of a cup of coffee a month for all the fantastic benefits."*
	UR008	*"I'm confused what the point of this app is? I know there is a premium version but i can't find any features that aren't locked. Even all of my recommended mediations, beginners mediations, I've scrolled through every category, all of it is locked. There is literally nothing ti do in the app if you don't pay. Why not just charge to download the app. All of your ads are very nice, but i can't even do a breathing exercise for free. By all means charge for your app but I would thing there would be at least something, anything*
Ease of use	UR009	*"App will log you out on its own after inactivity, but you can't log back in afterwards. Gives you an error message. Only way to log back in is to repeatedly uninstall and reinstall the app. What a waste. I was genuinely enjoying the content I'd gotten to use so far."*
	UR010	*"You don't know where to start and there is not a step by step guide. You just face a lot of videos/audios and will be confused. Not recommended."*
	UR011	*"I am really loving this app! it's easy to use and i like that there is a point system in place because I am a very rewards driven person and that helps me feel that I truly am making progress."*

Table 10: User reviews categorized by the ChatGPT API were shown to participants.

Similarly, Participant P001 noted that concerns raised in the reviews regarding excessive data collection and users' vulnerability made them uncomfortable, ultimately influencing their decision not to proceed with the app.

> *"It's because most of the reviews say that they're providing more information that the app developers might not need for this specific app. So it's very likely that they're taking additional information from you. And like not giving you enough privacy for your space." - P001*

These examples underscore the importance of user reviews in providing valuable insights that may not be readily apparent from app descriptions or expert opinions. By highlighting real user experiences and concerns, these reviews empower individuals to make more informed decisions about the applications they choose to integrate into their mental health management routines.

Furthermore, our participants disclosed that they can discern whether the initial set of reviews is sufficient or if they should explore into more reviews, thus alleviating concerns about making decisions based on a limited number of reviews. They explained that if they cannot find decisive factors upon initial review or if they remain unsure about an app, they adopt different approaches. Some participants, unaffected by a few reviews, opt to try out the application for a trial period or seek out additional reviews to gain a more comprehensive understanding of the situation.

> *"I would probably I would probably look for a few more reviews, and then see if the if there are more good, like positive or not. Because so far, what I'm seeing is that is kind of like App Usage issue in terms of, you know, the crashing after logging you out after an activity, that seems like a little bit of a problem. But there are like there is also a positive point. So I'm looking for like a little bit more positives. And if that was an isolated event or not. Maybe that's just happening to that specific person for some reason, or maybe the device that they're using. But if it's if I see a more positive reviews, I would probably download it. So I would look for some like a couple of more reviews." - P007*

We also observed varied opinions among participants regarding the same set of reviews. While a particular review might deter one user from proceeding with the app, it might prompt another user to give the application a try. For instance, Participant P006 expressed frustration about the inconvenience of the log-in and log-out process (**UR009**) in a specific app, considering how it might amplify their mental health concerns.

> *"we actually appreciate if we don't have to log in every time you will know and and there*

> *should be clear instructions about what to do next in the app. You know, if I if I get you I've told you earlier the people actually the people who are used will be using it. They might be going through a difficult time of his or her life. So actually, we don't want to get them more confused about using the app, you know." - P006*

Conversely, Participant P003 indicated that while they would take note of such issues, they would be willing to wait or check for recent updates or reviews to see if the problem has been resolved (**UR009**). They believe that such issues are not deal-breakers as they can typically be resolved by developers.

Similarly, opinions diverged on other reviews, such as one suggesting the inclusion of video instructions (**UR010**). While one participant agreed with the suggestion, another participant downplayed its significance.

Likewise, when it comes to reviews in the privacy category, opinions varied greatly. Some users expressed concern about the reviewS (**UR002, UR003**) that mentioned this app amount of information collected by a particular app (e.g., Participants P004 and P001), stating that they would not feel safe using it. However, a few participants (e.g. P006) acknowledged the necessity for apps to collect extensive data to provide personalized services, understanding the value of data in improving service quality.

> *"Okay, so actually, yes, the first two reviews actually I support with them because you know, the actual Yes, privacy is the topmost priority, as I've said, and they actually agree with them. Agree with me, I think, but I think to be able to give useful information on mental health. The mental health, the app has to gather much information, you know, because when you go to a mental health professional, you know, they ask questions, if, if the app doesn't ask questions. If the app doesn't have enough information there it can. It can't give you a proper counseling or proper, you know, helpful information, helpful feedback." - P006*

In summary, user reviews have the potential to convey crucial information not found elsewhere and play a vital role in reinforcing the decision-making process by empowering users. They provide users with agency and control over whether to proceed with an application, ultimately promoting informed decision-making.

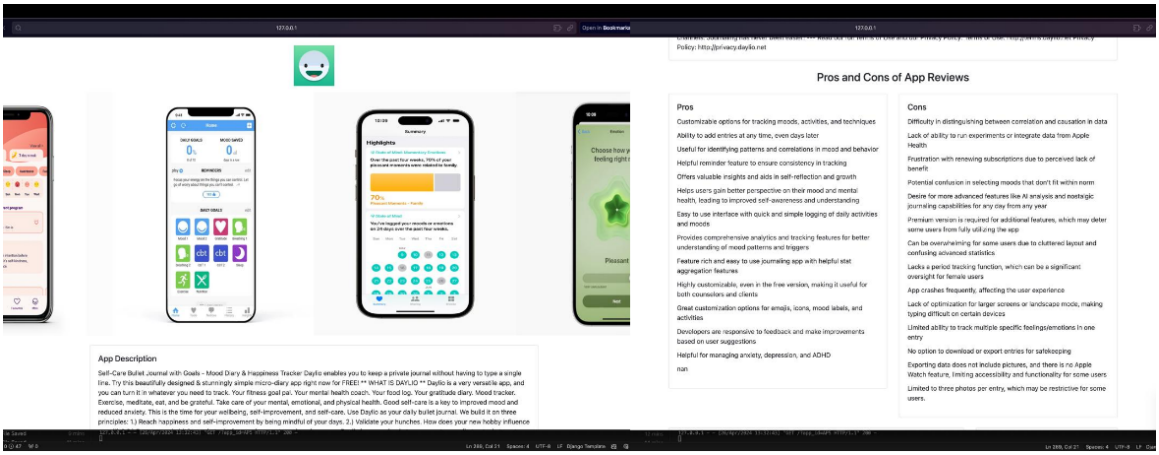


Figure 11: Images of the initial version of the Web-Based Recommender System for Mobile Mental Health Applications are displayed. On the right side of the screen, there are descriptions of the applications, while on the left side, there are lists of pros and cons derived from user reviews through the ChatGPT API.

5.6 Design a Web-Based Mobile Mental Health Applications Recommender System

For the final phase of the study, we design a pilot version of the web-based mobile mental health applications recommender system with one application called "Daylio". Daylio is a mobile application designed for diary and mood tracking, accessible on both iOS and Android platforms. At its core, Daylio allows users to log their mood using a five-point scale and optionally jot down accompanying journal entries. Users can also record their daily activities from a predefined list or create custom ones. Additionally, the app supports the inclusion of up to three photos with each diary entry. To encourage regular usage, Daylio employs various techniques such as a daily streak system, push notifications, and achievement rewards. It offers users the ability to visualize their mood data through graphs and statistics, including a monthly mood line graph and an average daily mood bar chart. These visualizations are dynamically updated as users input more data.

The web-based tool was constructed using a Flask-based application written in Python. Its front-end was crafted using HTML. Figures 11 and 12 showcase screenshots of the tool. Starting with the webpage walkthrough, the initial section displays screenshots sourced from various app marketplaces. This design choice emerged from our discussions with users, where we learned that visualizations play a crucial role in forming first impressions. Consequently, we prioritized these visuals at the top of the page. Following this, a condensed version of the app description, also sourced from app marketplaces, is provided. Utilizing the chatGPT 3.5 turbo API, renowned for its ability to produce

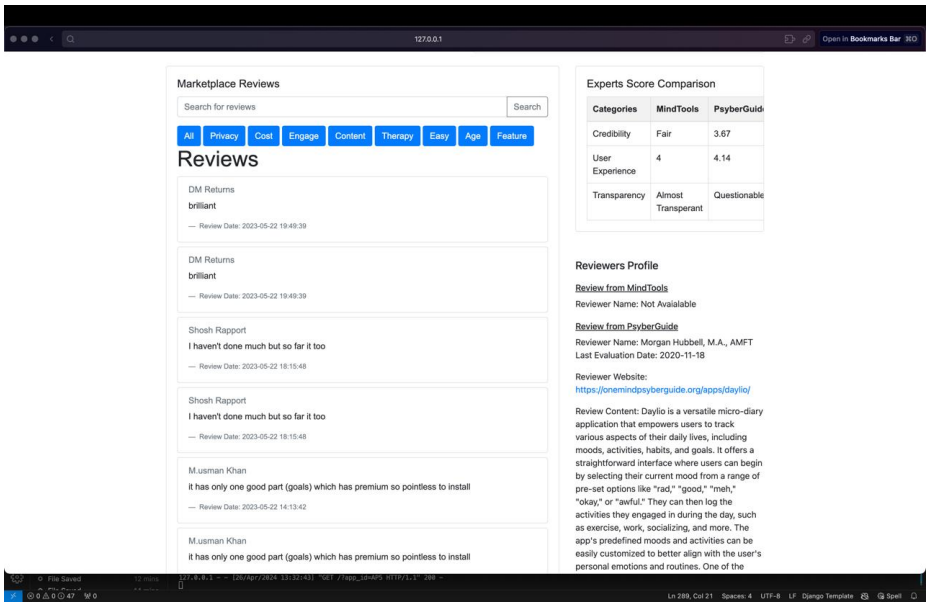


Figure 12: Images of the initial version of the Web-Based Recommender System for Mobile Mental Health Applications are displayed. Here, we can observe a visual layout where user reviews gathered from online marketplaces and expert opinions sourced from PsyberGuide are displayed together.

concise summaries, we opted for this tool to streamline the process. By leveraging chatGPT 3.5 turbo, we aimed to avoid unnecessary complexities associated with other language models. These app descriptions serve as a quick reference for users, allowing them to discern the targeted mental health concerns and prevention techniques offered by each MMHA.

Additionally, a list of pros and cons specific to each MMHA, derived from user reviews through the chatGPT 3.5 turbo API, is presented. To generate these pros and cons, we segmented 500 user reviews into chunks and extracted 10 pros and cons from each segment. Subsequently, through an iterative process, we identified common themes among these pros and cons, ensuring their accuracy and relevance. To maintain inter-reliability, manual verification was conducted alongside the automated process, validating the consistency of results provided by the chatGPT API.

Moving to the second part of the page depicted in Figure 12, user reviews from marketplaces and expert opinions from PsyberGuide and MINDTools are juxtaposed. Quantitative findings from expert opinions and qualitative reviews from experts are presented on the right-hand side, while the most recent marketplace reviews, collected between January 2020 and May 2022, are displayed on the left-hand side. Additionally, an open-ended search option, accompanied by suggestive tags generated from the chatGPT API, is provided. By clicking on any of these tags, users can automatically filter reviews based on specific themes or categories, facilitating ease of navigation and enabling informed

decision-making.

5.7 Discussion

In this section, we offer insights derived from our analysis regarding the necessity of adopting a user-centric approach in identifying app evaluation criteria. We also discuss the importance of striking a balance to present simplified information, aiming to create user-friendly app recommender systems.

Our analysis reveals significant gaps in current app evaluation frameworks like Psyberguide and the MIND framework, particularly in addressing user preferences and requirements. These frameworks often overlook the importance of diversity and inclusion, failing to ensure that mobile mental health applications (MMHAs) are culturally competent, gender-inclusive, and supportive of diverse user backgrounds. There's a critical need to shift from focusing solely on short-term scientific evidence to considering long-term effectiveness, user-reported outcomes, and real-world data. Additionally, transparency regarding costs, including development expenses, value propositions, and potential additional charges, is essential for user decision-making. Users also emphasize the importance of app integrity, desiring regular updates, responsive customer support, and stability to enhance their experience. Lastly, while privacy and security are crucial, addressing global privacy law variations and user safety perceptions is necessary for a more holistic and inclusive evaluation approach. Incorporating these user-centric considerations can bridge the gap between existing evaluation criteria and real-world user experiences, ensuring a more effective and inclusive digital mental health landscape.

These findings underscore the critical importance of integrating user perspectives into the design of app assessment and evaluation tools. Among the limited existing literature addressing how individuals select MMHAs, Khan et al. suggest that due to challenges related to digital literacy, collaboration between mental health professionals and primary stakeholders is essential to create user-centered approaches, particularly in understanding concepts surrounding privacy, security, diversity, and inclusion. To ensure user-friendliness, this knowledge must be conveyed in a manner comprehensible to individuals across varying levels of digital literacy (146). Furthermore, King et al. emphasize the significance of both clinician and user perspectives in effectively navigating the landscape of mobile mental health apps, highlighting the need for evaluating each app individually based on these varied viewpoints (149). Moreover, due to the discrepancies observed in reviews across rating frameworks, Carlo et al. recognize the importance of standardized behavioral health app guidance and the necessity of considering user preferences (48). Hudson et al. also found some variances when contrasting the requirements of professionals and consumers, emphasizing the im-

portance of platforms considering the perspectives of individuals with mental health issues when issuing reviews. This will ensure the reviews are highly relevant to those who are most inclined to utilize mental health applications (129). Given the identified discrepancies and gaps in contemporary evaluation and assessment tools regarding user opinions and requirements, our analysis underscores the necessity for successful navigation in app marketplaces and the disparities in consumer ratings. Prior literature has indicated that the absence of these factors can sometimes lead to adverse effects (112). Therefore, it becomes imperative to incorporate user opinions and preferences into evaluation and assessment tools to ensure both safety and user-friendliness. By integrating user perspectives, these tools can effectively address the nuanced needs and concerns of individuals, fostering a more holistic approach to app selection and ultimately enhancing the user experience.

The necessity of integrating feedback from both users and mental health experts (such as psychiatrists, psychologists, and therapists) into a comprehensive MMHA recommender system is evident. While user reviews are crucial for understanding personal experiences and influencing app selection, reliance solely on these can be problematic due to the prevalence of misleading reviews in the "wild west" of the MMHA marketplace, as described by Schueller et al. (194). This underscores the importance of exploring additional methods to gather user perspectives. Our analysis also highlighted several components crucial from expert opinions and existing assessment tools that are missing from user perspectives. Moreover, experts like healthcare professionals are used to suggesting suitable apps to their patients according to their custom needs. Therefore, a user-centric research to build a nuanced understanding of the aspects that users or experts or both prioritize is required for building future MMHA recommenders.

To develop a more user-friendly MMHA evaluation tool, it's essential to understand how individuals choose apps amidst the plethora of options and reviews. Investigating users' decision-making processes to create a cognitive map and addressing the specific needs of vulnerable populations can provide deeper insights for tool development. Moreover, the varied expectations users have of AI-based tools, as noted by Haque et al., highlight the need for research into user expectations to inform the design of future evaluation tools (115). By comprehensively understanding and incorporating user preferences and expectations, especially regarding AI, developers can create more effective and satisfying MMHA evaluation and recommendation systems.

5.8 Conclusion

This study sets the first cornerstone for developing an MMHA recommender systems: identifying app-related aspects from user and expert opinions that are not/poorly captured in the existing

tools. Moving forward, our goal is to develop a web-based recommendation system following the research and design implications described above. This application will prioritize presenting user perspectives in a more comprehensive and categorized format, aligning with the factors identified in our analysis. Moreover, our research has identified a notable gap concerning how users choose MMHAs. To address this, we first plan to investigate the mental models of both users and mental health professionals, exploring their decision-making processes in selecting the most suitable MMHA. The insights gleaned from these studies will be integrated into the application design, aiming to provide a more detailed and user-friendly experience. To conclude, this study conducted an extensive qualitative review of two prominent app evaluation frameworks, PsyberGuide and MIND, comparing them with user requirements extracted from 11 peer-reviewed articles. We explored both similarities and discrepancies between expert and user opinions on various criteria, while also highlighting gaps in existing app evaluation frameworks from the users' perspectives. These include the capability to provide emergency support, sensitivity towards vulnerable populations, and the efficacy of AI tools, among others. Our findings underscore the importance of adopting a user-centric approach in identifying app evaluation criteria and emphasize the need to strike a balance in design to simplify information for user-friendly app recommender systems.

CHAPTER 6: FINAL CHAPTER

6.1 Contributions

Mental health concerns affect nearly a billion people worldwide, with the COVID-19 pandemic exacerbating these issues. Mobile health technologies, including Mobile Mental Health Applications (MMHAs), have gained prominence as a means of providing non-emergency care and support. MMHAs offer various functions to address psychological concerns like anxiety and depression, including symptom assessment, tracking, and education on symptom management techniques. However, research on their effectiveness is still in its early stages, and many apps lack thorough evaluation. Challenges such as early abandonment, minimal user engagement, and usability issues contribute to their limited efficacy. There is a growing need to enhance our understanding of how mobile health apps are utilized in real-world settings.

Currently, there is limited knowledge about which apps are favored by consumers and whether popularity aligns with app quality. This underscores concerns about the subjectivity of contemporary app rating frameworks. Users face difficulty in choosing MMHAs due to the plethora of options available and rely on subjective star ratings and reviews, which may not align with their individual needs. This lack of guidance leads to high rates of app abandonment shortly after download. There are a few existing frameworks and expert evaluations could guide consumers and clinicians through the complex behavioral health app market. However, they often overlook consumer concerns and lead to inconsistencies in ratings. Definitive guidelines are needed to assist users and clinicians in making informed decisions about selecting MMHAs amidst the growing number of available options. consolidating these frameworks could enable more consistent decisions regarding mobile MH applications. To that end, we developed a web-based recommender system for mobile mental health apps. We created a comprehensive framework of app rating criteria integrating insights from user reviews and expert assessments. Additionally, we saught to understand the mental models guiding users' app selection process. This involved analyzing factors influencing decisions, exploring selection criteria, and understanding how users become aware of these criteria. Th findings ultimately resulted in a web-based interface for users to access data and make informed decisions about available MMHAs.

In order to create a web-based guidance tool for mobile mental health applications (MMHAs) that takes into account the viewpoints and feedback of users, we conducted a study involving the analysis

of user reviews from 164 MMHAs. The analysis underscores the critical importance of consistent moderation and community support within app-embedded communities. Poorly moderated environments can lead to biased behavior, harassment, and discrimination, undermining the intended supportive nature of these platforms. Future efforts should focus on tailored moderation policies, leveraging expertise and social control mechanisms to foster healthier online interactions. Moreover, automation via natural language processing could aid in content moderation, although challenges remain in ensuring fairness and transparency. Transparency regarding data collection and pricing is vital for user trust and satisfaction. Collaboration among stakeholders can help strike a balance between user desires and regulatory constraints. Developers should prioritize clear pricing structures and offer free trials to ensure users can assess the suitability of paid features. Affordable solutions should be accessible to all, regardless of insurance coverage. Improving usability and user experience (UX) is essential for mental health app efficacy. Users appreciate easy navigation, detailed explanations, and motivational content. Recommendation algorithms and transparent therapist qualifications can enhance user satisfaction and credibility.

Our analysis also revealed that the existing guidelines for developers of Mental Health (MH) apps are fragmented and lack comprehensiveness, leading to challenges in their adoption and implementation. To address this issue, efforts should be made to consolidate these guidelines into a centralized and accessible resource. Tools such as those developed by the Federal Trade Commission (FTC) and expert recommendations from organizations like the American Psychiatric Association (APA) and psyberguide can be integrated into interactive platforms accessible through common Integrated Development Environments (IDEs) to aid developers in navigating regulations and best practices. Additionally, there is a notable communication gap between users and developers, with user feedback often being disregarded. Promoting accountability by providing developers with snapshots of negative reviews and associated health consequences can help bridge this gap. Developers should also prioritize evidence-based approaches and conduct thorough testing before making apps publicly available. A comprehensive guideline for MH app development should include elements such as strategies for ensuring accountability in peer-to-peer support, enhancing transparency in app functionalities, adhering to data privacy regulations, and implementing automated approaches to predict and mitigate potential harm to users. Participatory research and design techniques should be employed to facilitate conversations among stakeholders and continuously evaluate app efficacy for ongoing improvement.

This study revealed that Chatbots have potential but require optimization for personalized, con-

textually relevant responses. Machine learning algorithms could assist in crisis detection and intervention, fostering safer interactions and enhancing user support. In order to gain a comprehensive understanding of the integration of chatbot technology, which has gained increasing popularity in recent years, we aimed to explore its implications within mobile mental health applications (MMHAs). Consequently, we initiated another study involving the analysis of user reviews from 10 MMHAs that utilize chatbot functionality. Our research underscores the significant potential of chatbots within mobile mental health applications as conversational companions, virtual friends, and immediate helpers. Users appreciate the human-like personality and communication skills of chatbots, which create a judgment-free environment for discussing their issues and concerns anytime, anywhere. However, there are several practical implications from our findings to enhance the user experience. Firstly, customization is crucial in MH chatbots to cater to individual user needs. Designers should consider the target age group and personalize interactions accordingly, avoiding elements like emojis and gifs that may not resonate with adult users. Additionally, incorporating user models and adapting interactions based on user types can enhance personalization. Furthermore, balanced persuasion strategies are essential to prevent user disengagement. While features like daily check-ins and reminders are helpful, excessive use may lead to users feeling guilt-tripped. Leveraging persuasive techniques to limit chatbot interaction and promote non-technical means of support can mitigate this issue. Building trust is paramount in MH chatbots, particularly regarding the collection and use of user information. Transparency about data collection practices and privacy policies is essential, along with providing information about the organization behind the app and the therapeutic methods used. This transparency can help establish a safe environment for user disclosure. However, it's crucial to recognize that chatbots should not replace human interaction for mental health support. While they can provide valuable support and guidance, they have limitations, particularly in crisis situations where human understanding and intervention are necessary. Setting clear boundaries and goals for chatbots, along with promoting in-person connections and professional support, is essential to manage user expectations and ensure their well-being.

Following the completion of our two studies, which shed light on various facets of user requirements and preferences concerning MMHAs, we aimed to engage directly with a vulnerable population. Military veterans represent one of the most vulnerable groups susceptible to mental health issues. Hence, we collaborated directly with military veterans and participated in the creation of design artifacts for mobile applications offering peer support to this demographic. Our initiative involved conducting a user-centered design investigation focused on developing persuasive technology, particularly

persuasive reminders, aimed at enhancing engagement and commitment to the treatment process facilitated by MMHAs. Traditionally, user-centered design follows a linear process of formative research, prototyping, and evaluation. However, in this study, an additional layer of validation before development was prioritized to ensure alignment with the needs and preferences of the community. This iterative approach emphasizes understanding the context and specifying user requirements comprehensively, validating language, motivational factors, and customization options before proceeding to evaluation.

Our findings contributed to context-specific design and the potential of semi-guided reminders in reducing cognitive loads on peer mentors within veteran peer-support interventions. The findings challenge established beliefs in the efficacy of personalized reminders using recipients' names as motivational tools, suggesting that this strategy might undermine trust in sensitive contexts such as mentor-mentee relationships involving USMVs with mental health issues. Additionally, a more straightforward communication strategy is advocated over gentle nudges, particularly with high-risk groups like veterans, whose responses to reminders may differ from those in traditional persuasion contexts. The implications for future research and design underscore the importance of context-specific design. Integrating veterans' perspectives into the design process is crucial for crafting reminders that resonate with this demographic. This involves avoiding certain terminology and employing tailored communication strategies that align with veterans' experiences. Furthermore, the study suggests a constrained approach to providing choices in reminders, focusing on simple, goal-oriented messages that reflect veterans' military backgrounds. The implementation of semi-guided reminders is proposed as a way to address cognitive loads while maintaining personalization. Peer mentors could customize communication styles and types of reminders, drawing from their deep understanding of their mentees' situations. Leveraging the "Battle Buddy" concept from the military, where peers support each other, can enhance peer support interventions by providing additional layers of support and intervention. Strategic areas where persuasive reminders can be employed to reduce cognitive loads on peer mentors are identified. By leveraging mobile technology to prompt actions such as check-ins and resource access, peer mentors can provide support more efficiently. Designing warning notifications for mentors regarding potential crises in mentees' lives and prompting mentees to seek immediate support adds an extra layer of safety to peer support interventions.

In our most recent study, we developed a web application. Our contribution to the design involved creating a framework for a guidance tool that considers input from both users and experts, as

well as gaining insights into the cognitive processes involved in users' selection of MMHAs. We found significant gaps in current app evaluation frameworks for mobile mental health applications (MMHAs), such as Psyberguide and the MIND framework, particularly in addressing user preferences and requirements. These frameworks often overlook the importance of diversity, inclusion, and long-term effectiveness, focusing solely on short-term scientific evidence. There's a need to shift towards considering user-reported outcomes, real-world data, and transparency regarding costs and app integrity. Privacy, security, and global privacy law variations are also crucial considerations for a more holistic evaluation approach.

Our analysis on identifying mental model on how users choose an MMHA revealed that integrating user perspectives into the design of app assessment and evaluation tools is critical. Collaboration between mental health professionals and primary stakeholders is essential to create user-centered approaches, particularly in understanding concepts surrounding privacy, security, diversity, and inclusion. Evaluating each app individually based on varied viewpoints from clinicians and users is necessary to effectively navigate the landscape of MMHAs. Standardized guidance and considering user preferences are vital to address discrepancies observed in reviews across rating frameworks.

The analysis underscores the necessity of successful navigation in app marketplaces and disparities in consumer ratings due to the absence of user opinions and preferences in evaluation and assessment tools. Incorporating user perspectives can address the nuanced needs and concerns of individuals, fostering a more holistic approach to app selection and enhancing the user experience. Integrating feedback from both users and mental health experts into a comprehensive MMHA recommender system is crucial. While user reviews are essential, additional methods to gather user perspectives are needed to address the prevalence of misleading reviews. Understanding users' decision-making processes and the specific needs of vulnerable populations can provide deeper insights for tool development. Research into user expectations, especially regarding AI-based tools, can inform the design of more effective MMHA evaluation and recommendation systems.

Integrating user needs and expert opinions into a holistic mental health app evaluation framework presents exciting opportunities, but also a crucial challenge: information overload. How do we present all this valuable information in a way that's easy to understand, rather than overwhelming users before they choose an app? We recommend a few different considerations:

- **NLP-powered summaries:** User reviews can be vast and varied. Leveraging NLP (Natural Language Processing) or LLMs (Large Language Models) to distill key points from numerous

reviews into concise, readable summaries can be a powerful tool (108).

- **Sentiment analysis:** By analyzing the common sentiment present in user and expert reviews, we can help users understand the overall 'vibe' surrounding different features of an app. This provides valuable insights beyond just factual summaries (36).
- **Side-by-side comparisons:** Empower users with informed decision-making by presenting user and expert opinions side-by-side. This allows for direct comparisons and a more nuanced understanding of different perspectives.

Moreover, the visual design should employ clear and intuitive visual elements like icons, charts, and color-coding to enhance information comprehension. It is also important to tailor the information presented based on individual user needs and preferences (e.g., search results based on both most frequent keywords or keywords prioritizing user-specific needs). The system should consider interactive features that allow users to explore specific topics or filter information based on their interests. By addressing these challenges and utilizing innovative design approaches, we can create user-friendly MMHA recommendation systems that empower individuals to make informed choices about their mental health apps. The goal is not just to provide information, but to do so in a way that empowers and guides these vulnerable users towards the right app suited for their unique needs.

6.2 Future Work

Our study has opened several promising avenues for future research. While our current project has introduced a pilot version of the recommender tool with a single application, we aim to expand its scope by incorporating a broader range of popular applications. This expansion will enable users to make more informed and consistent choices when utilizing the tool. To streamline the process with the introduction of a wider array of choices, our goal is to add a preliminary layer of recommendations before presenting user reviews and mental health professional reviews.

In future versions of the tool, we plan to include features that identify and recommend mobile mental health applications (MMHAs) based on users' specific symptoms and preferred intervention techniques. This enhancement will cater to the diverse needs and preferences of individuals seeking mental health support. Additionally, we aim to introduce criteria based on age, gender, and ethnicity. One of the key findings from our user study in Chapter 5 is that users prefer their mobile apps to be demographically aligned. To address this, we will first delve into the relevant demographic criteria for various communities, such as the Hispanic community and people from the global south. Our

goal is to integrate these criteria into the pre-selection process for users, providing them with a narrower choice of applications. This approach will allow them to delve deeper into user reviews and professional feedback to find the mobile mental health application best suited to their needs.

Following the development of the next iteration of the tool, it will be essential to conduct an evaluation study involving real users and mental health professionals. Mental health professionals possess extensive knowledge about what works for users and what might not. They also have a deeper understanding of not only what users typically prioritize when choosing an app but also what they should prioritize. Through our initial user study with mental health professionals, we learned that some mobile apps, though unpopular among users, are highly effective for individuals with specific symptoms.

Our goal will be to understand from both perspectives whether the comprehensive information about different applications influences their decision-making process. Evaluating the tool will also help identify key attributes that need to be incorporated to enhance its effectiveness. As a UX research scientist, I believe in making the design process iterative until the final product delivers an optimal user experience. The ongoing evaluation and refinement of the system to make it robust and effective will play a pivotal role in future research. This iterative approach will ensure that the tool not only meets users' needs but also integrates professional insights, ultimately improving mental health support.

In a more ambitious plan, we aim to make the system's API public, thereby creating opportunities for other researchers to build upon the concept we are promoting. The goal is to integrate user opinions and concerns into the recommender model, ensuring users are informed about what to expect from a usability perspective. A future objective is to engage with developers and marketplace owners, such as Apple and Google, to bridge the gap between user opinions and developers. Ideally, this recommender model would be incorporated directly into online marketplaces, providing users with comprehensive information in one place rather than requiring them to search multiple sources.

We also want to emphasize the sustainability of the project. One of the key attributes of this system is the continuous collection of knowledge from user reviews and mental health professionals, which evolves rapidly due to updates in applications or changes in user needs. To maintain the project's sustainability, our goal is to implement automatic updating methodologies or algorithms. This would ensure that the information in the system is refreshed every six months to a year, keeping it current and relevant.

Lastly, working with U.S. military veterans made me realize that different communities have unique needs, and it is crucial to personalize any tool we design according to those needs and values. As an individual from the global south, one of my personal research goals is to understand the requirements and values related to mobile mental health applications for people in the global south. Currently, these mobile apps and their features are primarily designed for individuals in developed countries. However, there is a significant need for such services in other parts of the world, which has not been adequately recognized or understood. My research aim is to amplify the voices from the global south and bring their opinions and concerns regarding mobile mental health applications to the forefront. By doing so, I hope to ensure that the design and functionality of these tools are inclusive and cater to the diverse needs of global populations, ultimately promoting better mental health support worldwide.

6.3 Conclusion

In conclusion, our project delved deeply into understanding user needs and perspectives regarding MMHAs, going beyond the traditional clinical framework. While clinical evidence remains essential for assessing the efficacy of mental health interventions like MMHAs, it's equally crucial to analyze users' opinions and real-life experiences. I advocate for research that involves multiple stakeholders, including both direct and indirect participants, as their input is vital for upholding key values in the design process.

To address this, our research initially focused on thoroughly examining user perspectives. We conducted an extensive analysis of user reviews from 164 mobile mental health applications sourced from platforms like the Google Play Store and Apple App Store. Additionally, we specifically investigated the integration of chatbots as a recent feature in MMHAs by analyzing user reviews of 10 chatbot-based applications. This analysis provided valuable insights into user acceptance of these innovative solutions and the features they prioritize.

To capture personal narratives detailing both successful and unsuccessful experiences with mental health apps, we conducted several studies involving US veterans, a group particularly vulnerable to mental health issues. Our findings highlighted deficiencies in existing apps, such as lacking visual appeal, inadequate privacy features, and limited support during critical moments, informing the direction for delivering digital interventions that better adapt to users' daily routines.

Furthermore, our research aligns with observations from Human-Computer Interaction (HCI) researchers, emphasizing the importance of a flexible, user-driven approach to digital mental health

interventions. This approach aims to tailor psychological interventions to better meet patients' mental health needs and daily lives.

Drawing on these insights, we designed an interactive tool to assist users in making informed decisions when selecting MMHAs. Through a web interface, user reviews and expert opinions are categorized based on preferences identified in previous studies. By providing access to both user and expert perspectives, our tool empowers users and healthcare providers to make consistent decisions about which MMHAs to choose.

Overall, our research contributes valuable insights and design implications, emphasizing the significance of user-centric design and evaluation, particularly for systems targeting vulnerable populations. Our study makes a significant contribution to this field by emphasizing the importance of incorporating user perspectives into the development of MMHA guidance tools. Moving forward, future evaluations will further refine our understanding of user needs and preferences, informing the design recommendations and practicality of such applications.

BIBLIOGRAPHY

[1] ACM Digital Library.

[2] EBSCOhost Research Platform | EBSCO.

[3] IEEE Xplore.

[4] PubMed.

[5] ScienceDirect.com | Science, health and medical journals, full text articles and books.

[6] Abd-Alrazaq, A. A., Alajlani, M., Alalwan, A. A., Bewick, B. M., Gardner, P., and Househ, M. (2019). An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132:103978.

[7] Abd-Alrazaq, A. A., Alajlani, M., Ali, N., Denecke, K., Bewick, B. M., and Househ, M. (2021). Perceptions and opinions of patients about mental health chatbots: Scoping review. *Journal of medical Internet research*, 23(1):e17828.

[8] Abras, C., Maloney-Krichmar, D., Preece, J., et al. (2004). User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4):445–456.

[9] Adaji, I. and Adisa, M. (2022). A review of the use of persuasive technologies to influence sustainable behaviour. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, pages 317–325.

[10] Adamopoulou, E. and Moussiades, L. (2020). An overview of chatbot technology. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 373–383. Springer.

[11] Ahmed, A., Ali, N., Aziz, S., Abd-Alrazaq, A. A., Hassan, A., Khalifa, M., Elhusein, B., Ahmed, M., Ahmed, M. A. S., and Househ, M. (2021). A review of mobile chatbot apps for anxiety and depression and their self-care features. *Computer Methods and Programs in Biomedicine Update*, 1:100012.

[12] Alahäivälä, T. and Oinas-Kukkonen, H. (2016). Understanding persuasion contexts in health gamification: A systematic analysis of gamified health behavior change support systems literature. *International journal of medical informatics*, 96:62–70.

[13] Albarracin, D. and Wyer Jr, R. S. (2000). The cognitive impact of past behavior: influences on beliefs, attitudes, and future behavioral decisions. *Journal of personality and social psychology*, 79(1):5.

[14] Alqahtani, F. and Orji, R. (2019). Usability issues in mental health applications. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, pages 343–348.

[15] Alqahtani, F. and Orji, R. (2020a). Insights from user reviews to improve mental health apps. *Health informatics journal*, 26(3):2042–2066.

[16] Alqahtani, F. and Orji, R. (2020b). Insights from user reviews to improve mental health apps. *Health informatics journal*, 26(3):2042–2066.

[17] American Psychiatric Association (APA), year = 2021, t. A. u. h. l. A.

[18] Andalibi, N., Flood, M. K., et al. (2021). Considerations in designing digital peer support for mental health: Interview study among users of a digital support system (buddy project). *JMIR mental health*, 8(1):e21819.

[19] Anita Nathania, P. and Yulia, V. Y. P. (2018). Android based chatbot and mobile application for tour and travel company. *IJCT*, page 21.

[20] Annapureddy, P., Franco, Z., Madiraju, P., Ahamed, S. I., Flower, M., Hossain, M. F., Haque, M. R., Johnson, N., Rubya, S., Baker, N. D., et al. (2021). Identifying precursors to long-term crisis in veterans using associative classifier. In *2021 IEEE International Conference on Big Data (Big Data)*, pages 4633–4642. IEEE.

[21] Annapureddy, P., Hossain, M. F., Kissane, T., Frydrychowicz, W., Nitu, P., Coelho, J., Johnson, N., Madiraju, P., Franco, Z., Hooyer, K., et al. (2020). Predicting ptsd severity in veterans from self-reports for early intervention: A machine learning approach. In *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*, pages 201–208. IEEE.

[22] Arean, P. A., Hallgren, K. A., Jordan, J. T., Gazzaley, A., Atkins, D. C., Heagerty, P. J., and Anguera, J. A. (2016). The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. *Journal of medical Internet research*, 18(12):e330.

[23] Aryana, B., Brewster, L., and Nocera, J. A. (2019). Design for mobile mental health: an exploratory review. *Health and Technology*, 9:401–424.

[24] Asensio-Cuesta, S., Blanes-Selva, V., Conejero, J. A., Frigola, A., Portolés, M. G., Merino-Torres, J. F., Almanza, M. R., Syed-Abdul, S., Li, Y.-C. J., Vilar-Mateo, R., et al. (2021). A user-centered chatbot (wakamola) to collect linked data in population networks to support studies of overweight and obesity causes: design and pilot study. *JMIR Medical Informatics*, 9(4):e17503.

[25] Ayanouz, S., Abdelhakim, B. A., and Benhmed, M. (2020). A smart chatbot architecture based nlp and machine learning for health care assistance. In *Proceedings of the 3rd International Conference on Networking, Information Systems & Security*, pages 1–6.

[26] Bae Brandtzæg, P. B., Skjuve, M., Kristoffer Dysthe, K. K., and Følstad, A. (2021). When the social becomes non-human: Young people's perception of social support in chatbots. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

[27] Bakker, D., Kazantzis, N., Rickwood, D., and Rickard, N. (2016). Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR mental health*, 3(1):e4984.

[28] Bartlett, Y. K., Webb, T. L., and Hawley, M. S. (2017). Using persuasive technology to increase physical activity in people with chronic obstructive pulmonary disease by encouraging regular walking: a mixed-methods study exploring opinions and preferences. *Journal of medical Internet research*, 19(4):e6616.

[29] Bates, M. (2019). Health care chatbots are here to help. *IEEE pulse*, 10(3):12–14.

[30] Battineni, G., Chintalapudi, N., and Amenta, F. (2020). Ai chatbot design during an epidemic like the novel coronavirus. In *Healthcare*, volume 8, page 154. MDPI.

[31] Bauer, M., Glenn, T., Geddes, J., Gitlin, M., Grof, P., Kessing, L. V., Monteith, S., Faurholt-Jepsen, M., Severus, E., and Whybrow, P. C. (2020). Smartphones in mental health: a critical review of background issues, current status and future concerns. *International journal of bipolar disorders*, 8(1):1–19.

[32] Becker, A. E. and Kleinman, A. (2013). Mental health and the global agenda. *New England Journal of Medicine*, 369(1):66–73.

[33] Beehler, S., LoFaro, C., Kreisel, C., Dorsey Holliman, B., and Mohatt, N. V. (2021). Veteran peer suicide prevention: A community-based peer prevention model. *Suicide and Life-Threatening Behavior*, 51(2):358–367.

[34] Bell, S., Wood, C., and Sarkar, A. (2019). Perceptions of chatbots in therapy. In *Extended*

Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, CHI EA '19, page 1–6, New York, NY, USA. Association for Computing Machinery.

[35] Bendig, E., Erb, B., Schulze-Thuesing, L., and Baumeister, H. (2019). The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review. *Verhaltenstherapie*, pages 1–13.

[36] Bhatt, A., Patel, A., Chheda, H., and Gawande, K. (2015). Amazon review classification and sentiment analysis. *International Journal of Computer Science and Information Technologies*, 6(6):5107–5110.

[37] Bhirud, N., Tataale, S., Randive, S., and Nahar, S. (2019). A literature review on chatbots in healthcare domain. *International journal of scientific & technology research*, 8(7):225–231.

[38] Björgvinsson, E., Ehn, P., and Hillgren, P.-A. (2012). Design things and design thinking: Contemporary participatory design challenges. *Design issues*, 28(3):101–116.

[39] Blomberg, D. J. and Karasti, H. (2012). Ethnography: Positioning ethnography within participatory design. In *Routledge international handbook of participatory design*, pages 106–136. Routledge.

[40] Bowie-DaBreo, D., Sas, C., Iles-Smith, H., and Sünram-Lea, S. (2022). User perspectives and ethical experiences of apps for depression: A qualitative analysis of user reviews. In *CHI Conference on Human Factors in Computing Systems*, pages 1–24.

[41] Brandtzaeg, P. B. and Følstad, A. (2018). Chatbots: changing user needs and motivations. *interactions*, 25(5):38–43.

[42] Brister, T. (2018). Navigating a mental health crisis: A nami resource guide for those experiencing a mental health emergency. *National Alliance on Mental Illness*, pages 1–25.

[43] Bucci, S., Berry, N., Morris, R., Berry, K., Haddock, G., Lewis, S., and Edge, D. (2019). "they are not hard-to-reach clients. we have just got hard-to-reach services." staff views of digital health tools in specialist mental health services. *Frontiers in psychiatry*, 10:344.

[44] Burns, J. (2013). Game on: Exploring the impact of technologies on young men's mental health and wellbeing.

[45] Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., Armour, C.,

and McTear, M. (2017). Towards a chatbot for digital counselling. In *Proceedings of the 31st International BCS Human Computer Interaction Conference (HCI 2017) 31*, pages 1–7.

[46] Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., Armour, C., and McTear, M. (2018). Assessing the usability of a chatbot for mental health care. In *International Conference on Internet Science*, pages 121–132. Springer.

[47] Cardie, C., Wiebe, J., Wilson, T., and Litman, D. J. (2003). Combining low-level and summary representations of opinions for multi-perspective question answering. In *New directions in question answering*, pages 20–27.

[48] Carlo, A. D., Ghomi, R. H., Renn, B. N., and Areán, P. A. (2019). By the numbers: ratings and utilization of behavioral health mobile applications. *NPJ digital medicine*, 2(1):1–8.

[49] Chambless, D. L. and Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual review of psychology*, 52(1):685–716.

[50] Chan, S., Torous, J., Hinton, L., and Yellowlees, P. (2015). Towards a framework for evaluating mobile mental health apps. *Telemedicine and e-Health*, 21(12):1038–1041.

[51] Chaovalit, P. and Zhou, L. (2005). Movie review mining: A comparison between supervised and unsupervised classification approaches. In *Proceedings of the 38th annual Hawaii international conference on system sciences*, pages 112c–112c. IEEE.

[52] Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354.

[53] Chinman, M., George, P., Dougherty, R. H., Daniels, A. S., Ghose, S. S., Swift, A., and Delphin-Rittmon, M. E. (2014). Peer support services for individuals with serious mental illnesses: assessing the evidence. *Psychiatric Services*, 65(4):429–441.

[54] Chinman, M., Goldberg, R., Daniels, K., Muralidharan, A., Smith, J., McCarthy, S., Medoff, D., Peeples, A., Kuykendall, L., Vineyard, N., et al. (2021). Implementation of peer specialist services in va primary care: a cluster randomized trial on the impact of external facilitation. *Implementation Science*, 16(1):1–13.

[55] Chinman, M., Oberman, R. S., Hanusa, B. H., Cohen, A. N., Salyers, M. P., Twamley, E. W., and Young, A. S. (2015). A cluster randomized trial of adding peer specialists to intensive case management teams in the veterans health administration. *The journal of behavioral health services & research*, 42:109–121.

[56] Chinman, M., Salzer, M., and O'Brien-Mazza, D. (2012). National survey on implementation of peer specialists in the va: implications for training and facilitation. *Psychiatric Rehabilitation Journal*, 35(6):470.

[57] Chuanromanee, T. and Metoyer, R. (2020). Evaluation and comparison of four mobile breathing training visualizations. In *2020 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 1–12. IEEE.

[58] Civan, A. and Pratt, W. (2007). Threading together patient expertise. In *AMIA annual symposium proceedings*, volume 2007, page 140. American Medical Informatics Association.

[59] Clemons, E. K., Gao, G. G., and Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of management information systems*, 23(2):149–171.

[60] Connolly, S. L., Miller, C. J., Koenig, C. J., Zamora, K. A., Wright, P. B., Stanley, R. L., and Pyne, J. M. (2018). Veterans' attitudes toward smartphone app use for mental health care: qualitative study of rurality and age differences. *JMIR mHealth and uHealth*, 6(8):e10748.

[61] Cox, K. E., Simonds, L. M., and Moulton-Perkins, A. (2021). Therapist-targeted googling: Characteristics and consequences for the therapeutic relationship. *Professional Psychology: Research and Practice*.

[62] Crits-Christoph, P., Rieger, A., Gaines, A., and Gibbons, M. B. C. (2019). Trust and respect in the patient-clinician relationship: preliminary development of a new scale. *BMC psychology*, 7(1):1–8.

[63] Cui, L., Huang, S., Wei, F., Tan, C., Duan, C., and Zhou, M. (2017). Superagent: A customer service chatbot for e-commerce websites. In *Proceedings of ACL 2017, system demonstrations*, pages 97–102.

[64] Darkins, A., Ryan, P., Kobb, R., Foster, L., Edmonson, E., Wakefield, B., and Lancaster, A. E. (2008). Care coordination/home telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. *Telemedicine and e-Health*, 14(10):1118–1126.

[65] Database, M. (2023). M-health index navigation database.

[66] de Alva, F. E. M., Wadley, G., and Lederman, R. (2015). It feels different from real life: users'

opinions of mobile applications for mental health. In *Proceedings of the annual meeting of the Australian special interest group for computer human interaction*, pages 598–602.

[67] Dekker, I., De Jong, E. M., Schippers, M. C., De Bruijn-Smolders, M., Alexiou, A., and Giesbers, B. (2020). Optimizing students' mental health and academic performance: Ai-enhanced life crafting. *Frontiers in Psychology*, 11:1063.

[68] Denecke, K., Vaaheesan, S., and Arulnathan, A. (2020). A mental health chatbot for regulating emotions (sermo)-concept and usability test. *IEEE Transactions on Emerging Topics in Computing*, 9(3):1170–1182.

[69] Devakumar, A., Modh, J., Saket, B., Baumer, E. P., and De Choudhury, M. (2021). A review on strategies for data collection, reflection, and communication in eating disorder apps. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–19.

[70] DiMarzio, J. (2016). *Beginning Android Programming with Android Studio*. John Wiley & Sons.

[71] Divya, S., Indumathi, V., Ishwarya, S., Priyasankari, M., and Devi, S. K. (2018). A self-diagnosis medical chatbot using artificial intelligence. *Journal of Web Development and Web Designing*, 3(1):1–7.

[72] Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., and Christensen, H. (2013). Smartphones for smarter delivery of mental health programs: a systematic review. *Journal of medical Internet research*, 15(11):e2791.

[73] Dow, K. and Cutter, S. L. (2000). Public orders and personal opinions: Household strategies for hurricane risk assessment. *Global Environmental Change Part B: Environmental Hazards*, 2(4):143–155.

[74] Drabek, T. E. and McEntire, D. A. (2003). Emergent phenomena and the sociology of disaster: lessons, trends and opportunities from the research literature. *Disaster Prevention and Management: An International Journal*.

[75] Drapeau, C. and McIntosh, J. (2016). U. sa suicide 2012: Official final data. washington, dc: American association of suicidology; 2014.

[76] Drebing, C. E., Reilly, E., Henze, K. T., Kelly, M., Russo, A., Smolinsky, J., Gorman, J., and Penk, W. E. (2018). Using peer support groups to enhance community integration of veterans in transition. *Psychological Services*, 15(2):135.

[77] Druss, B. G., Zhao, L., von Esenwein, S. A., Bona, J. R., Fricks, L., Jenkins-Tucker, S., Sterling, E., DiClemente, R., and Lorig, K. (2010). The health and recovery peer (harp) program: a peer-led intervention to improve medical self-management for persons with serious mental illness. *Schizophrenia research*, 118(1-3):264–270.

[78] Duan, W., Cao, Q., Yu, Y., and Levy, S. (2013). Mining online user-generated content: using sentiment analysis technique to study hotel service quality. In *2013 46th Hawaii International Conference on System Sciences*, pages 3119–3128. IEEE.

[79] Dyer, T. A., Owens, J., and Robinson, P. G. (2014). The acceptability of care delegation in skill-mix: the salience of trust. *Health Policy*, 117(2):170–178.

[80] Elmasri, D. and Maeder, A. (2016). A conversational agent for an online mental health intervention. In *International conference on brain informatics*, pages 243–251. Springer.

[81] Eslami, M., Vaccaro, K., Lee, M. K., Elazari Bar On, A., Gilbert, E., and Karahalios, K. (2019). User attitudes towards algorithmic opacity and transparency in online reviewing platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

[82] Fadhil, A. (2018). Can a chatbot determine my diet?: Addressing challenges of chatbot application for meal recommendation. *arXiv preprint arXiv:1802.09100*.

[83] Fadhil, A., Schiavo, G., Wang, Y., and Yilma, B. A. (2018). The effect of emojis when interacting with conversational interface assisted health coaching system. In *Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare*, pages 378–383.

[84] Fang, X. and Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2(1):1–14.

[85] Faurholt-Jepsen, M., Frost, M., Ritz, C., Christensen, E. M., Jacoby, A., Mikkelsen, R. L., Knorr, U., Bardram, J., Vinberg, M., and Kessing, L. V. (2015). Daily electronic self-monitoring in bipolar disorder using smartphones–the monarca i trial: a randomized, placebo-controlled, single-blind, parallel group trial. *Psychological medicine*, 45(13):2691–2704.

[Federal Trade Commission] Federal Trade Commission, year = 2016, t. M. u. h. l. A.

[87] 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):e7785.

[88] Fleming, T., Bavin, L., Lucassen, M., Stasiak, K., Hopkins, S., Merry, S., et al. (2018). Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *Journal of medical Internet research*, 20(6):e9275.

[89] Fogg, B. J. (2002). Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002(December):2.

[90] Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information systems research*, 19(3):291–313.

[91] Fortuna, K., Barr, P., Goldstein, C., Walker, R., Brewer, L., Zagaria, A., and Bartels, S. (2019). Application of community-engaged research to inform the development and implementation of a peer-delivered mobile health intervention for adults with serious mental illness. *Journal of participatory medicine*, 11(1):e12380.

[92] Franco, Z., Baker, N., Okusanya, T. R., Haque, M. R., Gresser, J., Rubya, S., Madiraju, P., Curry, J., Winstead, O., Hooyer, K., et al. (2023). Customizing the battlepeer app: Connecting first responders with peer support to manage mental health crises.

[93] Franco, Z., Hooyer, K., Roushan, T., O'Brien, C., Johnson, N., Watson, B., Smith-Watson, N., Semaan, B. C., Flower, M., Tasse, J., et al. (2018). Detecting & visualizing crisis events in human systems: an mhealth approach with high risk veterans. In *ISCRAM*.

[94] Franco, Z., Hooyer, K., Ruffalo, L., and Frey-Ho Fung, R. A. (2021). Veterans health and well-being—collaborative research approaches: Toward veteran community engagement. *Journal of Humanistic Psychology*, 61(3):287–312.

[95] Franco, Z. E., Logan, C., Flower, M., Curry, B., Ruffalo, L., Brazauskas, R., and Whittle, J. (2016). Community veterans' decision to use va services: A multimethod veteran health partnership study. *Progress in community health partnerships: research, education, and action*, 10(1):31–44.

[96] Fu, B., Lin, J., Li, L., Faloutsos, C., Hong, J., and Sadeh, N. (2013). Why people hate your app: Making sense of user feedback in a mobile app store. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1276–1284.

[97] Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., Rauws, M., et al. (2018). Using psychological

artificial intelligence (tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health*, 5(4):e9782.

[98] Garg, S., Williams, N. L., Ip, A., and Dicker, A. P. (2018). Clinical integration of digital solutions in health care: an overview of the current landscape of digital technologies in cancer care. *JCO clinical cancer informatics*, 2:1–9.

[99] Geuens, J., Geurts, L., Swinnen, T. W., Westhovens, R., Van Mechelen, M., and Abeele, V. V. (2018). Turning tables: A structured focus group method to remediate unequal power during participatory design in health care. In *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial-Volume 2*, pages 1–5.

[100] Ginis, K. A. M., van der Ploeg, H. P., Foster, C., Lai, B., McBride, C. B., Ng, K., Pratt, M., Shirazipour, C. H., Smith, B., Vásquez, P. M., et al. (2021). Participation of people living with disabilities in physical activity: a global perspective. *The Lancet*, 398(10298):443–455.

[101] Gorwa, R., Binns, R., and Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1):2053951719897945.

[102] Gould, C. E., Kok, B. C., Ma, V. K., Zapata, A. M. L., Owen, J. E., and Kuhn, E. (2019). Veterans affairs and the department of defense mental health apps: A systematic literature review. *Psychological Services*, 16(2):196.

[103] Gournaris, M. J. et al. (2016). Certified peer support specialists: Advancing peer support services in deaf mental health care. *JADARA*, 50(1):1.

[104] Green, J., Huberty, J., Puzia, M., Stecher, C., et al. (2021). The effect of meditation and physical activity on the mental health impact of covid-19–related stress and attention to news among mobile app users in the united states: Cross-sectional survey. *JMIR mental health*, 8(4):e28479.

[105] Greer, S., Ramo, D., Chang, Y.-J., Fu, M., Moskowitz, J., Haritatos, J., et al. (2019). Use of the chatbot "vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. *JMIR mHealth and uHealth*, 7(10):e15018.

[106] Gui, X., Chen, Y., Kou, Y., Pine, K., and Chen, Y. (2017). Investigating support seeking from peers for pregnancy in online health communities. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–19.

[107] Gupta, J., Singh, V., and Kumar, I. (2021). Florence-a health care chatbot. In *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 1, pages 504–508. IEEE.

[108] Guzman, E., Oliveira, L., Steiner, Y., Wagner, L. C., and Glinz, M. (2018). User feedback in the app store: a cross-cultural study. In *2018 IEEE/ACM 40th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, pages 13–22. IEEE.

[109] Haddad, S. M., Souza, R. T., and Cecatti, J. G. (2019). Mobile technology in health (mhealth) and antenatal care–searching for apps and available solutions: a systematic review. *International journal of medical informatics*, 127:1–8.

[110] Haldar, S., Mishra, S. R., Khelifi, M., Pollack, A. H., and Pratt, W. (2017). Opportunities and design considerations for peer support in a hospital setting. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 867–879.

[111] Haque, M. R., Franco, Z., Hossain, M. F., Frydrychowicz, W., Madiraju, P., Baker, N. D., Hooyer, K., Ahamed, S. I., Winstead, O., Curry, R., and Rubya, S. (2023). Perceptions of mental health crisis among u.s. military veteran peer mentors and potential of mobile-based peer-support interventions. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '23 Companion, page 33–38, New York, NY, USA. Association for Computing Machinery.

[112] Haque, M. R. and Rubya, S. (2022a). " for an app supposed to make its users feel better, it sure is a joke"-an analysis of user reviews of mobile mental health applications. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–29.

[113] Haque, M. R. and Rubya, S. (2022b). "for an app supposed to make its users feel better, it sure is a joke" - an analysis of user reviews of mobile mental health applications. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2).

[114] Haque, M. R. and Rubya, S. (2022c). "for an app supposed to make its users feel better, it sure is a joke" - an analysis of user reviews of mobile mental health applications. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2).

[115] Haque, M. R. and Rubya, S. (2023). An overview of chatbot-based mobile mental health apps: Insights from app description and user reviews. *JMIR mHealth and uHealth*, 11(1):e44838.

[116] Harjumaa, M. and Oinas-Kukkonen, H. (2007). Persuasion theories and it design. In *Persuasive*

Technology: Second International Conference on Persuasive Technology, PERSUASIVE 2007, Palo Alto, CA, USA, April 26-27, 2007, Revised Selected Papers 2, pages 311–314. Springer.

[117] Harrison, V., Proudfoot, J., Wee, P. P., Parker, G., Pavlovic, D. H., and Manicavasagar, V. (2011). Mobile mental health: review of the emerging field and proof of concept study. *Journal of mental health*, 20(6):509–524.

[118] Hartzler, A. and Pratt, W. (2011). Managing the personal side of health: how patient expertise differs from the expertise of clinicians. *Journal of medical Internet research*, 13(3):e62.

[119] Hawkins, R. P., Kreuter, M., Resnicow, K., Fishbein, M., and Dijkstra, A. (2008). Understanding tailoring in communicating about health. *Health education research*, 23(3):454–466.

[120] He, D., Pan, M., Hong, K., Cheng, Y., Chan, S., Liu, X., and Guizani, N. (2020). Fake review detection based on pu learning and behavior density. *IEEE Network*, 34(4):298–303.

[121] Himmelfarb, N., Yaeger, D., and Mintz, J. (2006). Posttraumatic stress disorder in female veterans with military and civilian sexual trauma. *Journal of Traumatic Stress: Official Publication of The International Society for Traumatic Stress Studies*, 19(6):837–846.

[122] Hoch, D. B., Norris, D., Lester, J. E., and Marcus, A. D. (1999). Information exchange in an epilepsy forum on the world wide web. *Seizure*, 8(1):30–34.

[123] Hollis, V., Pekurovsky, A., Wu, E., and Whittaker, S. (2018). On being told how we feel: how algorithmic sensor feedback influences emotion perception. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–31.

[124] Hoon, L., Vasa, R., Schneider, J.-G., and Mouzakis, K. (2012). A preliminary analysis of vocabulary in mobile app user reviews. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, pages 245–248.

[125] Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '04, page 168–177, New York, NY, USA. Association for Computing Machinery.

[126] Huang, C.-Y., Yang, M.-C., Huang, C.-Y., Chen, Y.-J., Wu, M.-L., and Chen, K.-W. (2018). A chatbot-supported smart wireless interactive healthcare system for weight control and health promotion. In *2018 IEEE international conference on industrial engineering and engineering management (IEEM)*, pages 1791–1795. IEEE.

[127] Huang, H.-Y., Bashir, M., et al. (2017). Users' adoption of mental health apps: examining the impact of information cues. *JMIR mHealth and uHealth*, 5(6):e6827.

[128] Huckvale, K., Prieto, J. T., Tilney, M., Benghozi, P.-J., and Car, J. (2015). Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC medicine*, 13(1):1–13.

[129] Hudson, G., Negbenose, E., Neary, M., Jansli, S. M., Schueller, S. M., Wykes, T., Jilka, S., et al. (2022). Comparing professional and consumer ratings of mental health apps: mixed methods study. *JMIR formative research*, 6(9):e39813.

[130] Huh, J., Marmor, R., and Jiang, X. (2016). Lessons learned for online health community moderator roles: a mixed-methods study of moderators resigning from webmd communities. *Journal of medical Internet research*, 18(9):e247.

[131] Hundt, N. E., Robinson, A., Arney, J., Stanley, M. A., and Cully, J. A. (2015). Veterans' perspectives on benefits and drawbacks of peer support for posttraumatic stress disorder. *Military medicine*, 180(8):851–856.

[132] Hwang, W. J., Ha, J. S., and Kim, M. J. (2021). Research trends on mobile mental health application for general population: a scoping review. *International Journal of Environmental Research and Public Health*, 18(5):2459.

[133] Inkster, B., Sarda, S., Subramanian, V., et al. (2018). An empathy-driven, conversational artificial intelligence agent (wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth*, 6(11):e12106.

[JingYi] JingYi, S. Y. The chatbot as a medium of communication.

[135] Judson, T. J., Odisho, A. Y., Young, J. J., Bigazzi, O., Steuer, D., Gonzales, R., and Neinstein, A. B. (2020). Implementation of a digital chatbot to screen health system employees during the covid-19 pandemic. *Journal of the American Medical Informatics Association*, 27(9):1450–1455.

[136] Kabacińska, K., McLeod, K., MacKenzie, A., Vu, K., Cianfrone, M., Tugwell, A., and Robillard, J. M. (2022). What criteria are young people using to select mobile mental health applications? a nominal group study. *Digital Health*, 8:20552076221102775.

[137] Kamal, A. (2013). Subjectivity classification using machine learning techniques for mining feature-opinion pairs from web opinion sources. *arXiv preprint arXiv:1312.6962*.

[138] Kanthawala, S. and Peng, W. (2021). Credibility in online health communities: Effects of moderator credentials and endorsement cues. *Journalism and Media*, 2(3):379–396.

[139] Kaplan, B. (2020). Seeing through health information technology: the need for transparency in software, algorithms, data privacy, and regulation. *Journal of Law and the Biosciences*, 7(1):lsaa062.

[140] Kaptein, M. and Van Halteren, A. (2013). Adaptive persuasive messaging to increase service retention: using persuasion profiles to increase the effectiveness of email reminders. *Personal and Ubiquitous Computing*, 17:1173–1185.

[141] Karusala, N., Seeh, D. O., Mugo, C., Guthrie, B., Moreno, M. A., John-Stewart, G., Inwani, I., Anderson, R., and Ronen, K. (2021). "that courage to encourage": Participation and aspirations in chat-based peer support for youth living with hiv. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17.

[142] Kelders, S. M., Kok, R. N., Ossebaard, H. C., and Van Gemert-Pijnen, J. E. (2012). Persuasive system design does matter: a systematic review of adherence to web-based interventions. *Journal of medical Internet research*, 14(6):e2104.

[143] Kelly, M. and Nozzi, J. (2013). *Mastering Xcode: Develop and Design*. Peachpit Press.

[144] Khan, S. and Peña, J. (2020). Using persuasive messages to increase engagement with mental health video game apps. In *Technology and Health*, pages 353–376. Elsevier.

[145] Khan, S. and Rabbani, M. R. (2021). Artificial intelligence and nlp-based chatbot for islamic banking and finance. *International Journal of Information Retrieval Research (IJIRR)*, 11(3):65–77.

[146] Khan, W., Jebanesan, B., Ahmed, S., Trimmer, C., Agic, B., Safa, F., Ashraf, A., Tuck, A., Kavic, K., Wadhawan, S., et al. (2023). Stakeholders' views and opinions on existing guidelines on "how to choose mental health apps". *Frontiers in Public Health*, 11.

[147] Kiger, M. E. and Varpio, L. (2020). Thematic analysis of qualitative data: Amee guide no. 131. *Medical teacher*, 42(8):846–854.

[148] Kim, J., Kim, Y., Kim, B., Yun, S., Kim, M., and Lee, J. (2018). Can a machine tend to teenagers' emotional needs? a study with conversational agents. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–6.

[149] King, D. R., Emerson, M. R., Tartaglia, J., Nanda, G., and Tatro, N. A. (2023). Methods for navigating the mobile mental health app landscape for clinical use. *Current Treatment Options in Psychiatry*, pages 1–15.

[150] Kirch, D. G. (2020). Physician mental health: My personal journey and professional plea. *Academic Medicine*, 96(5):618–620.

[151] Kirch, D. G. (2021). Physician mental health: my personal journey and professional plea. *Academic Medicine*, 96(5):618–620.

[152] Koh, J., Tng, G. Y., and Hartanto, A. (2022). Potential and pitfalls of mobile mental health apps in traditional treatment: An umbrella review. *Journal of Personalized Medicine*, 12(9):1376.

[153] Kopka, M., Camacho, E., Kwon, S., and Torous, J. (2023). Exploring how informed mental health app selection may impact user engagement and satisfaction. *PLOS Digital Health*, 2(3):e0000219.

[154] Kostkova, P., Brewer, H., de Lusignan, S., Fottrell, E., Goldacre, B., Hart, G., Koczan, P., Knight, P., Marsolier, C., McKendry, R. A., et al. (2016). Who owns the data? open data for healthcare. *Frontiers in public health*, 4:7.

[155] Kowatsch, T., Nißen, M., Shih, C.-H. I., Rüegger, D., Volland, D., Filler, A., Künzler, F., Barata, F., Büchter, D., Brogle, B., et al. (2017). Text-based healthcare chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity.

[156] Koyani, S. et al. (2004). Research-based web design & usability guidelines (p. 232). *National Cancer Institute*.

[157] Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I., and Group, N. Y. P. A. (2019). Can your phone be your therapist? young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical informatics insights*, 11:1178222619829083.

[158] Kumar, N. and Benbasat, I. (2006). Research note: the influence of recommendations and consumer reviews on evaluations of websites. *Information Systems Research*, 17(4):425–439.

[159] Kwan, Y., Cheng, T., Yoon, S., Ho, L., Huang, C., Chew, E., Thumboo, J., Østbye, T., and Low, L. (2020). A systematic review of nudge theories and strategies used to influence adult health behaviour and outcome in diabetes management. *Diabetes & metabolism*, 46(6):450–460.

[160] Kwon, W., Lee, M., and Back, K.-J. (2020). Exploring the underlying factors of customer value in restaurants: A machine learning approach. *International Journal of Hospitality Management*, 91:102643.

[161] Lagan, S., Aquino, P., Emerson, M. R., Fortuna, K., Walker, R., and Torous, J. (2020). Actionable health app evaluation: translating expert frameworks into objective metrics. *NPJ digital medicine*, 3(1):100.

[162] Lagan, S., Sandler, L., and Torous, J. (2021). Evaluating evaluation frameworks: a scoping review of frameworks for assessing health apps. *BMJ Open*, 11(3).

[163] Lambton-Howard, D., Simpson, E., Quimby, K., Kharrufa, A., Hoi Ming Ng, H., Foster, E., and Olivier, P. (2021). Blending into everyday life: designing a social media-based peer support system. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

[164] Langrial, S. and Oinas-Kukkonen, H. (2012). Less fizzy drinks: a multi-method study of persuasive reminders. In *Persuasive Technology. Design for Health and Safety: 7th International Conference, PERSUASIVE 2012, Linköping, Sweden, June 6-8, 2012. Proceedings 7*, pages 256–261. Springer.

[165] Leal, D. d. C., Krüger, M., Reynolds-Cuéllar, P., Caicedo, A., Gómez, C., Randall, D., and Wulf, V. (2021). Growing together, remaining apart: The role of digital technology in former guerrilla fighters' social capital. *Proc. ACM Hum.-Comput. Interact.*, 5(CSCW2).

[166] Lee, J. and Kim, J. (2018). Method of app selection for healthcare providers based on consumer needs. *CIN: Computers, Informatics, Nursing*, 36(1):45–54.

[167] Lee, M., Ackermans, S., Van As, N., Chang, H., Lucas, E., and IJsselsteijn, W. (2019). Caring for vincent: a chatbot for self-compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

[168] Lee, Y.-C., Yamashita, N., and Huang, Y. (2020a). Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW1).

[169] Lee, Y.-C., Yamashita, N., and Huang, Y. (2020b). Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):1–27.

[170] Leigh, S., Ouyang, J., and Mimnagh, C. (2017). Effective? engaging? secure? applying the orcha-24 framework to evaluate apps for chronic insomnia disorder. *BMJ Ment Health*, 20(4):e20–e20.

[171] Luxford, K., Safran, D. G., and Delbanco, T. (2011). Promoting patient-centered care: a qualitative study of facilitators and barriers in healthcare organizations with a reputation for improving the patient experience. *International Journal for Quality in Health Care*, 23(5):510–515.

[172] MacLeod, H., Oakes, K., Geisler, D., Connelly, K., and Siek, K. (2015). Rare world: Towards technology for rare diseases. In *Proceedings of the 33rd Annual ACM Conference on human factors in computing systems*, pages 1145–1154.

[173] Mahatody, T., Sagar, M., and Kolski, C. (2010). State of the art on the cognitive walkthrough method, its variants and evolutions. *Intl. Journal of Human–Computer Interaction*, 26(8):741–785.

[174] Majchrzak, A., Jarvenpaa, S. L., and Hollingshead, A. B. (2007). Coordinating expertise among emergent groups responding to disasters. *Organization science*, 18(1):147–161.

[175] Manzini, E. and Rizzo, F. (2011). Small projects/large changes: Participatory design as an open participated process. *CoDesign*, 7(3-4):199–215.

[176] Martens, D. and Maalej, W. (2019). Towards understanding and detecting fake reviews in app stores. *Empirical Software Engineering*, 24(6):3316–3355.

[177] Matthews, J., Win, K. T., Oinas-Kukkonen, H., and Freeman, M. (2016). Persuasive technology in mobile applications promoting physical activity: a systematic review. *Journal of medical systems*, 40:1–13.

[178] Matzat, U. and Rooks, G. (2014). Styles of moderation in online health and support communities: An experimental comparison of their acceptance and effectiveness. *Computers in Human Behavior*, 36:65–75.

[179] McLean, A. (2020). Mhealth apps as effective persuasive health technology: Contextualizing the "necessary" functionalities. *JMIR nursing*, 3(1):e19302.

[180] Mehta, A., Niles, A. N., Vargas, J. H., Marafon, T., Couto, D. D., Gross, J. J., et al. (2021). Acceptability and effectiveness of artificial intelligence therapy for anxiety and depression (youper): Longitudinal observational study. *Journal of Medical Internet Research*, 23(6):e26771.

[181] Mercier, J.-M., Hosseiny, F., Rodrigues, S., Friio, A., Brémault-Phillips, S., Shields, D. M., and Dupuis, G. (2023). Peer support activities for veterans, serving members, and their families: Results of a scoping review. *International journal of environmental research and public health*, 20(4):3628.

[182] Meyerowitz-Katz, G., Ravi, S., Arnolda, L., Feng, X., Maberly, G., and Astell-Burt, T. (2020). Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. *Journal of Medical Internet Research*, 22(9):e20283.

[Mindtools.io] Mindtools.io, year = 2022, t. W. u. h. l. S.

[184] Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., and Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA internal medicine*, 176(5):619–625.

[185] Montena, A. L., Possemato, K., Kuhn, E., Carlson, E. B., McGovern, M., Smith, J., and Blonigen, D. (2021). Barriers and facilitators to peer-supported implementation of mental health mobile applications with veterans in primary care. *Journal of Technology in Behavioral Science*, pages 1–12.

[186] Mozgai, S. A., Femminella, B., Hartholt, A., and Rizzo, A. (2021a). User-centered design model for mobile health (mhealth) applications: A military case study in rapid assessment process (rap). In *Extended abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–8.

[187] Mozgai, S. A., Femminella, B., Hartholt, A., and Rizzo, A. (2021b). User-centered design model for mobile health (mhealth) applications: A military case study in rapid assessment process (rap). In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21, New York, NY, USA. Association for Computing Machinery.

[188] Mudambi, S. M. and Schuff, D. (2010). Research note: What makes a helpful online review? a study of customer reviews on amazon. com. *MIS quarterly*, pages 185–200.

[189] Munson, S. A. and Consolvo, S. (2012). Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In *2012 6th international conference on pervasive computing technologies for healthcare (pervasivehealth) and workshops*, pages 25–32. IEEE.

[190] Nadarzynski, T., Miles, O., Cowie, A., and Ridge, D. (2019). Acceptability of artificial

intelligence (ai)-led chatbot services in healthcare: A mixed-methods study. *Digital health*, 5:2055207619871808.

[191] Nair, G., Johnson, S., and Sathya, V. (2018). Chatbot as a personal assistant. *International Journal of Applied Engineering Research*, 13(20):14644–14649.

[192] Narain, J., Quach, T., Davey, M., Park, H. W., Breazeal, C., and Picard, R. (2020). Promoting wellbeing with sunny, a chatbot that facilitates positive messages within social groups. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8.

[193] National Academies of Sciences, E., Medicine, et al. (2018). Evaluation of the department of veterans affairs mental health services.

[194] Neary, M., Bunyi, J., Palomares, K., Mohr, D. C., Powell, A., Ruzek, J., Williams, L. M., Wykes, T., and Schueller, S. M. (2021). A process for reviewing mental health apps: using the one mind psyberguide credibility rating system. *Digital Health*, 7:20552076211053690.

[195] Neary, M. and Schueller, S. M. (2018). State of the field of mental health apps. *Cognitive and Behavioral Practice*, 25(4):531–537.

[196] Ng, M. M., Firth, J., Minen, M., and Torous, J. (2019). User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatric Services*, 70(7):538–544.

[197] Nicholas, J., Boydell, K., and Christensen, H. (2017a). Beyond symptom monitoring: Consumer needs for bipolar disorder self-management using smartphones. *European Psychiatry*, 44:210–216.

[198] Nicholas, J., Fogarty, A. S., Boydell, K., and Christensen, H. (2017b). The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. *Journal of medical Internet research*, 19(4):e105.

[199] Noei, E., Zhang, F., and Zou, Y. (2019). Too many user-reviews! what should app developers look at first? *IEEE Transactions on Software Engineering*, 47(2):367–378.

[200] Nowell, L. S., Norris, J. M., White, D. E., and Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International journal of qualitative methods*, 16(1):1609406917733847.

[201] of America, D. (2023a). Battlepeer.

[202] of America, D. (2023b). Dryhootch of america: Peer support by veterans for veterans.

[203] Oh, K.-J., Lee, D., Ko, B., and Choi, H.-J. (2017). A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In *2017 18th IEEE International Conference on Mobile Data Management (MDM)*, pages 371–375. IEEE.

[204] Oh, Y. J., Zhang, J., Fang, M.-L., and Fukuoka, Y. (2021). A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. *International Journal of Behavioral Nutrition and Physical Activity*, 18(1):1–25.

[205] O'Leary, K., Bhattacharya, A., Munson, S. A., Wobbrock, J. O., and Pratt, W. (2017). Design opportunities for mental health peer support technologies. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1470–1484.

[206] O'Leary, K., Schueller, S. M., Wobbrock, J. O., and Pratt, W. (2018). "suddenly, we got to become therapists for each other" designing peer support chats for mental health. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

[One Mind PsyberGuide] One Mind PsyberGuide, year = 2021, t. T. u. h. l. A.

[208] Organization, W. H. (2001). The world health report 2001: Mental health: new understanding, new hope.

[209] Orji, R. and Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health informatics journal*, 24(1):66–91.

[210] O'Leary, K., Liu, L., McClure, J. B., Ralston, J., and Pratt, W. (2016). Persuasive reminders for health self-management. In *AMIA Annual Symposium Proceedings*, volume 2016, page 994. American Medical Informatics Association.

[211] O'Leary, K., Tanghe, D., Pratt, W., and Ralston, J. (2018). Collaborative health reminders and notifications: Insights from prototypes. In *AMIA Annual Symposium Proceedings*, volume 2018, page 837. American Medical Informatics Association.

[212] Pagano, D. and Maalej, W. (2013). User feedback in the appstore: An empirical study. In *2013 21st IEEE international requirements engineering conference (RE)*, pages 125–134. IEEE.

[213] Paknejad, S. (2018). Sentiment classification on amazon reviews using machine learning approaches.

[214] Parker, L., Karliychuk, T., Gillies, D., Mintzes, B., Raven, M., and Grundy, Q. (2017). A

health app developer's guide to law and policy: a multi-sector policy analysis. *BMC medical informatics and decision making*, 17(1):1–13.

[215] Patoz, M.-C., Hidalgo-Mazzei, D., Pereira, B., Blanc, O., de Chazeron, I., Murru, A., Verdolini, N., Pacchiarotti, I., Vieta, E., Llorca, P.-M., et al. (2021). Patients' adherence to smartphone apps in the management of bipolar disorder: a systematic review. *International journal of bipolar disorders*, 9(1):1–15.

[216] Petty, R. E., Cacioppo, J. T., Petty, R. E., and Cacioppo, J. T. (1986). *The elaboration likelihood model of persuasion*. Springer.

[217] Platzer, E. (2011). Opportunities of automated motive-based user review analysis in the context of mobile app acceptance. In *Central European Conference on Information and Intelligent Systems*, page 309. Faculty of Organization and Informatics Varazdin.

[218] Poole, R., Smith, D., and Simpson, S. (2015). How patients contribute to an online psychoeducation forum for bipolar disorder: A virtual participant observation study. *JMIR mental health*, 2(3):e21.

[219] Possemato, K., Wu, J., Greene, C., MacQueen, R., Blonigen, D., Wade, M., Owen, J., Keane, T., Brief, D., Lindley, S., et al. (2022). Web-based problem-solving training with and without peer support in veterans with unmet mental health needs: pilot study of feasibility, user acceptability, and participant engagement. *Journal of Medical Internet Research*, 24(1):e29559.

[220] Powell, J., Clarke, A., et al. (2007). Investigating internet use by mental health service users: interview study. In *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems*, page 1112. IOS Press.

[221] Priester, M. A., Browne, T., Iachini, A., Clone, S., DeHart, D., and Seay, K. D. (2016). Treatment access barriers and disparities among individuals with co-occurring mental health and substance use disorders: an integrative literature review. *Journal of substance abuse treatment*, 61:47–59.

[222] Rahman, R., Rahman, M. R., Tripto, N. I., Ali, M. E., Apon, S. H., and Shahriyar, R. (2021). Adolescentbot: Understanding opportunities for chatbots in combating adolescent sexual and reproductive health problems in bangladesh. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15.

[223] Ralston, A. L., Andrews III, A. R., and Hope, D. A. (2019). Fulfilling the promise of mental

health technology to reduce public health disparities: Review and research agenda. *Clinical Psychology: Science and Practice*, 26(1):e12277.

[224] Ramos, G., Ponting, C., Labao, J. P., and Sobowale, K. (2021). Considerations of diversity, equity, and inclusion in mental health apps: a scoping review of evaluation frameworks. *Behaviour research and therapy*, 147:103990.

[225] Ramsberger, P. F., Legree, P., and Mills, L. (2002). Evaluation of the buddy team assignment program. *United States Army Research Institute for the Behavioral and Social Sciences*.

[226] Redmond, S., Wilcox, S. L., Campbell, S., Kim, A., Finney, K., Barr, K., and Hassan, A. M. (2015). A brief introduction to the military workplace culture. *Work*, 50(1):9–20.

[227] Rizia, R., Franco, Z., Hooyer, K., Johnson, N., Patwary, A. K., Ahsan, G. M. T., Curry, B., Flower, M., and Ahamed, S. I. (2015a). ipeer: A sociotechnical systems approach for helping veterans with civilian reintegration. In *Proceedings of the 2015 Annual Symposium on Computing for Development*, pages 85–93.

[228] Rizia, R., Franco, Z., Johnson, N., Hooyer, K., Patwary, A. K., Ahsan, G. T., Flower, M., Curry, B., and Ahamed, S. I. (2015b). Collaborative design with veterans: Identifying challenges of designing mhealth solution for veterans. In *2015 17th International Conference on E-health Networking, Application & Services (HealthCom)*, pages 358–362. IEEE.

[229] Rizia, R., Johnson, N., Patwary, A. K., Ahsan, G. M. T., Ahamed, S. I., Franco, Z., Hooyer, K., Curry, B., and Flower, M. (2014). Mobile peer-mentoring: An approach to make veterans seek mental health-care support a normality. In *10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pages 437–441. IEEE.

[230] Rohani, D. A., Tuxen, N., Lopategui, A. Q., Faurholt-Jepsen, M., Kessing, L. V., and Bardram, J. E. (2019). Personalizing mental health: a feasibility study of a mobile behavioral activation tool for depressed patients. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 282–291.

[231] Rubya, S., Wang, X., and Yarosh, S. (2019). Hair: Towards developing a global self-updating peer support group meeting list using human-aided information retrieval. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, CHIIR '19, page 83–92, New York, NY, USA. Association for Computing Machinery.

[232] Rubya, S. and Yarosh, S. (2017). Video-mediated peer support in an online community for

recovery from substance use disorders. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 1454–1469.

[233] Rueger, J., Dolfsma, W., and Aalbers, R. (2021). Perception of peer advice in online health communities: Access to lay expertise. *Social Science & Medicine*, 277:113117.

[234] Saha, K., Ernala, S. K., Dutta, S., Sharma, E., and De Choudhury, M. (2020). Understanding moderation in online mental health communities. In *International Conference on Human-Computer Interaction*, pages 87–107. Springer.

[235] Samarajiva, R. (2005). Policy commentary: mobilizing information and communications technologies for effective disaster warning: lessons from the 2004 tsunami. *New Media & Society*, 7(6):731–747.

[236] Sannon, S., Stoll, B., DiFranzo, D., Jung, M., and Bazarova, N. N. (2018). How personification and interactivity influence stress-related disclosures to conversational agents. In *companion of the 2018 ACM conference on computer supported cooperative work and social computing*, pages 285–288.

[237] Schlosser, D., Campellone, T., Kim, D., Truong, B., Vergani, S., Ward, C., and Vinogradov, S. (2016). Feasibility of prime: a cognitive neuroscience-informed mobile app intervention to enhance motivated behavior and improve quality of life in recent onset schizophrenia. *JMIR research protocols*, 5(2):e5450.

[238] Seal, K. H., Metzler, T. J., Gima, K. S., Bertenthal, D., Maguen, S., and Marmar, C. R. (2009). Trends and risk factors for mental health diagnoses among iraq and afghanistan veterans using department of veterans affairs health care, 2002–2008. *American journal of public health*, 99(9):1651–1658.

[239] Sells, D., Davidson, L., Jewell, C., Falzer, P., and Rowe, M. (2006). The treatment relationship in peer-based and regular case management for clients with severe mental illness. *Psychiatric services*, 57(8):1179–1184.

[240] Semaan, B., Britton, L. M., and Dosono, B. (2017). Military masculinity and the travails of transitioning: Disclosure in social media. In *Proceedings of the 2017 ACM Conference on computer supported cooperative work and social computing*, pages 387–403.

[241] Semaan, B. C., Britton, L. M., and Dosono, B. (2016). Transition resilience with icts: 'identity awareness' in veteran re-integration. In *Proceedings of the 2016 CHI Conference on Human Factors*

in Computing Systems, CHI '16, page 2882–2894, New York, NY, USA. Association for Computing Machinery.

[242] Shen, N., Levitan, M.-J., Johnson, A., Bender, J. L., Hamilton-Page, M., Jadad, A. A. R., and Wiljer, D. (2015). Finding a depression app: a review and content analysis of the depression app marketplace. *JMIR mHealth and uHealth*, 3(1):e3713.

[243] Shepardson, R. L., Johnson, E. M., Possemato, K., Arigo, D., and Funderburk, J. S. (2019). Perceived barriers and facilitators to implementation of peer support in veterans health administration primary care-mental health integration settings. *Psychological Services*, 16(3):433.

[244] Silverstone, P. H. and Salsali, M. (2003). Low self-esteem and psychiatric patients: Part i–the relationship between low self-esteem and psychiatric diagnosis. *Annals of general hospital psychiatry*, 2(1):1–9.

[245] Singh, V., Kumar, A., and Gupta, S. (2022). Mental health prevention and promotion—a narrative review. *Frontiers in Psychiatry*, 13:898009.

[246] Skousen, T., Safadi, H., Young, C., Karahanna, E., Safadi, S., and Chebib, F. (2020). Successful moderation in online patient communities: inductive case study. *Journal of medical Internet research*, 22(3):e15983.

[247] Smith, D., Menon, S., and Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of interactive marketing*, 19(3):15–37.

[248] Spagnuelo, D., Ferreira, A., and Lenzini, G. (2019). Accomplishing transparency within the general data protection regulation. In *ICISSP*, pages 114–125.

[249] Srisopha, K., Link, D., Swami, D., and Boehm, B. (2020). Learning features that predict developer responses for ios app store reviews. In *Proceedings of the 14th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pages 1–11.

[250] Stolz, T., Schulz, A., Krieger, T., Vincent, A., Urech, A., Moser, C., Westermann, S., and Berger, T. (2018). A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and pc-based guided self-help interventions. *Journal of Consulting and Clinical Psychology*, 86(6):493.

[251] Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., and Mani, M. (2015). Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth*, 3(1):e3422.

[252] Sunyaev, A., Dehling, T., Taylor, P. L., and Mandl, K. D. (2015). Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*, 22(e1):e28–e33.

[253] Suzuki, T. (2016). Reminder game: Indirectness in persuasion. *Games and Economic Behavior*, 100:240–256.

[254] Sweeney, C., Potts, C., Ennis, E., Bond, R., Mulvenna, M. D., O'neill, S., Malcolm, M., Kuosmanen, L., Kostenius, C., Vakaloudis, A., et al. (2021). Can chatbots help support a person's mental health? perceptions and views from mental healthcare professionals and experts. *ACM Transactions on Computing for Healthcare*, 2(3):1–15.

[255] Ta, V., Griffith, C., Boatfield, C., Wang, X., Civitello, M., Bader, H., DeCero, E., Loggarakis, A., et al. (2020). User experiences of social support from companion chatbots in everyday contexts: thematic analysis. *Journal of medical Internet research*, 22(3):e16235.

[256] Tan, F. T. C. and Vasa, R. (2011). Toward a social media usage policy. In *Proc. ACIS*, pages 84–89.

[257] Teo, A. R., Metcalf, E. E., Strange, W., Call, A. A., Tuepker, A., Dobscha, S. K., and Kaboli, P. J. (2021). Enhancing usability of appointment reminders: qualitative interviews of patients receiving care in the veterans health administration. *Journal of general internal medicine*, 36:121–128.

[258] Thach, K. S. (2019a). A qualitative analysis of user reviews on mental health apps: Who used it? for what? and why? In *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, pages 1–4. IEEE.

[259] Thach, K. S. (2019b). A qualitative analysis of user reviews on mental health apps: Who used it? for what? and why? In *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, pages 1–4. IEEE.

[260] Torous, J., Luo, J., and Chan, S. R. (2018a). Mental health apps: what to tell patients. *Curr Psychiatry*, 17(3):21–25.

[261] Torous, J., Nicholas, J., Larsen, M. E., Firth, J., and Christensen, H. (2018b). Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evidence-based mental health*, 21(3):116–119.

[262] Torous, J. and Roberts, L. W. (2017). Needed innovation in digital health and smartphone applications for mental health: transparency and trust. *JAMA psychiatry*, 74(5):437–438.

[263] Torous, J. B., Chan, S. R., Gipson, S. Y.-M. T., Kim, J. W., Nguyen, T.-Q., Luo, J., and Wang, P. (2018c). A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatric Services*, 69(5):498–500.

[264] Toyama, K. (2011). Technology as amplifier in international development. In *Proceedings of the 2011 iConference*, pages 75–82.

[265] Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., and Torous, J. B. (2019). Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, 64(7):456–464.

[266] Valtolina, S. and Hu, L. (2021). Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness. In *CHItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter*, pages 1–5.

[267] Vande Velde, F., Overgaard, H. J., and Bastien, S. (2021). Nudge strategies for behavior-based prevention and control of neglected tropical diseases: A scoping review and ethical assessment. *PLOS Neglected Tropical Diseases*, 15(11):e0009239.

[268] Vasa, R., Hoon, L., Mouzakis, K., and Noguchi, A. (2012). A preliminary analysis of mobile app user reviews. In *Proceedings of the 24th Australian computer-human interaction conference*, pages 241–244.

[269] Walji, M. F. and Zhang, J. (2008). Human-centered design of persuasive appointment reminders. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, pages 236–236. IEEE.

[270] Wallerstein, N. and Duran, B. (2017). The theoretical, historical and practice roots of cbpr. *Community-based participatory research for health: Advancing social and health equity*, pages 17–29.

[271] Wang, K., Varma, D. S., and Prosperi, M. (2018). A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *Journal of psychiatric research*, 107:73–78.

[272] Wang, X., Markert, C., and Sasangohar, F. (2021). Investigating popular mental health

mobile application downloads and activity during the covid-19 pandemic. *Human Factors*, page 0018720821998110.

[273] Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., and Andrews, G. (2013). Cbt for depression: a pilot rct comparing mobile phone vs. computer. *BMC psychiatry*, 13(1):1–9.

[274] Wildeboer, G., Kelders, S. M., and van Gemert-Pijnen, J. E. (2016). The relationship between persuasive technology principles, adherence and effect of web-based interventions for mental health: a meta-analysis. *International journal of medical informatics*, 96:71–85.

[275] Windler, C., Clair, M., Long, C., Boyle, L., and Radovic, A. (2019). Role of moderators on engagement of adolescents with depression or anxiety in a social media intervention: content analysis of web-based interactions. *JMIR mental health*, 6(9):e13467.

[276] Winkler, R. and Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In *Academy of Management Annual Meeting (AOM)*.

[277] Yadav, D., Malik, P., Dabas, K., and Singh, P. (2019). Feedpal: Understanding opportunities for chatbots in breastfeeding education of women in india. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW).

[278] Ye, Q., Zhang, Z., and Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert systems with applications*, 36(3):6527–6535.

[279] You, Y. and Gui, X. (2020). Self-diagnosis through ai-enabled chatbot-based symptom checkers: user experiences and design considerations. In *AMIA Annual Symposium Proceedings*, volume 2020, page 1354. American Medical Informatics Association.

[280] Zhang, J., Oh, Y. J., Lange, P., Yu, Z., Fukuoka, Y., et al. (2020). Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet. *Journal of medical Internet research*, 22(9):e22845.

[281] Zhang, R., E. Ringland, K., Paan, M., C. Mohr, D., and Reddy, M. (2021). Designing for emotional well-being: Integrating persuasion and customization into mental health technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

[282] Zhou, J., Saha, K., Lopez Carron, I. M., Yoo, D. W., Deeter, C. R., De Choudhury, M., and

Arriaga, R. I. (2022). Veteran critical theory as a lens to understand veterans' needs and support on social media. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW1).

[283] Zimmer, M. (2010). "but the data is already public": on the ethics of research in facebook. *Ethics and information technology*, 12(4):313–325.

APPENDICES

APPENDIX A: List of 164 mobile mental health applications considered in chapter 2

A.1 List of mobile mental health applications available in Google Play Store

- Self-manage Depression: Daily exercise (GGDE)
- Stress Control Norbu: game, breathing, meditation
- uMore — Stress Measuring and Mood Tracker Monitor
- PsyTests
- MoodSpace - Stress, anxiety, and low mood self-help
- Pocketcoach - Reduce Anxiety
- CBT Companion: (Cognitive Behavioral Therapy app)
- Anxiety Tracker - Stress and Anxiety Log
- Panic Diary: A anxiety tracker app
- Aurum - best therapy app, stress, anxiety.
- Bearable - Symptoms and Mood tracker
- Zulu Logical Puzzle and Anxiety Relief Game
- My Oasis: Calming, Relaxing and Anxiety Relief Game
- Calmerry Counseling and Therapy
- Managing your stress and anxiety
- Cerebral - Mental Health
- Whispers from God - Christian Meditation
- Anxiety Calculator
- WellTrack - Interactive Self-Help Therapy

- Mind journal: Diary, Mood tracker and Gratitude
- Cognitive Behavioral Therapy - CBT
- Yoga for Anxiety
- CBT Tools for Healthy Living, Self-help Mood Diary
- CBT Guide to Depression Self-help: Mood Log, Diary
- Stop Panic and Anxiety Self-Help
- Dealing with Depression
- Anxiety Test
- FeelJoy - Mental Well-Being for Anxiety
- Flow - Depression
- GG Self Care: Create a Positive Headspace
- HearMe — Empathy Not Therapy
- Online Therapy, Emotional diary, Mindfulness tools
- Relax and Sleep Well: Hypnosis and Meditation
- Lift Depression and Anxiety
- BestHelp Therapy online, mental help
- A Day in the Life - Living With Depression
- I Meet Myself
- Depression: The Game
- MDLIVE: Talk to a Doctor 24/7
- feel better - Mood and CBT therapy to manifest goals
- UP! - Depression, Bipolar and Borderline Management
- Mind health AI Friend Free Counseling Therapy Chat
- Mood Potatoes - Mood Tracker — Mental Health Diary
- MoodLinks - Anxiety and Depression

- MoodMission - Cope with Stress, Moods and Anxiety
- Moodflow: Mood and Symptom Tracker, Diary.
- Mood Time - Mood Tracker, Improves Mental Health
- Depression Test
- My Online Therapy: For Mental Health and Self Care
- MyPossibleSelf: Mental Health
- OCD Test
- Online Counseling - Feel less depression and anxiety
- Mental Health Tracker: Mental Disorders Test
- Carpe Diem - Depression and Anxiety Forum
- Depression Test
- Psychiatry Pro-Diagnosis,Info,Treatment,CBT and DBT
- WeAreMore: Find a Therapist + Free Peer Support
- Reflectly: Diary, Gratitude Journal and Mood Tracker
- Steady - Breathing Meditation For Stress and Anxiety
- Moodfit - Fitness for Your Mental Health
- Serenity: Guided Mental Health
- Shadow's Edge - Resilience, Selfcare and Anti Stress
- SilverCloud
- ezeCBT - CBT Self Help Thought Diary
- Mind Ease: Anxiety Relief
- Anxiety Relief Hypnosis - Stress, Panic Attacks
- Talkspace Counseling and Therapy
- TherapyChat - Online therapy and counselling
- InnerHour Self-Care Therapy: Anxiety and Depression

- NOCD: OCD Treatment App
- Anti Stress Games, Relaxing, Stress Anxiety Relief
- MindReset Free your mind in moments!
- Panic Attack Anxiety Relief: Breathing Exercises
- Wisdo - Empathic Emotional Support Community
- Beat Depression
- Skills Games to cope with stress
- Depression App
- Pi Journal: anxiety relief therapy and mood tracker
- Mood Tracker Journal. Mental Health, Depression
- Worrydolls
- Insight: For Stress and Anxiety Relief
- Mood Patterns - Mood Tracker and Diary with Privacy
- Mindspa: Self-Care, Mental Health and Intellect
- Pinkymind for anxiety and stress relief
- Moodily - Mood Tracker, Depression Support
- eMoods Bipolar Mood Tracker
- STOPP app
- Clear Fear

A.2 List of mobile mental health applications available in Apple App Store

- MyLife Meditation: Meditate, Relax Sleep Better
- Calm Harm - manages self harm
- MOODnotes- mood tracker
- Cove - music for mental health
- Moods - Mental Helath Tracking
- Aetheria
- Wellness: Mood Meds Panic Relief
- Joyable" An AbleTo Program
- Mend: Go within
- Meditation and Sleep Sounds+
- Stress guide: HRV Meditation
- Sibly
- Primed Mind: Mindset Hypnosys
- Mood app: Journal
- Happy Not Perfect: Mind Gym
- Aloe Bud
- Mindset health: Hypotherapy
- Jour: Daly self-care expert
- Sayana: Daily self-care guide
- Quirk CBT
- Breathe2Relax
- Moodkit
- RR: Eating disorder
- Ginger emotional support

- iMood Journal
- superbetter
- Moodtrack social diary
- smiling mind
- Nomo - Sobriety clocks
- Rootd - Panic Attack Anxiety Relief
- 12 Steps - AAA Alcoholic
- CBT i-coach
- mindful - motivational widget
- Guided mind
- psychology - magic brain psycho
- I AM sober
- stress Anxiety comapinion
- Thinkup: positive affirmations
- quit that - Habit tracker
- breethe: Meditation and sleep
- Remente" Self help and wellbeing
- Feelmo: Mental health support
- reallifechange
- Anxiety test relief
- Whil: Mindfullness meditation
- Ten percent happier mediattion
- 7mind - meditation and sleep

A.3 List of mobile mental health applications available in both platforms

- GG OCD - Anxiety and Depression
- Wysa: Sleep Depression Support
- Youper AI Therapy for Anxiety Depression
- Sanity Self: anxiety stress relief, sleep sounds
- Pocketcoach - Reduce Anxiety
- DBT Coach: (Dialectical Behavior Therapy app)
- Bloom : Better You
- TalkLife for Anxiety, Depression Stress
- BetterHelp: Online Counseling Therapy
- Mindshift
- Depression test
- Happify
- What's Up? - A Mental Health App
- MoodTools - Depression Aid
- Gratitude: Journal, Affirmations Vision Board
- Sanvello: Anxiety Depression
- 7 Cups: Online Therapy for Mental Health Anxiety
- Shine: Calm Anxiety Stress
- Simple Habit: Meditation, Sleep
- stoic.
- Talkspace Counseling Therapy
- Self-help Anxiety Management
- Booster Buddy
- Woebot: Your Self-Care Expert

- MindDoc: Your Mental Health Companion
- Mindfulness Coach
- Dare: Anxiety and pain relief
- Mental Health Tests
- Calm

APPENDIX B: User study materials for chapter 4

B.1 Study protocol for focus group study done in chapter 4

The study was divided into a few sessions. The first session began with a 10-minute presentation about the goals and motivations for the project. Then a discussion was led based on the following questions, but I want to stress that the goal was to be an open discussion –

- Would you please tell us about your experience of working with your mentees? Particularly, what are the scenarios or specific situations your mentees have contacted you or you contacted them? How often do they contact you? Do you think that's sufficient? I think most of the mentors meet in person with their mentees or call or text them. Which do you prefer? Does the mode of communication change base on your mentee's needs? How? Other than calling and texting, do you have any other mode of communication, like any other apps or systems? If so, why?
- How do you remind your peers in doing daily support-related activities, daily surveys, and check-ins? Do you set yourself any reminders? Do you have any preferred technology to set those reminders? Is there any time and place that motivates you to complete these tasks? Why? What motivates you to complete these tasks?
- Are there specific behaviors or signals from your mentees that make you think that they are on the verge of a crisis? In those cases, what steps do you take? How do you feel about an automated early detection of crisis? For example, from your mentees' actions and symptoms, your phone will automatically send you a text or an email alerting you that they may be at risk. What actions and behaviors of your mentees should be considered to automatically detect this type of risk? How would you like to be alerted? For example, text, email, a notification with a loud sound?
- Are there things about being a peer mentor that are difficult for you? How do you stay motivated to help veterans even if you are feeling overburdened?

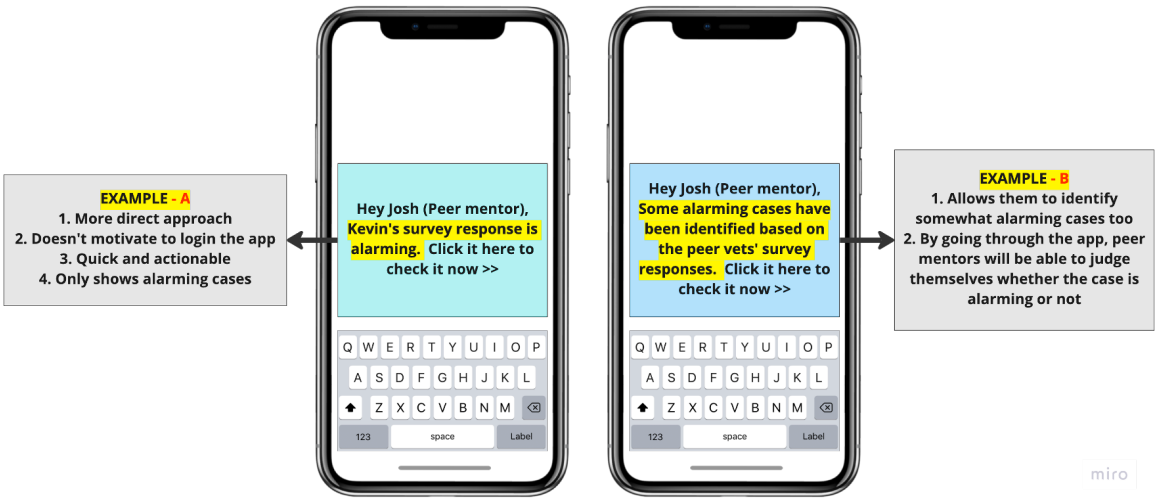


Figure A1: Scenario 1

B.2 List of scenarios presented to peer mentors for participatory design study done in chapter 4

The participatory design session began with a 20-minute presentation on a plain explanation of persuasive technology in the QRF apps. It is assumed that only a few mentors may have used the app if they participated in a prior study. Therefore, the presenter agve an overview of the current version of the apps and the team's plans on adding reminders for facilitating communication. Then the participants was being presented with some actual scenarios (based on the prior studies our team has done) where these persuasive and alert reminders can be pushed. A few examples of such scenarios are -

Figure A2: Scenario 2

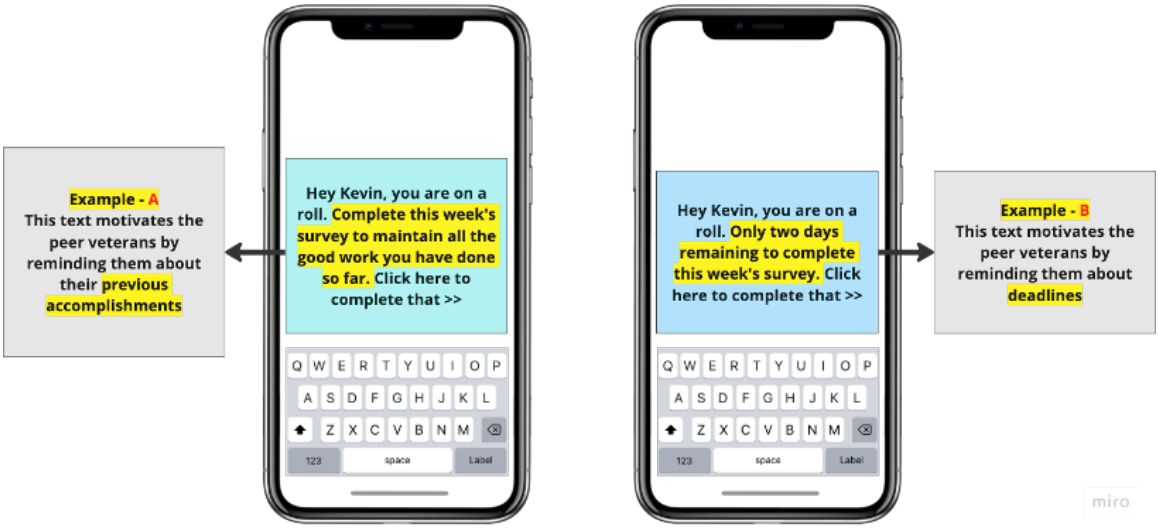


Figure A3: Scenario 3

Figure A4: Scenario 4

Figure A5: Scenario 5

APPENDIX C: User study materials for chapter 4

C.1 Study protocol for user study done in chapter 5

For this study, we recruited 10 participants. The study session was done in two phases.

First Phase:

In the first phase, at first, the participants was given a short presentation on what our objective for the study is. Then, one of the researchers had an open discussion about how they choose mobile mental health applications when needed based on the following semi-structured questions.

- How familiar are you with mobile mental health applications? Have you used any? What features do you look at while downloading a mental health application?
- What are your top priorities while downloading a mobile application?
- Can you give me a step-by-step on how you choose a mobile mental health application?
- Do you look at user reviews while downloading an application? In the reviews, what particularly do you look at while downloading a mobile application?
- In your experience as a user, is there enough information on the app store to make a uniform decision about the app? What additional information in the app store that might help you in making a decision?
- Do you look for any external sources like blogs, or websites to make a decision about using a mental health mobile app? Are the helful? Please explain
- How do you identify the credibility of an app? e.g., how do you know that this app will work for you?
- How important is it that the app is backed by science or research? Why? What constitutes a "Backed by science or research" to you?
- Does user experience matter to you when choosing an application? How do you identify the user experience of an app?
- How is the cost of an application important to you? Do you measure the cost-to-benefit ratio while downloading an application? How do you measure it?

- How important is privacy to you using these mobile mental health applications? What privacy controls would you like to have while using these apps?
- Do you check privacy policy while downloading mobile mental health applications? If you do, then what particularly do you look there? If you don't, why?
- Do you check any requirements or features specific to your gender in the applications? If yes, what particularly do you look for before downloading any applications? Can you think of any gender-based useful features that will make your experience as a user better? Please explain.
- Same questions with the age. Do you check any requirements or features specific to your age in the applications? If yes, what particularly do you look for before downloading any applications? Can you think of any age-based useful features that will make your experience as a user better? Please explain.
- In your experience as a user, how important is it for the app to be culturally appropriate? Care to elaborate?
- Do you have any experience with applications that have artificial intelligence such as chatbots as a feature? If yes, what particularly do you find useful and what is not very useful? What is your opinion on using AI tools in mobile mental health applications?

Second Phase:

In the second phase of the session, we showed the participants some actual user reviews from the mobile app stores (e.g. Apple App Store or Google Play Store). Our goal was to gather an initial thought on whether showing them these reviews is useful. The reviews are based on a few identifiable features from prior literature that users usually look for while downloading mobile mental health applications.

Scenario 1: Suppose you are concerned about the privacy regulations of an app. Does seeing these three reviews help you in making a decision?

> *Review 1: "Ok so as a teenager its like really good to have a daily journal especially bc u can keep track of how ur doing on some days and u can keep track of how ur health is on this app too. Only complaint is that I have to have a premium subscription to get a pin number lock on the app?? It shouldn't be like that. If I don't want my cousins or nephews and nieces or anyone looking at this I would have to pay? Its not fair. PLEASE whoever works for this apps company let the password feature be unlocked for people without a*

subscription. Privacy is a very important thing and as I said, we should not have to pay for our own privacy."

Review 2: "I don't like apps that ask personal information and then have to link them to an email account. It's too stressful- no privacy- when it asked for my email I deleted it- never found out if it was good or not.

Review 3: "This app collects too much information about me that is not necessary for me to use their app. Look at their new privacy info in the App Store. I value my privacy and wish they didn't collect so much info and use it."

- What is your opinion about the app seeing these reviews now?
- How helpful was it in deciding on an app that values user's privacy?
- What information about privacy is missing in these user reviews?
- What are your next steps if you see these three user reviews?

Scenario 2: Suppose you are concerned about the cost of a mobile mental health applications. Does seeing these five reviews help you in making a decision?

Review 1: "But once it's up, almost everything is behind a paywall. If you guys want to sell a paid app, just do it up front. I know it sounds counter-intuitive to meditation, but I'd even watch ads before the sessions if that would help them offset production costs. Super expensive otherwise."

Review 2: "I like it a lot, but I wish I didn't have to pay for sessions. I have a problem with sleep, I can't get enough sleep so, that makes a big impact on school. I need and want to have a good education. I would buy it but I need money for other things. In conclusion, it's good but costs money."

Review 3: "I love this app. The daily activities and games help me stay focused and positive. I went ahead and paid for the subscription after testing out the free version. For me, it's worth the cost of a cup of coffee a month for all the fantastic benefits."

Review 4: "Watch out for the auto renew option, it's not clear and canceling subscription was not easy, had to research a little since other services like Spotify, HBO Now, etc... make it very easy to unsubscribe.DO NOT TRUST developers that want to keep you engaged at any cost."

Review 5: "I'm confused what the point of this app is? I know there is a premium version but i can't find any features that aren't locked. Even all of my recommended mediations, beginners mediations, I've scrolled through every category, all of it is locked. There is literally nothing ti do in the app if you don't pay. Why not just charge to download the app. All of your ads are very nice, but i can't even do a breathing exercise for free. By all means charge for your app but I would thing there would be at least something, anything available."

- What is your opinion about the app seeing these reviews now?
- How helpful was it in deciding on an app that values user's privacy?
- What information about privacy is missing in these user reviews?
- What are your next steps if you see these three user reviews?

Scenario 3: Suppose you are trying to see how easy this app to use. Does seeing these three reviews help you in making a decision?

Review 1: "App will log you out on its own after inactivity, but you can't log back in afterwards. Gives you an error message. Only way to log back in is to repeatedly uninstall and reinstall the app. What a waste. I was genuinely enjoying the content I'd gotten to use so far."

Review 2: "You don't know where to start and there is not a step by step guid. You just face a lot of videos/audios and will be confused. Not recommended."

Review 3: "I am really loving this app! it's easy to use and i like that there is a point system in place because I am a very rewards driven person and that helps me feel that I truly am making progress."

- What is your opinion about the app seeing these reviews now?
- How helpful was it in deciding on an app that values user's privacy?
- What information about privacy is missing in these user reviews?
- What are your next steps if you see these three user reviews?