

User Perspectives and Ethical Experiences of Apps for Depression: A Qualitative Analysis of User Reviews

Dionne Bowie-Dabreo
Department of Psychology
Lancaster University
d.bowie@lancaster.ac.uk

Corina Sas
School of Computing and Communications
Lancaster University
c.sas@lancaster.ac.uk

Heather Iles-Smith
Research and Innovation, Salford Royal NHS
Foundation Trust
H.M.Iles-Smith@salford.ac.uk

Sandra I Sünram-Lea
Department of Psychology
Lancaster University
s.sunram-lea@lancaster.ac.uk

Apps for depression have the potential to innovate mental health care and increase access to treatment. Yet, concerns abound with disparities between academic development of apps and those available direct-to-consumers through the app marketplace. Reviews have highlighted ethical shortcomings of these self-management tools, with a need for greater insight into how ethical issues may be experienced by users. We addressed these gaps by exploring user reviews of apps for depression to better understand user experiences and ethical issues. We conducted a thematic analysis of 2,217 user reviews sampled from 40 depression apps in Google Play and Apple App Store, totaling over 77,500 words. Users reported both positive and negative experiences, with ethical implications evident in areas of benefits, adverse events, access, usability and design, support, commercial models, autonomy, privacy, and transparency. We conclude by presenting an ethical framework for developing apps for depression and navigating their ethical tensions.

CCS CONCEPTS • **Social and professional topics~Professional topics~Computing profession~Codes of ethics • Human-centered computing~Human computer interaction (HCI)~HCI design and evaluation methods• Applied computing~Law, social and behavioral sciences~Psychology**

Additional Keywords and Phrases: Mobile mental health, Depression, User experiences, Ethics, User reviews

1 INTRODUCTION

Mental health is a global health concern, with one in six people estimated to experience a common mental health problem in any given week [46]. Among these common mental health problems, depression, defined by persistent negative mood and diminished positive affect [59], has been shown to be particularly disabling, standing as the second leading cause of disability worldwide and a significant contributor to the global burden of disease. More specifically, people experiencing depression may present decline in functioning associated with the presence of depressed mood or the loss of interest or pleasure, accompanied by physiological changes (weight loss or gain, fatigue), insomnia or hypersomnia, psychomotor agitation or retardation, feelings of worthlessness or guilt, or cognitive changes (reduced concentration, indecisiveness, recurrent thoughts of death, or suicide ideation) [2]. Given the level of impairment and high prevalence, it is not surprising that an increasing body of work both in academia and in commercial settings has focused on the design and development

of more accessible, cost-effective digital interventions for depression, with potential to remove situational and attitudinal barriers which often accompany traditional interventions.

Digital mental health is the use of technology for mental health care, support, and resources [77]. This commonly includes online resources for mental health information and psychoeducation; online therapy; blended care integrating technology-based self-care with in-person support; smartphone applications; and biosensors and wearables for detecting, monitoring, and tracking mental health symptoms [8, 33, 70].

While such technological approaches to the treatment of depression have the potential to revolutionize care, much remains unknown about their long-term effectiveness, risks and implications, and broader impact on individuals, healthcare systems, and society [54, 76]. Notable concerns have been raised surrounding the ethical implications of digital mental health and the potential for unmitigated risks or misuse, often due to a lack of adequate ethical guidelines, regulations, and evaluations [34, 38, 44, 61]. These concerns are particularly relevant for the subfield of mobile mental health, specifically publicly available mobile mental health apps accessed through app stores such as the Google Play store or Apple App Store. So far, these major app marketplaces remain largely ungoverned, offering direct-to-users, as consumers, mental health products for unsupported self-care [28, 38, 39].

While work in this space has started to emerge focusing on evaluating the content [13, 14, 28, 45, 53], functions [37, 51, 65], and quality [38, 67, 76, 78] of apps for depression, there is a strong need to better understand the breadth of ethical implications of the depression app marketplace and to provide clearer design, development, and deployment guidelines to increase ethical practices. Moreover, there is a poverty of research exploring user experiences of publicly available apps for depression and the potential impact of ethical issues on their use and adoption.

To address these gaps in ethical understanding and user experiences we report a study of 2,217 user reviews sampled from 40 apps for depression listed in Google Play and Apple App Store (UK) to understand how apps for depression can be better designed to account for users' feedback in their reviews. We sought to answer the following research questions:

1. What are users' experiences of publicly available apps for depression?
2. What ethical issues are evident in app store user reviews of apps for depression?
3. Based on user experiences, what are the key elements of ethically designed apps for depression?

The contribution of our work is three-fold. First, we provide diverse user perspectives of apps for depression and how aspects of their design and development impact not only users' experiences of the apps, but their wellbeing. Second, our analysis and discussion of ethical issues of apps for depression is framed within user reviews, resulting in concrete examples of ethical concepts rather than abstract and often ambiguous concepts from ethics theory that may be difficult to apply in design. Third, we generated several implications for designing better ethics-informed apps for depression. We integrate these implications within a framework for the design of ethical apps for depression offering guidance on the navigation of tensions among key ethical areas.

2 RELATED WORK

There has been an expanse of innovation and research on digital health technologies for depression, with apps being one of the most widely developed and used. Research has shown the potential benefits of mobile mental health, but there are ethical concerns and limited understanding of how these impact users in the wild. Here, we outline the literature on mobile mental health for depression and the ethical challenges in this field.

2.1 HCI research on depression

HCI work on depression has grown significantly in the past decade, ranging from exploring the impact of depression on one's use of technology [17, 26] and social media behaviors [4, 22] to detecting or predicting depressive symptomatology from social media data [42, 60, 72].

Efforts to develop more objective methods of detecting depression include multimodal systems [12, 23] integrating for instance audio with facial or body data [75]. HCI work on the detection of depression also includes the use of commonly available technologies such as smartphone sensors [6, 47, 73] for passive or active data collections. Beyond prediction and detection, a growing body of HCI work centers on the design of technology for the treatment or management of depression symptoms, leveraging memory technologies [52], game interventions [20], virtual reality [7], social robots [56], or chatbots [25]. Most commonly, technological interventions for depression are delivered online or via apps [58, 74, 80]. While much of the HCI work in this area has focused on technologies for the self-management of depression, others have sought to enhance face-to-face treatments by using technology in existing healthcare systems [11] or in blended care [64].

The design and development of technologies for depression can be a difficult space to navigate in HCI, with challenges ranging from access to and co-design with vulnerable user groups, to the potential impact of sensitive content on the wellbeing of designers and researchers [61]. To address such challenges, researchers have explored novel design methods to increase understanding of users' lived experiences of depression, real-world contexts for use and adoption, and the potential impact of proposed technologies [32, 49, 62].

HCI work in depression has also sought to improve digital intervention by exploring factors impacting engagement and adoption [21, 43] and improved clinical outcomes [18]. Factors impacting use and effectiveness of mental health technologies and self-management include (but are not limited to) appropriate client support [18, 55], managing expectations [30], provision of social support [40], and designing for flexibility in use [21] and fluctuations in symptoms [35].

2.2 Mobile mental health for depression

With the advent of smartphones and commercial apps, mobile apps are among the most used technologies for depression, often throughout the entire lifespan of depression, from screening and detection to treatment and relapse prevention. This is supported by the extensive body of research in mobile mental health, where clinical trials [5, 24] and systematic reviews [54, 68] have highlighted the efficacy of app interventions for depression. Apps for depression have been reported to significantly reduce depressive symptoms and improve overall wellbeing, with common depression app functionalities including psychoeducation [28, 45, 37, 38], screening and assessment [28, 37, 51, 63], symptom management [28, 53, 63], interactive interventions [37], and tracking of moods, thoughts, or behaviors [24, 45, 53]. Publicly available app interventions for depression may be based on existing evidence-based treatments such as cognitive behavioral therapy [24, 68], and behavioral activation and dialectical behaviour therapy [68]. However, few of these apps demonstrated high fidelity to the adopted treatment approach [14, 65], causing them to be more aptly described as evidence-informed, rather than truly evidence-based. There is also high prevalence of complementary and alternative treatments for depression, including mindfulness meditation, hypnosis, and sound or music therapy [14].

Publicly available app interventions for depression also appear to adopt innovative uses of traditional treatment approaches and strategies, with frequent use of integrative, multitheoretical, or transdiagnostic approaches [14]. These innovations reflect the eclectic delivery of clinical interventions in the wild [41] but unfortunately often lack research evidence supporting apps' specific design and use in treating depression [14]. The poverty of research on apps for depression in the wild also limits knowledge of use and adoption beyond academic research.

There are also concerns with high rates of attrition and issues with treatment adherence of apps for depression [9, 57]. Studies have suggested users may engage with such apps for short-term symptom management but discontinue their use once the symptom is no longer an immediate concern, with some users describing apps as a ‘crutch’ to help them cope until they find a more sustained means of managing their difficulties [19]. Despite the growing body of work on apps for depression, only a few studies explored user experiences through app store user reviews [1, 48, 65]. These studies show that user reviews outline both positive and negative aspects of user experiences, with users appreciating flexible access to care, variety in app functionalities, and engagement features such as customization and notifications. Negative user experiences of mental health apps were typically related to poor usability, concerns with content, issues with privacy and security, poor customer service, and issues with costs and billing. While most studies discussed user views on privacy and security, only one [48] made explicit mention of ethical issues which were limited to privacy violations and excessive in-app advertisements. There is therefore a need to amplify the user voice in discussions of the ethics of mobile mental health, and to account for their lived experiences in the provision of guidelines and recommendations.

2.3 The ethics of mobile mental health

Mental health professionals have long been guided in their work by ethical principles and codes of conduct aimed at ensuring good and fair delivery of care in the best interests of the client, the profession, and wider society [3, 16]. There has been increased discussion of the ethics and evaluation of mobile mental health, largely in the form of expert commentaries and reviews [31, 34, 51]. Authors have highlighted key ethical issues in mobile mental health, most commonly privacy and data security, risks and safety concerns, and benefits and evidence [27, 36, 44, 61]. These discussions often explored related issues of transparency, trust, and informed consent [69, 79]. More targeted ethical issues emerging from the literature included the importance of user involvement within the ethical development of mental health technologies, respect for human rights and diversity, and challenges with standards and regulation.

Few researchers however have framed these discussions within existing ethical frameworks. Sanches et al. [61] used bioethics [10] as a lens to present their review of the ethics of HCI and affective health research. This included discussion of the ethical principles of autonomy (respect for the decision-making capacity of autonomous people), beneficence (providing benefits and balancing risks), nonmaleficence (avoiding harm), and justice (fairness in distribution of benefits and risks for all people), with a focus on how these principles were reflected in HCI research and design for affective health. Comparatively, Jones and Moffitt [31] and Karcher and Presser [34] referenced the professional ethical principles of the American Psychological Association [3] to provide guidance for app development and the use of mobile health in clinical practice, respectively. Bowie-DaBreo et al. [14] adapted these frameworks in their review of ethical issues within app store descriptions of apps for depression [13], with their findings highlighting the relevance of principles of beneficence, nonmaleficence, responsibility, integrity, autonomy, and justice. They advocated for the application of these ideals using a responsible innovation approach [50], which encourages a process of anticipation, reflection, inclusive deliberation, and responsiveness in the design and development of new technologies [71].

Principlism (i.e. the principles of biomedical ethics) [10] is typically used as the foundation for professional ethical codes of conduct [3, 16] as the theory provides a structured approach for ethical guidance and practice, particularly when compared to more abstract theories such as consequentialism [the greatest good (outcomes) for the greater number], deontology (focus on actions, duty, and responsibility, not outcomes), and virtue ethics (how one’s character or values should be) [66]. While principlism is widely used in the practical application of ethics, some consider it to be too prescriptive and encourage integration with other ethical theories and ideals [29]. This is relevant for HCI work in depression and mobile mental health, as researchers and designers can often feel limited when discussions of ethics and

associated guidelines arise. The present study therefore aims to amplify the users' voice and experiences of ethical issues of apps for depression, and to use this perspective to shape accessible and applicable guidance for the design and development of ethical mobile mental health. We approached ethics in the broadest sense as relating to individual and social good and universal standards of right and wrong [66]. This often relates to but is not limited to issues of harm, fairness, and rights.

3 APP REVIEW STUDY

3.1 Sampling method

We now describe the method for sampling the apps and for sampling the user reviews. The search for apps for depression was conducted on the two main app stores (UK version): Google Play Store and Apple App Store, during October–November 2018, guided by methods used by Shen et al. [63] and Stawarz et al. [65]. Separate searches were performed using the terms “depression” and “mental health”, as well as a hand-search for apps for depression which were reported in previous research but not returned in the searches. For this research, apps for depression were defined as apps with app store listings mentioning depression or depressive symptoms. Apps were included in the review if they met the following criteria: (1) app description included terms “depression”, “low mood/mood disorder”, “mood management”, “negative thoughts”, or “distress”; and (2) app store listing was in English. Apps were excluded from review if they: (1) did not mention depression or depressive symptoms, (2) were for professional training, (3) only provided depression quotes or wallpapers, or (4) were duplicates, i.e., copies of an app listed within the same app store. Apps were not excluded from review if they targeted another mental health problem (e.g., anxiety) once they mentioned depression or depressive symptoms, as outlined in inclusion criteria (1). This returned a total of 353 unique apps for depression for which we captured the number of users rating them, number of downloads, and users' ratings (from 1 to 5).

From this large set of apps, we decided to focus on a subset of them, to allow for the in-depth analysis of a rather large user reviews data. To include the breadth of user reviews, we aimed for both positive and negative reviews, that were written by large numbers of users. For this, we ranked all 353 apps according first to the numbers of users rating them, and second according to users' ratings. Thus, we identified the 20 most rated apps for depression, or those with the highest number of user ratings across the app store, ranging from 160,019 to 4,082 raters, and whose Google Play ratings range from 3.8 to 4.8. These apps are those that users frequently downloaded with number of downloads ranging from over 100,000 to more than 10,000,000. We also identified the 20 lowest rated apps for depression whose Google Play ratings ranged from 2.5 to 3.7, and which were downloaded by at least 1,000 users (range 1,000 to more than 100,000). So, while the former 20 apps were among the most popular and positive rated ones, the latter 20 apps were least popular and more negatively rated but still widely downloaded.

Apps were removed from selection and replaced by the next app in the category if they were no longer listed in the app stores or if the app had no user reviews. This resulted in six apps being removed from the lowest rated apps. The final 40 sampled apps are listed in Appendices A.1 and A.2.

Then, user reviews were purposively sampled in December 2018. For each app the 50 ‘*most helpful reviews*’ (determined by the app store filter for sorting reviews) were extracted from each platform, with a total of 100 reviews being sampled for apps listed in both stores. Additionally, the 50 ‘*most critical reviews*’ (also an app store filter) were extracted from apps listed in Apple App Store; this categorization was not available in Google Play. To ensure richness of data, user reviews were excluded from selection if they lacked content, for example reviews consisting of only ‘emoji’ symbols or single word expressions such as “Bad”. This led to a final set of 2,217 valid user reviews with an average of 35 words per

review, totaling over 77,500 words. These reviews were extracted verbatim and exported to ATLAS.ti for analysis (see Figure 1 for sampling flowchart).

3.2 Thematic analysis

Thematic analysis was used to explore user experiences and ethical issues of apps for depression, using the methods and guidelines for thematic analysis outlined in Braun and Clarke [15]. The user reviews were first coded as ‘positive’, ‘negative’, and ‘ambivalent’. ‘Positive’ and ‘negative’ were defined as reviews which only discussed the positive or negative parts of the app. Reviews were coded as ‘ambivalent’ when users discussed both positive and negative aspects of the app. The user’s numerical rating of the app was also recorded. Ratings are scored on a scale of 1 to 5, with 1 being the lowest rating. Inductive codes were generated from the user reviews to capture the content, context, and ideas expressed. This included the use of in vivo codes to reflect important concepts and the user’s voice. This iterative process involved ongoing review of quotes and consolidation of codes. The final code list was then categorized into themes guided by the inductive codes and the idea of ethics as standards of right and wrong that apps for depression should encompass. Thematic development and mapping were led by the first author with ongoing discussion with all authors over six months, until consensus was reached. Our findings present a sample of quotes from user reviews, with minor edits to improve readability and to protect user anonymity. This research received institutional ethics approval.

4 FINDINGS

4.1 Summary of user reviews

Over half of all user reviews (53%, 1178/2217) were positive, with 27% (592/2217) classed as negative, and 20% (447/2217) as ambivalent. Just over half of all reviews had a user rating of 5 (51%, 1129/2217). The next most common rating was 1 (22%, 482/2217), then 4 (14%, 312/2217), 3 (7%, 164/2217), and 2 (6%, 130/2217).

4.2 Benefits and harms of apps for depression

4.2.1 Benefits of apps for depression

Almost half of all user reviews (42%, 921/2217) across 36 of the 40 apps sampled described a benefit of using apps for depression (Table 1). Most commonly, users described benefits to their mental health and wellbeing, while some benefited from using apps during difficult situations:

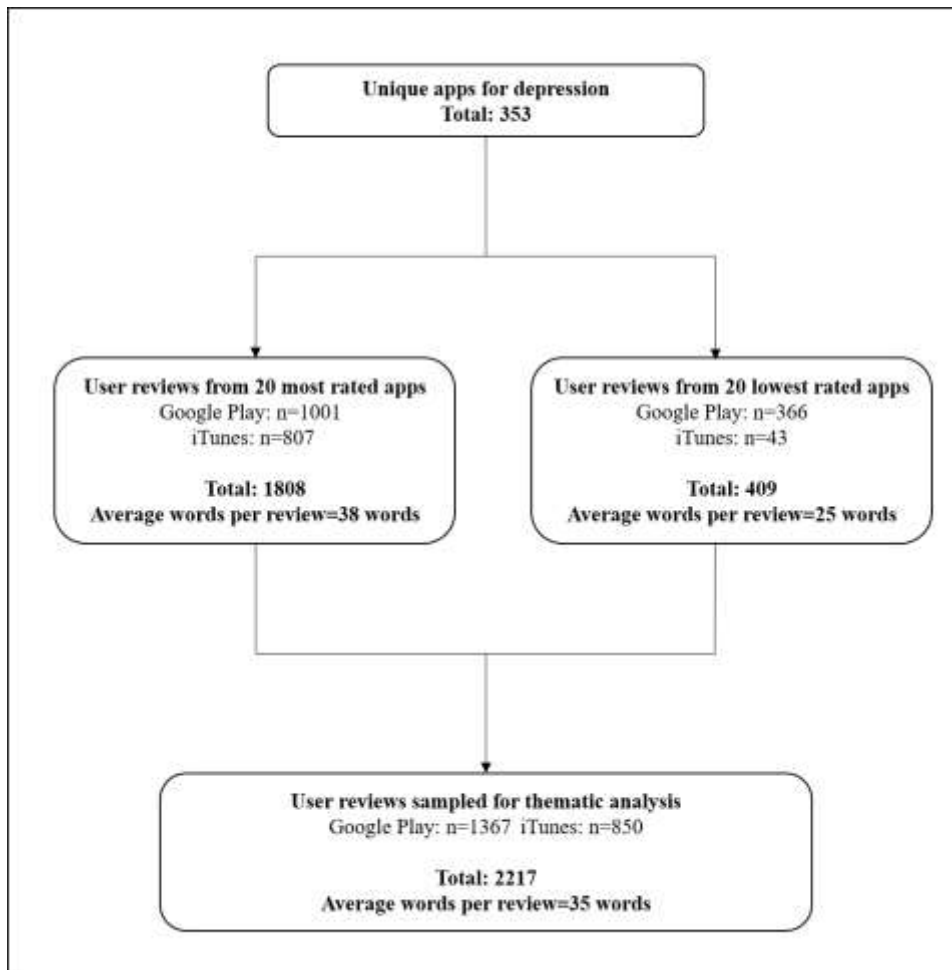


Figure 1: Sampling flowchart for user reviews

“[This app] is extremely helpful in the sense it’s someone to talk to at 3 AM when everyone else is asleep. It gave me a mantra I really needed to hear when admitting I wanted to self-harm and even suggested a hotline.”

Users often expressed feeling they had no one else to turn to and sought comfort in the app to help them cope with challenges or moments of distress. More generally, apps helped users with their overall wellbeing, including personal development, which at times led to increased confidence, self-efficacy, and insight:

“This is an amazing app that will help you focus more on yourself and your own happiness. Starting with the fact that it lets you rate your mood and write about your day and ending at letting you set tasks for yourself and motivating you. It has really helped me to love myself and learn how to have a better lifestyle...I recommend this app to anyone and everyone, because writing your feelings down and learning to love yourself really will help you gain more confidence and trust in yourself. You will feel more secure.”

Apps for depression also had positive effects on users' behaviors, thoughts, and emotions. For some, apps helped with behavioral activation through encouragement of simple activities and small steps to feeling better:

"I was getting into one of the worst wrecks of depression and anxiety and this really helped me. Not only have I been taking walks every day and enjoying the beauty in the world instead of the negativity, I've also been more mindful, healthy, and getting back into my love of exercising and yoga."

Many users described how using apps for depression helped them identify and understand their patterns of thinking, leading to better management of thoughts, more positivity, and change in perspective. Users reported improvements in emotion regulation, from greater emotional awareness to better management of affect.

"It helps to be able to track my moods and visually see that I'm happier, calmer, and more positive than I believe! And when I am down or anxious, it helps to track what is affecting my mood and understand my mind and thought process better than I ever have!"

Emotional and bodily awareness was achieved through diverse methods including actively tracking thoughts and moods to increase insight, to mindfulness approaches.

4.2.2 Adverse events, risks, and safety

Less commonly, user reviews highlighted negative outcomes, risks, and safety concerns of apps for depression (10%, 220/2217). Negative feelings associated with use of apps were attributed to a range of factors, including discomfort with completing aspects of the intervention, difficulties accessing the intervention, poor app or intervention design, and poor quality of support.

"I'm going through a particularly tough time so I'm having to uninstall this app for now. I only made it through about 7 days of tracking my moods, because seeing such a long streak of very sad to mediocre moods has actually made me feel worse about my situation."

Some users reported experiences of bullying or mistreatment by in-app support:

"So, in theory this is a brilliant app, but I do have some queries about the idea of a 'trained' listener. I understand that it's free to use and therefore don't expect anything miraculous, but to be told by a listener that I deserved to be bullied, that there are so many people in the world worse off than me, and that I need to open my eyes to the world because there are people with no money, no legs etc. really didn't make me feel best pleased...Maybe the training needs to be better or conversations need some sort of monitoring?"

Table 1: User perspectives and experiences of the benefits and harms of apps for depression

Ethical Issue	User experiences
Benefits	<p>Benefits to mental health and wellbeing</p> <ul style="list-style-type: none"> • Help managing mental health disorders • Help at different stages of disorder (prevention, treatment, recovery, relapse prevention) • Support or connection to services during crises • Increased emotion regulation • Skills building and improved problem-solving and resilience • Increased sense of balance, focus, and motivation • Increased gratitude, self-acceptance, and compassion • Benefits to spirituality <p>Benefits to physical health</p> <ul style="list-style-type: none"> • Help managing or treating general health concerns • Help managing chronic pain • Improved sleep hygiene
Adverse events	<p>Negative feelings associated with use</p> <ul style="list-style-type: none"> • Feeling stressed while using app • Feeling worse after using app • Triggering and worsening mental health problems through hearing others' mental health difficulties via peer support • Feeling rejected after using app <p>Bullying or harassment</p> <ul style="list-style-type: none"> • Unwanted sexual advances from in-app support or communities • Rude, judgmental, or inappropriate comments from others • Trolling behaviour
Perceived or experienced risks	<p>Unhelpful or harmful advice, including inappropriate in-app or peer support</p> <p>Misdiagnosis</p> <ul style="list-style-type: none"> • Errors in collection and reporting of user data • App-generated diagnosis based only on user-reported symptoms <p>Failure to deliver elements of intervention, at times because of technical difficulties</p> <p>Overreliance on app and over access to therapist or app creating dependency</p>
Safety and safeguarding	<p>Monitoring and regulation of apps and app stores</p> <ul style="list-style-type: none"> • Insufficient moderation of in-app communities • Misuse of peer support apps and online communities by others • Insufficient monitoring and response to safety concerns in user reviews • Insufficient regulation of exploitative or unsafe apps <p>Transparency and awareness of app limitations</p> <ul style="list-style-type: none"> • Knowledge of risks and limits to use • App disclaimers or cautions • Recognizing that apps do not replace in-person care <p>Signposting and safeguarding</p> <ul style="list-style-type: none"> • Assessing user risk and needs • Referral to appropriate services • Lack of transparency over refusal of care

These types of negative experiences were especially evident in apps with peer support. While some users used peer support communities to offload their emotional burdens by venting, others expressed wanting more support and responsiveness, leading to feelings of rejection when not received. Feelings of rejection were also voiced when users were

declined treatment or when in-app support was delayed. While some users described adverse effects of apps for depression, others outlined risks that could lead to potential harms. These included the provision of harmful advice and the potential for misdiagnosis:

“I'm not sure why it's been saying I'm 'hardly distressed' for the past few days when I've been nothing but the opposite! Also confused as to why my doctor's letter says I haven't been having problems with sleep when I certainly do.”

App errors posed additional risks in the failure to deliver essential elements of interventions, such as prompts to take medication:

“This used to work fine, but now it doesn't remind me when it is supposed to remind me. Sometimes it is off by an hour. Sometimes it is off by SEVERAL HOURS. Sometimes I never get alerts. What's the point in having the app if it doesn't remind me??”

A lesser reported risk was the potential for apps to create overreliance in users, which may negatively affect intervention effectiveness, outcomes, and user autonomy:

“You should have spaced out and structured sessions with suitable reflection time in between. Also, having a 'therapist' readily available just creates dependency and doesn't help you to recover. There's a need for boundaries.”

As this user quote states, dependency on apps for depression has the potential to negatively impact intervention effectiveness, mental health outcomes, and user autonomy that is crucial in self-care.

Risks and adverse events highlight the great importance of safety and safeguarding in apps for depression. In this context, safety is ensuring apps are safe and free from known risks and harms, while safeguarding refers to specific measures to protect vulnerable people from harm or abuse. Few users described apps as a safe space or designed with safety in mind. Several reviews expressed concerns that apps were being misused by some users:

“I came on here for genuine help, and it seemed nobody was capable of taking me seriously, and the group chats are full of 12-year-olds who are making jokes of mental health issues.”

Some users highlighted the importance of being aware of app limitations. Potential users were advised to seek in-person care as needed, or to contact emergency services when in crisis. Few users posted triggering content in app store reviews. It was unclear whether users who posted reviews containing safety concerns were contacted privately by developers or app stores. A small number of users called on app stores to take more responsibility to regulate or ban apps that are exploitative or unsafe.

“More deceptive crap. Says free, but then turns the screws to make you pay up \$69. It is shameful that Google Play doesn't enforce honesty by its vendors”

Users also demanded greater protections within apps given their use by vulnerable people. This was not limited to instances of bullying and abuse, but also included a need for appropriate referrals and explanations for why users were refused care. The latter may reflect attempts to safeguard users who were not suited to an app intervention, but typically left them feeling rejected and helpless:

“It is a horrible feeling when the problems that are dwelling inside of you seem to demand an anonymous form of assistance. But when you seek and find a solution you are told they won’t help you. They do not even give a reason, just to seek face-to-face therapy. If I felt that was the solution, then I wouldn’t have tried to download the app. To give people hope when they are at their lowest then to refuse on an unknown basis just compounds the feelings of hopelessness and worthlessness they are already feeling.”

4.3 Facilitators of benefits and harms

4.3.1 Usability

Benefits, avoidance of harms and risks, and safety are the foundation of ethical apps for depression. In mobile mental health, these areas are also indirectly affected by usability, design, and support (Table 2). Usability was the most common theme emerging from user reviews (43%, 940/2217), with users reporting both positive and negative experiences. Usability had implications beyond app functioning, with poor usability impacting access to support, reliable delivery of interventions, and accurate data collection and reporting. These issues pose potential risks and harms to users. Thus, while usability is not a conventional ethical concept, it can be considered a structural factor with ethical implications.

Most reviews on app stability highlighted issues with technical difficulties affecting the use of the app. When apps worked as intended, many users found them to be simple and easy to use. This was not the experience for all users with some apps described as too complex. User experiences were also impacted by app updates or lack thereof. A small number of user reviews expressed appreciation for updates which fixed technical issues, improved app content and design, and enhanced app functioning. For some apps, lack of updates negatively affected performance. For others, updates brought changes that negatively impacted the app or intervention:

“Ever since the newest update it has been getting my moods wrong. I said “very bad” and it told me I was very good. I then put “bad” and it said good. It’s really annoying and it’s saying I’m hardly distressed even though I have been feeling horribly recently, it’s like you’re mocking me.”

Discussions on usability also pertained to device functionality. This includes how the app functioned on specific mobile devices as well as how apps affected a device’s functioning. Few reviews described issues with devices or other apps affecting app functionality:

“Lost progress that took a long time because I made a phone call between step 2 and 3.”

Loss of progress or app data was also related to poor error recovery. Some users voiced frustration at being unable to edit user data such as messages, diary entries, and logs:

“Fantastic app but no undo button! Extremely risky - I spent days logging my meds - all perfect, lovely interface, then ONE TIME accidentally clicked the wrong thing and there is NO WAY to undo this - so then you're on your own remembering your meds!”

Other difficulties in this area included being unable to restart tasks, with some users feeling restricted by app inflexibility and permanence. Some users appreciated the range and flexibility of in-app customizations, but a minority thought apps had too many options leading to unnecessary complexity and negative user experience.

Table 2: User perspectives and experiences of the usability, design, and support of apps for depression

Ethical Issue	User experiences
Usability	<p>App stability and functioning</p> <ul style="list-style-type: none"> • Difficulties opening, signing up, or logging in to app • Glitches, crashes, or freezing • Poor responsiveness • Element not working, e.g., sounds, notifications, recording of data, in-app communications • Issues with app updates • Poor error recovery <p>Ease of use</p> <ul style="list-style-type: none"> • Overly complex app or intervention • Inadequate guidance on how to use • Difficulty navigating interface and features <p>Device functionality</p> <ul style="list-style-type: none"> • Challenges with app working on specific mobile platforms • Difficulties with app operating on specific devices • App interferes with basic phone functions, e.g., calls, wakes up screen when phone not in use, sounds and volume settings, battery life and memory • App disabled device, requiring reboot or uninstall • Other apps on phone impaired functioning of app for depression <p>Customization</p> <ul style="list-style-type: none"> • Options to change app aesthetic, sounds, and notifications • Flexibility with intervention content and delivery • Need for greater customization and control over in-app elements
Design	<p>Design and content of the intervention</p> <ul style="list-style-type: none"> • Appreciation for the intervention’s concept, content, and resources • Intervention too generic and lacking depth • Concerns with app validity, theoretical orientation, or evidence base • Concerns with assessment questions and validity of results • Need to be comparable to other digital or in-person interventions <p>Specific elements and features</p> <ul style="list-style-type: none"> • Persuasive design features • Need for more meaningful data and options for data collection and storage
Support	<p>Developer support</p> <ul style="list-style-type: none"> • Addresses issues reported in user reviews or in-app error reports • Use of app store to reply to reviews and to get more information about issues • Fix usability issues and provide customer service support • Lack of support and updates suggested developers abandoned app <p>Therapeutic support</p> <ul style="list-style-type: none"> • Need for human support • Need for more frequent and consistent delivery of support • Need for support at specific points in intervention, e.g., after self-assessment or adversity • Inadequacies in the quality and competence of support • Need for greater vetting and regulation of sources of support • Inappropriate chatbot responses showing a lack of understanding and inadequate support <p>Social support</p> <ul style="list-style-type: none"> • Facilitates connection with others in-app or real-world settings • Receiving help from ‘like-minded’ people which helps users to feel less alone • Providing support via peer support resulting in empowerment, empathy, and better relations • Improved relations with and support from loved ones

4.3.2 Design

Like usability, design (30%, 657/2217) had indirect ethical implications related to concepts such as validity, accuracy, and safety. Some negative experiences in this area stemmed from disagreement with the intervention content:

“Waste of space on your phone! Basic and generic content with no actual information about mental health services. Under the crisis section it lists the same patronizing advice as the crisis team offer, i.e., go for a walk or have a bath...it doesn't even give the contact numbers for the local crisis teams. Snake game won't help you relax when it stops every 3 apples to give you a 'top tip'. Written by those who've obviously never experienced mental illness.”

While some users believed in the validity of the apps, others expressed concerns with the theoretical orientation or evidence base, while others critiqued the design of assessment tools:

“There needs to be a 'non-applicable' option on the answers. False scores can be generated when I'm answering 'not at all' when it's just because I haven't been in the situation.”

There were also some concerns with using apps for self-diagnosis, with a few users highlighting the potential for invalid results due to false reporting by users. Some user reviews highlighted areas of app design which enhanced user experiences and treatment delivery. This included persuasive design techniques such as notifications and reminders, tailored interventions, tunnelling, gamification, or self-monitoring and surveillance. Some users expressed a desire for more persuasive design features. A subset of users wished for more meaningful data, ranging from data on app usage, before-after data to monitor the effects of intervention activities, more open-ended data entry for better self-expression and accuracy, and improved data storage and long-term data collection.

4.3.3 Support

Another factor which greatly impacted benefits and harms of apps for depression was the provision of support (38%, 838/2217). Support for users fell within four broad categories: developer support, therapeutic support, social support, and support for real-world care.

User reviews provided users with a platform to not only voice their experiences and concerns with apps, but to reach out to developers in the hope that they may address these difficulties. Some users reported positive experiences of developer support, via response to their app store reviews or other in-app methods of reporting issues. Developer support was typically needed to address issues with usability but also extended to help resolving errors with payment. Several users reported difficulty in accessing developer assistance, with some expressing frustration with the lack of responses to queries. This, along with app instability and lack of updates, led users to believe some developers had abandoned the apps and their user base.

Users also equally discussed the importance of therapeutic support. Therapeutic support in apps for depression included online therapy with qualified counsellors, peer support, and in-app support provided by chatbots. Several users were satisfied with the therapeutic support received. Others noted shortcomings and additional support needs:

“Advice is general and mediocre at best. Do yourself a favor and go see a real expert. It seems to me that these 'therapists', due to their lack of knowledge and expertise, don't have many patients and they are not very successful in their job, so they have to make money in this app. The customer service was also terrible. I requested for therapist change, but it didn't happen.”

As a result, some users emphasized that in-app support was not equivalent to in-person care. Concerns about the competence of in-app support were not limited to therapists, with a small number of reviews outlining concerns with chatbot communications:

“I like this app other than the fact that it doesn't understand what the heck I'm saying. Like we were talking about how my day was and I said it was good and the AI asked why and I said 'idk.' Then the AI was like 'That's not good. Come on, let it all out' and I was like 'Huh?' Because that makes NO sense. And the AI was like 'That can't be easy. Is there more?' This is really not good because this shows me that lots of the responses don't look at what you said and just say things responding to what they think you'll say.”

This led some to describe interactions with chatbots as scripted or robotic. More generally, some users expressed dislike of therapeutic support which they found to be patronizing, impersonal, or inauthentic. However, more users described having a positive therapeutic alliance which was genuine and made them feel heard and understood.

For some users, apps helped to increase their social support and connections with others. Feelings of increased understanding from others, social connections, and community were common benefits of apps with peer support. These benefits extended beyond the app to impact some users' real-world connections. For a subset of reviewers, using apps for depression made it easier to talk about their mental health difficulties with loved ones. Apps also helped some users to feel more comfortable sharing information about their mental health with their real-world healthcare provider. Users described previously having difficulty getting their primary care provider to understand their challenges, and felt that the app made this easier:

“The referral letter at the end of my first week really helped as it got my GP to listen to me and I am now on a waiting list for professional psychological help.”

4.4 Justice and rights

4.4.1 Autonomy

Autonomy (the capacity to make informed decisions free from coercion or deception) is an important concept in self-care and mobile mental health, as reflected in its presence as a major theme in user reviews (33%, 734/2217). In this context, the concept of autonomy centered on four main areas: app choice, treatment options, in-app options and customization, and the user voice.

The depression app marketplace lets individuals take an active role in researching and selecting app interventions. Some users embraced this freedom of choice and wished to shop around for the best app to meet their needs. Several users reported having tried similar apps before finding the one that they preferred, with some users expressing a desire to try apps before buying:

“No thanks. It says free but everything I clicked on was not available for preview. It should say 'lite' so I know it's a preview and it should walk me through what I can do with the full version. This app is annoying. I'm not paying 69.99 until I know it can deliver what I need.”

Apps also facilitated users' engagement in treatment planning, with options to select treatment paths, goals, in-app support, or frequency and duration of use. This placed some responsibility on users to take an active role in treatment decisions and fit of care:

“As far as choosing a therapist, you should read their bios to see if they are a fit for you. I chose a therapist whose professional background is related to mine and who is the same demographic as I am so that I knew I could relate to her. While everyone doesn't have that luxury, I think [this app] is for people who has at least a little bit of self-awareness.”

Apps viewed as lacking in treatment options or flexibility resulted in some users feeling forced to complete aspects of the intervention against their will:

“I find this a really useful little app for downloading thoughts and feelings. Would be even more useful if you could set the prompt timings yourself rather than be a slave to the app.”

Outside of treatment choices, apps also offered users options for customization of features (discussed under ‘usability’) which let users make apps better fit their needs and individual preferences. Again, a lack of options in this area led some to voice frustration and feelings of being limited in choice.

A unique aspect of publicly available mental health apps is the opportunity for users to express themselves in app store user reviews. This gives users a voice to share their treatment experiences and needs, to report grievances, to help others in selecting treatment options, and to potentially influence future app design and development. User reviews helped potential users to select app interventions, often providing explicit recommendations for use, or recommendations for alternative treatments. In this manner, user reviews functioned as a community of peer referrals and support.

4.4.2 Access

Apps for depression provided many users with increased access to care with almost 30% of users (645/2217) discussing some aspect of this theme in their review (Table 3). This reflected the potential of mobile mental health to reduce barriers to care that may affect the more vulnerable groups. Many users sought help for their mental health difficulties via apps due to difficulties accessing standard care. Apps were also a preferred alternative to in-person care, for people who wished to avoid treatment as usual or human support. App interventions were said to be accessible whenever and wherever needed, with many describing them as a therapist in their pocket. Users appreciated the expediency of apps compared to standard care and the frequency of contact from in-app support. However, some users expressed disappointment when they did not receive support as promptly as expected:

“My therapist seemed well intentioned, but I had to keep poking him to check in or talk, and my experience confirmed a lot of what bad reviews said about it. In no way can I speak with certainty, but I just got the feeling I was just one name on a big list and so the care wasn't personalized. I informed my therapist of my intention to quit and at the time of writing he hasn't responded.”

Table 3: User perspectives and experiences of access and privacy of apps for depression

Ethical Issue	User experiences
Access	<p>Difficulties accessing standard care</p> <ul style="list-style-type: none"> • Difficulty physically accessing in-person care • Time demands such as childcare or work • Treatment costs <p>Preference of apps over in-person care</p> <ul style="list-style-type: none"> • Past negative experiences of standard care • Avoidance of human support or social anxiety • Comfort of online therapy <p>Barriers to apps for depression</p> <ul style="list-style-type: none"> • Asynchronous support and delayed therapist response • Disruptions to continuity of intervention, e.g., technical difficulties, app updates, unnotified termination of support, and unexpected costs
Privacy	<p>Issues related to the protection of personal details and identifying user data</p> <ul style="list-style-type: none"> • Secure passwords and app lock and save • Anonymity • Collection of too much sensitive user data • Concerns with data security and the sharing of data with third parties • User control over data and what to share with whom

Access to care was also negatively impacted by disruptions to the continuity of care. Some users reported disruptions to app interventions including unexpected costs or paywalls:

“I really needed to talk to a professional therapist. I do not have money to see one currently, so I was going to try out the 3-day free trial. Until I was told I would have to pay \$150 dollars up front for the 3-day ****FREE**** TRIAL. It upsets me because people like me need the help and want it and cannot get it. Because no one cares about you unless you have money to give.”

Other barriers included age restrictions and refusal of care by apps, presumably due to concerns with safety or suitability of care:

“Apparently a 15-year-old female isn’t allowed to need help...it says 12+ in the description so I don’t know why it didn’t let me get matched, false advertising. I wouldn’t trust this app.”

4.4.3 Commerce

Over 20% (21%, 463/2217) of user reviews commented on matters related to apps’ costs, business models, and consumer rights. This was one of the most passionately discussed themes with users expressing strong opinions over the pricing and billing practices of apps. Some users believed apps for depression were more affordable than standard of care, but others thought apps were too costly. This was associated with an unwillingness or an inability to pay for treatment, with many believing that mental healthcare should be free:

“I love this app and it has helped me in many ways, but I am **DISGUSTED** that you want me to pay for good mental health.”

Some users compared app pricing with other apps or online interventions, expressing an unwillingness to pay more when they believed they could access similar content for free elsewhere. Users expressed a desire for more free content in apps, longer trial periods, and greater flexibility in payments.

Apps' costs and billing practices form part of their business models, with some developers disclosing that users' financial support (via payment) was needed to maintain app operations. Other apps included ads to partially fund costs. The adoption of a commercial business model to healthcare was not always well received by users:

“This app WAS incredibly helpful. And then it updated! Now you hit a money wall at every turn...Now every time I'm on it I feel worse by the time I'm done...Pull down some of the money walls and actually help people. Instead of money-grabbing. You were once an app I believed wanted to help people oh, now, not so much.”

Apps business models and practices greatly influenced how users perceived the app, its developers, and their intentions. This was not always negative, with users expressing appreciation for apps with financial aid or flexible pricing options.

4.4.4 Privacy

Less than 5% of user reviews (4.8%, 107/2217) made mention of privacy. In the context of apps for depression, privacy pertains to the respect and protection of users' information, including personal details, identifying user data, intervention data (whether collected actively through user entry or passively via apps), and usage data.

A minority of users praised apps for keeping their details private and considered apps and their data to be secure. Users appreciated anonymity which they believed helped them to be more open in expressing themselves and seeking help and made them feel safe. This was also reflected in reviews themselves, with many users sharing personal and sensitive information in app store reviews under anonymous usernames.

However, most user reviews discussing privacy highlighted concerns. While some users found app privacy policies to be accessible and easy to understand, they did not always agree with the practices outlined:

“Looked like it could be really helpful with the two-week tracking and depression/anxiety scales, but I never even got that far since the first thing you have to do is accept a ‘privacy’ policy that includes, among other things, using your data for Facebook advertising and anonymous research. The latter is not a huge issue with me, but it could be with others. I definitely have an issue with using my data on a very personal app for advertising though, especially when there's no obvious way out.”

A key aspect of privacy was the desire of users to be in control of their data, from what is collected, to how it is stored and shared. Some users needed increased data protection, while others overlooked potential privacy issues with requests for cloud storage:

“I would love to rate it 5 stars but it's missing out on one feature, I really wished I could back my diary up in my Google Drive, so I don't lose all of my diary entries.”

4.4.5 Respect

Although respect was a minor theme in user reviews (1%, 31/2217), it is an important element of ethical apps for depression relating to inclusiveness, accessibility, and respect for the rights and dignity of all people. Few apps were praised for their inclusiveness and efforts to improve accessibility:

“Loved it. The app even has a chat for LGBTQ+ and teens especially.”

More often, reviews highlighted issues in these areas, with users expressing need for greater cultural awareness and suitability of support, increased language options, and accessibility for users with impairments:

“When you ask to chat, you are connected with a person from Asia...They have no understanding of European life or European socioeconomic problems. They have a completely different culture and lifestyle...so they are unable through no fault of their own to give you any better advice.”

“This app needs support for the hard of hearing like myself. I struggle to hear the voices clearly with the ambience in the background due to the lack of subtitle support for deaf and hard of hearing users. Please consider making this app accessible to those who don’t have ears as good as yours.”

4.5 Virtue of apps for depression

4.5.1 Transparency and Trust

The themes of transparency and trust emerged in almost a quarter of our user reviews (23%, 509/2217), with largely negative experiences being reported (Table 4). Several users outlined an insufficiency of information regarding app costs and billing practices, treatment processes and access to care, or elements of support. With respect to costs, users’ main concerns surrounded hidden costs, paywalls, and unexpected charges. This had implications on access to care:

“They make you write out personal statements on your mental health and fill out a questionnaire before telling you must pay \$70 a week to get help. Absolutely cruel.”

Table 4: User perspectives and experiences of transparency and trust in apps for depression

Ethical Issue	User experiences
Transparency and trust	<p>Insufficient information</p> <ul style="list-style-type: none"> • App costs and billing practices • Treatment processes and access to care • Unexpected changes to support • Unnotified changes to app design or content <p>Reduced trust in apps or developers</p> <ul style="list-style-type: none"> • Lack of fidelity, i.e., the app did not do as it said it would • Experiences impacting perceptions of developer’s motives • Ethical or legal concerns, such as illegal billing practices

Few reviews highlighted the importance of information, and awareness of the limitations of apps for depression to manage expectations and experiences. While user reviews helped potential users gain knowledge and perspective about apps for depression, transparent information on app costs, treatment details, expected outcomes, and limitations should be readily available from the app developers and the app itself.

Issues with transparency affected some users’ trust in apps and developers. Several reviews showed users’ trust to be impacted by fidelity. This included apps providing the support promised and achieving the expected results. In cases where the app was not as promised, some users questioned app or developers’ motives. Negative views of developers’ motives were most often related to beliefs that the app was created to exploit vulnerable people for financial gain:

“Full of trolls and mean people, support system does absolutely nothing to support you other than ask you to donate money to them in exchange for you to further customize your personal posts. Whole app is a scam to try to make money off of mentally ill people and does not provide any help for you at all.”

Users were more likely to express trust in the app and positive perceptions of motives when they had positive experiences or outcomes from using the app. In these cases, users described developers as humanitarians working for the greater good:

“Nice to have someone who always wishes the best for us. Great respect and admiration to the developers for showing so much passion towards making the world mentally healthy. Sometimes, the fact that people like you exist gives me so much faith in humanity.”

4.5.2 Social impact

A minor theme in user reviews was the wider social impact of apps for depression beyond the individual user (0.8%, 18/2217). These users believed apps for depression had the potential to positively impact attitudes towards mental health difficulties, reducing stigma and normalizing mental healthcare:

“The trouble is most people don't understand the giant issue we have as a race. Mental illness doesn't mean you're crazy, it means you got a cold, you need to heal now. That's a metaphor to help get the point across. Apps like this can actually save a life.”

There was belief that apps had the potential to positively impact communities and wider society:

“Overall, I believe [this app's] concept can bring significant changes in mental health globally and provide basic support and education to the majority of population.”

As such, developers were urged to consider their civic duty and the social impact of apps they develop:

“I think if [this app] considers its global impact more, perhaps it could make an even greater difference.”

5 DISCUSSION

5.1 Summary of findings

We reviewed and analyzed user reviews of publicly available apps for depression to capture user experiences, evidence of ethical issues in app store user reviews, and key elements of ethically designed apps for depression. This study provides a novel contribution to the literature on the ethics of mobile mental health. To the best of our knowledge, our study is the first to analyze user reviews of apps for depression for themes related to user experiences of ethical issues. Findings captured diverse user perspectives of apps for depression and how their design, development, and delivery impacted user experiences and wellbeing. The framing of ethical issues within user reviews provided concrete examples of ethical concepts which can sometimes be too abstract and ambiguous for everyday application. This research captured key elements that should be included in ethical apps for depression as reflected in reviews of real-world users and people with lived experience of mental health difficulties. These were: benefits, anticipation of risks, safety and safeguarding, usability, design, support, access, autonomy, privacy, fair commerce, transparency and trust, and social impact. These findings and their implications are discussed in greater detail in the following sections.

5.2 User experiences of apps for depression

The sample of user reviews of apps for depression was largely positive, with less than half of all reviews categorized as negative or ambivalent. Findings showed several factors that impacted user experiences and provided insight into what users considered to be elements of ‘good’ apps for depression. Some of these elements reflected common themes in user reviews of mental health apps, notably mention of app usability, design, costs, developer support, and privacy [1, 48]. Beyond this, our study captured themes specific to apps for depression, extending the findings of the one previous user review study in this area [65]. Limiting their review to user experiences of CBT apps for depression, Stawarz et al. [65] highlighted themes related to context of use; privacy, security, and trust; engagement features; and attitudes to non-CBT therapeutic features. By sampling from all apps for depression regardless of treatment approach, our study provides insight into user experiences across the spectrum of treatment options in the depression app marketplace.

Our findings showed how user experiences were impacted by more than just the app itself (i.e., the product), with users also commenting on an app’s purpose or developers’ intentions, the impact of specific processes in the app or intervention, and outcomes of use. Users believed apps should have a clear purpose centered on helping those in need, with positive experiences, gratitude, and support voiced for apps which were thought to be designed from a place of care and good intentions. This was captured in our theme of virtue, related to ideas of transparency and trust. The concepts of transparency and trust were more nuanced in user reviews than in the literature which tends to focus on issues of privacy, security, benefits, and safety [69, 79]. While these elements also emerged in our thematic analysis, users’ trust in apps and developers were often tied to users’ perceptions of developers’ motives, commonly influenced by their views of app costs, business models, and developer support. Some users alleged apps were scams based on their negative experiences of payments and subscriptions, while others made this accusation for apps with paywalls, limited trials, and misleading offers of free content. For some, paid apps reflected developer greed, with more trust in free apps which were thought to arise from developers’ good will. When compared to the lower prevalence of discussions on privacy and security, the findings show a disparity between users’ concerns in this area and the literature [27, 44]. This highlights a need for further research into the concept of trust in mobile mental health in the wild, to explore the many factors impacting user trust and their interrelations, as well as their impact on the use and adoption of mobile mental health.

User experiences were also impacted by specific processes in apps for depression, including steps required to access app interventions, safety and safeguarding measures, and methods of reporting and receiving support for issues and concerns. In general, the easier and more transparent processes were, the more positively they were experienced and reviewed by users. Users reported positive experiences of apps which provided guidance and support in these areas through in-app moderation, transparent safeguarding practices, and developer response to user concerns. Despite this, user reviews highlighted the need for more processes and measures to protect users from risks and harms and to provide them with enough support, especially considering the increased user responsibility evident in user reviews. Users’ responsibilities included choosing app interventions from the app marketplace, selecting treatment options and in-app support, providing support to others in peer support apps, ensuring appropriate use of the app and correct data entry, and reporting errors to developers. Some user reviews even suggested users were responsible for safeguarding oneself by understanding the limitations of apps before use. While some research has discussed user responsibility as a benefit of mental health apps [19], little is known of the potential impact of this responsibility on treatment outcomes, potential risks and harms, and the degree of support needed to ensure safe and effective use of apps. User reviews also demonstrated the importance of clear and complete communication regarding app and intervention processes, particularly in areas impacting access to treatment. Users appeared to be more understanding of limits to care when clearly explained but expressed strong disapproval and feelings of rejection when this was unclear. In line with traditional mental healthcare standards [3, 16], it is crucial to

actively involve users in safeguarding decisions and practices. This becomes even more relevant given the increased user autonomy and responsibility of mental health apps, where users are active participants in their treatment processes and decisions.

Ease of use and good product usability were also key to positive user experiences, with apps for depression well received when they were thought to be interactive, enjoyable, and easy to use. This was an important aspect of apps for depression, with user reviews prioritizing both usability and design [48]. Despite the prominence of these themes in reviews, users were found to be forgiving of errors and app instability when they had an overall positive experience of using the app, at times due to a positive alliance with in-app support or to positive outcomes. Not surprisingly, the effects of using apps for depression had a prominent impact on whether apps were perceived favorably, with users who benefited from use describing positive user experiences, while the inverse was true for perceived risks and adverse events. User experiences are therefore complex and influenced by the interplay of several factors. Findings suggest a possible hierarchy of how these elements are valued by users relative to their individual needs and preferences, with support and benefits seemingly the most important factors for positive experiences of apps for depression. More research is needed in user values in mobile mental health, and how these impact user expectations, use, and adoption of apps for depression.

5.3 Ethically designed apps for depression

The framing of user experiences according to purpose, processes, product, and outcomes aligns with ideas in responsible innovation [50, 71], which encourages a broader consideration of how the design of technological innovations creates positive impact. As our findings indicate, it is not enough to solely focus on the app itself, but rather developers must consider the interrelated elements around apps for depression that contribute to user experiences and ethical implications.

Ethics can be a daunting topic, at times presented too abstractly for practical application and other times too rigidly with the use of prescriptive ethical principles and codes [29]. Our work provides a fresh approach in its use of user reviews to explore users’ perspectives and experiences of ethical issues in apps for depression. We present these concepts as elements of ethical apps for depression and identify heuristics for supporting their design (Table 5).

Table 5: Elements of ethically designed apps for depression

Element	Description
Benefits	Apps for depression should provide direct benefits to individual users and indirect benefits to communities and wider society
Anticipation of risks	Designers should anticipate and avoid foreseeable risks and harms. Risk anticipation and management should be an iterative and responsive process occurring throughout the app lifespan
Safety and safeguarding	Apps for depression should be designed with safety in mind, with clear measures in place to protect vulnerable people from harm
Access	Apps for depression should increase access to care through the removal of situational and attitudinal barriers
Usability	Apps for depression should be technological stable, easy to use and amend, and should not interfere with a device’s normal functioning
Design	App interventions for depression should be valid, reliable, and evidence based. Apps should utilize appropriate persuasive design features and technological enhancements to increase adherence and adoption
Support	Apps for depression should provide users with adequate developer and therapeutic support and should facilitate improved social support either in-app or in users’ daily lives. Apps should support connections to in-person care should it be appropriate or needed for the individual user
Autonomy	Apps for depression should enable and respect user autonomy

Element	Description
Privacy	Apps for depression should respect and protect users' privacy with clear privacy policies, requests for user data proportionate to need, and robust data security
Fair commerce	Apps for depression should employ fair and ethical business models which avoid conflicts of interest and respect users' commercial and human rights
Transparency	All information and processes should be transparent and easy to understand, including costs, billing, risks, privacy policies, etc.
Trust	Apps and developers should be truthful and trustworthy. Developers should avoid any intentions or actions which may be fraudulent, deceptive or exploitative
Respect	Apps for depression should demonstrate respect for all people, including respect for human rights, diversity, cultural differences, and disabilities
Social impact	Designers and developers of apps for depression should consider their broader social impact and civic duty in the design and marketing of apps

These elements are not intended as strict rules for app design and development, but heuristics to promote reflection throughout the design and development process. They align with and advance the literature on the ethics of mobile mental health [34, 38, 44, 61], and highlight the importance of lesser discussed elements such as access, support, respect, and social impact. These concepts are often overlooked in discussions on the ethics of mobile mental health, where much of the focus is on privacy, data security, benefits, and risks [27, 36, 44, 61]. Our findings are unique in capturing the interplay of ethics of mobile mental health in the wild and the effect ethical elements have on users' experiences and wellbeing.

A key finding from the analysis of user reviews was the interrelations between ethical themes, such that a shortcoming in one element often negatively impacted others. For example, a lack of app updates affecting app functioning pertained to poor usability but also demonstrated inadequate developer support. Similarly, high app costs not only affected users' abilities to access care but may also impact their perceptions of developers' intentions and motives, leading to diminished trust in the app, which may be further generalized across all mental health apps. This has the potential to create both situational and attitudinal barriers to care and thus impact users' help-seeking behaviors and mental wellbeing. As the user quotes showed, these scenarios are not hypotheticals but reflect challenges faced by real people seeking help for real concerns.

These interrelations are further complicated by potential mediating factors. Using the previous example of poor usability, our findings showed that issues in this element impacted user safety and wellbeing if technical difficulties limited access to care or accuracy of user data. Likewise, app costs were more likely to reduce user trust when impacted by limited transparency and inadequate notification of payment processes and business models. In seeking to design ethical apps for depression, there must be greater reflection on, and understanding of how all elements of apps (from purpose to outcomes) affect user experiences and outcomes. The interrelation of ethical themes in mobile mental health is an under researched area warranting greater attention and guidance in navigating ethical design and tensions.

5.4 Ethical tensions in the design of apps for depression

The interrelations between elements of ethically designed apps for depression convey both positive and negative associations. In the case of positive associations, successful implementation of one element (e.g., developer support) would be expected to enhance related elements (e.g., usability). Designers should therefore consider how elements are positively related and use this to strengthen the ethical design of apps.

However, cases of negative associations may prove more difficult to navigate and overcome as designers are faced with conflicting elements, both of which represent an important aspect of 'good' apps for depression. This may result in ethical

tensions akin to moral conflicts or dilemmas [10, 71]. An example of this can be seen in the tensions between access to care and safety and safeguarding. Developers may prioritize access to care by allowing all age groups to use the app without restriction. This may have potential risks and safety concerns if vulnerable groups (such as children and young people) use the app without appropriate guidance or protections [14, 51]. Risks may also increase if app content or interventions are not specifically designed for these groups, e.g., adult-standardized assessment measures. In this scenario, increased access may reduce safety and benefits. Similarly, developers may prioritize safety and safeguarding by implementing strict criteria for access, with users not meeting these criteria not being granted access to the intervention. In this case, an increase in safety potentially reduces access to care, as was the case with apps whose screening intake resulted in many users being refused treatment. Designers and developers may seek to resolve ethical tensions by prioritizing one element over the other (e.g., access vs. safety). This approach, however, may result in ethical shortcomings which may potentially impact other ethical elements. This is an even greater risk given the poverty of research into the interrelations and mediations of ethical elements to guide which factors may more greatly impact outcomes.

Alternatively, we favor the responsible innovation approach [50] which encourages designers to use moral conflicts to inspire, rather than hinder innovation [71]. Ethical tensions present important design opportunities for development teams to resolve through innovative technological design. For example, designers wishing to increase access to care for young people while ensuring safety may incorporate a way to assess a young person's understanding and competence in making decisions regarding their care. This capacity assessment could then be used to determine whether the young person can make an informed choice for care, as per Gillick competence, and granting access based on this finding. Although this competence assessment would likely still restrict access for some users (perhaps with parental consent required in those cases), it would increase both access and safety (and likely autonomy), resulting in a more ethically designed app.

5.5 Implication for the ethics of apps for depression

This research provides several key findings relevant to the design and development of apps for depression. Firstly, users' descriptions of their perceptions and experiences of apps for depression highlight the importance of considering ethical issues across all areas of apps, from their purpose, processes, product, and outcomes. Using a responsible innovation approach [50, 71], developers are encouraged to anticipate potential risks and ethical implications from the stage of conceptualization and throughout the entire app lifespan. User reviews also demonstrated the importance of user involvement in this process, encouraging inclusive deliberation and reflection of these issues and the resolution of ethical tensions. Users must be actively involved in not only the design of apps, but the development of ethical guidance particularly given their increased responsibilities. Crucially, these processes must be iterative, reflecting responsiveness that is key to responsible innovation, to ensure consistent delivery of care and prompt resolution of issues as they arise. It is hoped that the elements of ethically designed apps for depression assist in the design and evaluation of existing and future mental health technologies, and act as a tool to facilitate increased reflection and understanding of these concepts.

Secondly, the findings from the thematic analysis of user reviews have implications for how we view and discuss the ethics of mobile mental health. Few authors have attempted to apply existing ethical frameworks to the evaluation and discussion of ethics in this area [14, 31, 34, 61], typically using principlism [10] or variations thereof [3]. Our findings captured ethical elements aligned with the principles of bioethics. However, user reviews also demonstrated the utility of considering ethics beyond this lens, with value in incorporating other ethical theories which also align with users' experiences and concerns. For example, a notable theme in the analysis was trust and beliefs of developers' motives and morals. Users therefore judged apps and developers for their presumed values, akin to virtue ethics [66]. Users also believed that apps, developers, in-app support, and app stores had a duty to users and responsibilities beyond outcomes alone, akin

to ideas in deontology. Ethical frameworks in mobile mental health should therefore consider the inclusion of values and motives, duty and responsibilities, and principles and elements for ethically designed technologies.

5.6 Limitations

App store user reviews provided a valuable dataset for the exploration of user experiences of publicly available apps for depression. However, this was not without limitations. Firstly, by sampling pre-existing data this thematic analysis was confined to the content and context of the user reviews. Unlike traditional qualitative methods such as interviews or focus groups, we were unable to probe user statements, confirm interpretations of user statements, or further explore specific themes. As such, this study provides a good start for future studies exploring these findings and the interrelations of ethical elements in greater detail.

Moreover, the sampling of pre-existing data prevented the consistent collection of information from all users. User reviews therefore reflect content about apps for depression that users deemed to be important. This varied across users, limiting the ability to make conclusions for the entire sample of reviews. While we used frequencies to determine major and minor themes in user reviews, it is important to note that these figures do not necessarily represent all user experiences as some users may not have commented on all elements experienced. It is therefore important to interpret frequencies as the number of users who discussed specific elements in their reviews, rather than the number of users experienced these elements.

Lastly, this research aimed to sample a cross-section of reviews from apps for depression to capture the range and complexity of user experiences. As such, the sample included both the most rated and the lowest rated apps for depression, and from these samples of the most helpful and most critical reviews as determined by app stores. Due to the nature of app usage and user behaviors, there were a greater number of app store reviews for the most rated apps, with the lowest rated apps being less reviewed and having shorter user reviews. It is possible that this impacted the proportion of positive, negative, and ambivalent reviews in the sample. However, we consider this to reflect the reality of app stores, with an imbalance in how apps are rated, downloaded, and reviewed. This disparity should be kept in mind when interpreting the valence in user reviews.

6 CONCLUSION

Mental health apps have potential benefits in the treatment of depression and the increased access to care. This was reflected in our study, with many reviews expressing positive views of apps and a range of benefits to mental health and wellbeing. User reviews also provided invaluable insight into the challenges users experience when using apps for depression, and the ethical issues encountered. Our study demonstrated the complex interrelations between ethical elements of apps for depression, and the need for designers and developers to consider the entirety of apps and the role they play in users' lives. We presented key elements to be considered in the design of ethical apps for depression and encourage a responsible innovation approach to overcome ethical tensions through thoughtful design and user involvement. This study is novel in its capturing of user voices and experiences of ethical issues in mobile mental health in the wild. It provides developers with a framework and context to guide their design and conceptualization of new ethical mental health technologies.

REFERENCES

- [1] Felwah Alqahtani and Rita Orji. 2020. Insights from user reviews to improve mental health apps. *Health Informatics Journal* 26, 3 (January 2020), 2042-2066. <https://doi.org/10.1177/1460458219896492>
- [2] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders (5th ed.)*. American Psychiatric Association, Washington DC.

- [3] American Psychological Association. 2017. *Ethical principles of psychologists and code of conduct*. American Psychological Association, Washington DC.
- [4] Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2017. Sensitive self-disclosures, responses, and social support on Instagram: the case of #depression. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1485–1500. <https://doi.org/10.1145/2998181.2998243>
- [5] Patricia Arean, Kevin Hallgren, Joshua Jordan, Adam Gazzaley, David Atkins, Patrick Heagerty, and Joaquin Anguera. 2016. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. *J Med Internet Res* 18, 12 (December 2018), e330. <https://doi.org/10.2196/jmir.6482>
- [6] Kennedy Opoku Asare, Aku Visuri, and Denzil Ferreira. 2019. Towards early detection of depression through smartphone sensing. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1158–1161. <https://doi.org/10.1145/3341162.3347075>
- [7] Nilufar Baghaei, Lehan Stemmet, Andrej Hlasnik, Konstantin Emanov, Sylvia Hach, John Naslund, Mark Billingham, Imran Khaliq, and Hai-Ning Liang. 2020. Time to get personal: individualised virtual reality for mental health. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3334480.3382932>
- [8] Sonal Batra, Ross Baker, Tao Wang, Felicia Forma, Faith DiBiasi, and Timothy Peters-Strickland. 2017. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. *Medical Devices: Evidence and Research* 10 (October 2017), 237-251. <https://doi.org/10.2147/MDER.S144158>
- [9] Michael Bauer, Tasha Glenn, John Geddes, Michael Gitlin, Paul Grof, Lars V. Kessing, Scott Monteith, Maria Faurholt-Jepsen, Emanuel Severus, and Peter Whybrow. 2020. Smartphones in mental health: a critical review of background issues, current status and future concerns. *International Journal of Bipolar Disorders* 8, 2 (January 2020). <https://doi.org/10.1186/s40345-019-0164-x>
- [10] Tom Beauchamp. 2003. Methods and principles in biomedical ethics. *J Med Ethics* 29, 5, 269-274. <https://doi.org/10.1136/jme.29.5.269>
- [11] Alfred Philip Benigno, Samuel Lomugang, Patrick Daniel Santos, and Mary Jane Samonte. 2020. E-Psych: a psychiatry support system. In *Proceedings of the 2020 The 6th International Conference on Frontiers of Educational Technologies (ICFET 2020)*. Association for Computing Machinery, New York, NY, USA, 72–76. <https://doi.org/10.1145/3404709.3404744>
- [12] Shalini Bhatia, Munawar Hayat, and Roland Goecke. 2017. A multimodal system to characterise melancholia: cascaded bag of words approach. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI '17)*. Association for Computing Machinery, New York, NY, USA, 274–280. <https://doi.org/10.1145/3136755.3136766>
- [13] Dionne Bowie-DaBreo, Corina Sas, Sandra I Sünram-Lea, and Heather Iles-Smith. 2020. A call for responsible innovation in mobile mental health: content analysis of the depression app marketplace. *Annual Review of Cybertherapy and Telemedicine* 18, 11-16.
- [14] Dionne Bowie-DaBreo, Sandra I Sünram-Lea, Corina Sas, and Heather Iles-Smith. 2020. Evaluation of treatment descriptions and alignment with clinical guidance of apps for depression on app stores: systematic search and content analysis. *JMIR Form Res* 4, 11 (November 2020), e14988. <https://doi.org/10.2196/14988>
- [15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2, 77-101. <https://doi.org/10.1191/1478088706qp0630a>
- [16] The British Psychological Society. 2018. *Codes of ethics and conduct*. The British Psychological Society, Leicester, UK
- [17] Eleanor Burgess. 2019. Collaborative self-management of depression. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 38–42. <https://doi.org/10.1145/3311957.3361851>
- [18] Prerna Chikersal, Danielle Belgrave, Gavin Doherty, Angel Enrique, Jorge Palacios, Derek Richards, and Anja Thieme. 2020. Understanding client support strategies to improve clinical outcomes in an online mental health intervention. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3313831.3376341>
- [19] Liam Crosby and Oliver Bonnington. 2020. Experiences and implications of smartphone apps for depression and anxiety. *Social Health Illn* 42 (March 2020), 925-942. <https://doi.org/10.1111/1467-9566.13076>
- [20] Steven Denisevich and Jichen Zhu. 2019. Interweaving narrative and gameplay to cultivate empathy for anxiety and depression. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY '19 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 247–252. <https://doi.org/10.1145/3341215.3358245>
- [21] Gavin Doherty, David Coyle, and John Sharry. 2012. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)*. Association for Computing Machinery, New York, NY, USA, 1421–1430. <https://doi.org/10.1145/2207676.2208602>
- [22] Jordan Eschler, Eleanor Burgess, Madhu Reddy, and David Mohr. 2020. Emergent self-regulation practices in technology and social media use of individuals living with depression. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376773>
- [23] Weiquan Fan, Zhiwei He, Xiaofen Xing, Bolun Cai, and Weirui Lu. 2019. Multi-modality depression detection via multi-scale temporal dilated CNNs. In *Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop (AVEC '19)*. Association for Computing Machinery, New York, NY, USA, 73–80. <https://doi.org/10.1145/3347320.3357695>
- [24] Joseph Firth, John Torous, Jennifer Nicholas, Rebekah Carney, Abhishek Pratap, Simon Rosenbaum, and Jerome Sarris. 2017. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 16, 3

(October 2017), 287–298. <https://doi.org/10.1002/wps.20472>

- [25] Nidhin Harilal, Rushil Shah, Saumitra Sharma, and Vedanta Bhutani. 2020. CARO: an empathetic health conversational chatbot for people with major depression. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD (CoDS COMAD 2020)*. Association for Computing Machinery, New York, NY, USA, 349–350. <https://doi.org/10.1145/3371158.3371220>
- [26] Ari Hautasaari, Naomi Yamashita, and Takashi Kudo. 2017. Role of CMC in emotional support for depressed foreign students in Japan. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 2614–2621. <https://doi.org/10.1145/3027063.3053197>
- [27] Huckvale K, Torous J, Larsen ME. 2019. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Netw Open* 2, 4 (April 2019), e192542. <https://doi.org/10.1001/jamanetworkopen.2019.2542>
- [28] Anna Huguet, Sanjay Rao, Patrick J McGrath, Lori Wozney, Mike Wheaton, Jill Conrod, and Sharlene Rozario. 2016. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS One* 11, 5 (May 2016), e0154248. <https://doi.org/10.1371/journal.pone.0154248>
- [29] Richard Huxtable. 2013. For and against the four principles of biomedical ethics. *Clinical Ethics* 8, 2/3 (September 2013), 39-43. <https://doi.org/10.1177/147750913486245>
- [30] Jacinta Jardine, Caroline Earley, Derek Richards, Ladislav Timulak, Jorge E. Palacios, Daniel Duffy, Karen Tierney, and Gavin Doherty. 2020. The experience of guided online therapy: a longitudinal, qualitative analysis of client feedback in a naturalistic RCT. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3313831.3376254>
- [31] Nick Jones and Matthew Moffitt. 2016. Ethical guidelines for mobile app development within health and mental health fields. *Prof Psychol Res Pr* 47, 2, 155–162. <https://doi.org/10.1037/pro0000069>
- [32] Hyunggu Jung, Woosuk Seo, and Michelle Cha. 2017. Personas and scenarios to design technologies for North Korean defectors with depression. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17 Companion)*. Association for Computing Machinery, New York, NY, USA, 215–218. <https://doi.org/10.1145/3022198.3026308>
- [33] Nadi Kaonga and Jonathan Morgan. 2019. Common themes and emerging trends for the use of technology to support mental health and psychosocial well-being in limited resource settings: a review of the literature. *Psychiatry Research* 281 (November 2019), 112594. <https://doi.org/10.1016/j.psychres.2019.112594>
- [34] Nicole Karcher and Nan Presser. 2016. Ethical and legal issues addressing the use of mobile health (mHealth) as an adjunct to psychotherapy. *Ethics & Behavior* 28, 1 (October 2016), 1–22. <https://doi.org/10.1080/10508422.2016.1229187>
- [35] Rachel Kornfield, Renwen Zhang, Jennifer Nicholas, Stephen Schueller, Scott Cambo, David Mohr, and Madhu Reddy. 2020. "Energy is a finite resource": designing technology to support individuals across fluctuating symptoms of depression. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–17. <https://doi.org/10.1145/3313831.3376309>
- [36] Kira Kretzschmar, Holly Tyroll, Gabriela Pavarini, Arianna Manzini, Ilina Singh, and NeurOx Young People's Advisory Group. 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. *Biomed Inform Insights* 11, 1 (March 2019), 1–9. <https://doi.org/10.1177/1178222619829083>
- [37] Satish Kumar and Seema Mehrotra. 2017. Free mobile apps on depression for Indian users: a brief overview and critique. *Asian J Psychiatr* 28 (August 2017), 124–130. <https://doi.org/10.1016/j.ajp.2017.03.031>
- [38] Mark Erik Larsen, Kit Huckvale, Jennifer Nicholas, John Torous, Louise Birrell, Emily Li, and Bill Reda. 2019. Using science to sell apps: evaluation of mental health app store quality claims. *npj Digital Medicine* 2, 18 (March 2019). <https://doi.org/10.1038/s41746-019-0093-1>
- [39] Mark Erik Larsen, Jennifer Nicholas, and Helen Christensen. 2016. Quantifying app store dynamics: longitudinal tracking of mental health apps. *JMIR Mhealth Uhealth* 4, 3 (August 2016), e96. <https://doi.org/10.2196/mhealth.6020>
- [40] Emily Lattie, Jennifer Nicholas, Ashley Knapp, Joshua Skerl, Susan Kaiser, and David Mohr. 2020. Opportunities for and tensions surrounding the use of technology-enabled mental health services in community mental health care. *Adm Policy Ment Health* 47 (January 2020), 138–149. <https://doi.org/10.1007/s10488-019-00979-2>
- [41] Cara C Lewis, C Nathan Marti, Brigid Marriott, Kelli Scott, and David Ayer. 2019. Patterns of practice in community mental health treatment of adult depression. *Psychother Res* 29, 1 (January 2019), 70-77. <https://doi.org/10.1080/10503307.2017.1303210>
- [42] Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li, Jing Mei, Jie Zhou, and Henry Leung. 2020. SenseMood: depression detection on social media. In *Proceedings of the 2020 International Conference on Multimedia Retrieval (ICMR '20)*. Association for Computing Machinery, New York, NY, USA, 407–411. <https://doi.org/10.1145/3372278.3391932>
- [43] Jessica Lipschitz, Christopher J Miller, Timothy P Hogan, Katherine Burdick, Rachel Lippin-Foster, Steven Simon, and James Burgess. 2019. Adoption of mobile apps for depression and anxiety: cross-sectional survey study on patient interest and barriers to engagement. *JMIR Ment Health* 6, 1 (January 2019), e11334. <https://doi.org/10.2196/11334>
- [44] Samuel Lustgarten and Jon Elhai. 2018. Technology use in mental health practice and research: legal and ethical risks. *Clinical Psychology Science and Practice* 25, 2 (March 2018), e12234. <https://doi.org/10.1111/cpsp.12234>
- [45] Joshua Magee, Sarah Adut, Kevin Brazill, Stephen Warnick. 2018. Mobile app tools for identifying and managing mental health disorders in primary care. *Curr Treat Options Psychiatry* 5, 3 (September 2018), 345–362. <https://doi.org/10.1007/s40501-018-0154-0>
- [46] Mental Health Foundation. 2016. *Fundamental facts about mental health 2016*. Mental Health Foundation, London.
- [47] Isaac Moshe, Yannik Terhorst, Kennedy Opoku Asare, Lasse Bosse Sander, Denzil Ferreira, Harald Baumeister, David Mohr, and Laura Pulkki-Räback. 2021. Predicting symptoms of depression and anxiety using smartphone and wearable data. *Front. Psychiatry* 12 (January 2021), 625247. <https://doi.org/10.3389/fpsy.2021.625247>

- [48] Oladapo Oyeboade, Felwah Alqahtani, and Rita Orji. 2020. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. *IEEE Access* 8 (June 2020), 111141 – 111158. <https://doi.org/10.1109/ACCESS.2020.3002176>.
- [49] Maria Ludmilla Oliveira, Luis Rivero, José Nunes de Oliveira Neto, Rodrigo Santos, and Davi Viana. 2018. Developing an application for dealing with depression through the analysis of information and requirements found in groups from a social network. In *Proceedings of the XIV Brazilian Symposium on Information Systems (SBSI'18)*. Association for Computing Machinery, New York, NY, USA, Article 8, 1–10. <https://doi.org/10.1145/3229345.3229354>
- [50] Richard Owen, Jack Stilgoe, Phil Macnaghten, Mike Gorman, Erik Fisher, and Dave Guston. 2013. A framework for responsible innovation. In Richard Owen, John Bessant, and Maggy Heintz (Eds) *Responsible Innovation: Managing the responsible emergence of science and innovation in society*. <https://doi.org/10.1002/9781118551424.ch2>
- [51] Chengcheng Qu, Corina Sas, Claudia Daudén Roquet, and Gavin Doherty. (2020) Functionality of top-rated mobile apps for depression: systematic search and evaluation. *JMIR Ment Health* 7, 1 (January 2020), e15321. <https://doi.org/10.2196/15321>
- [52] Chengcheng Qu, Corina Sas, and Gavin Doherty. 2019. Exploring and designing for memory impairments in depression. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Paper 510, 1–15. <https://doi.org/10.1145/3290605.3300740>
- [53] Ana Radovic, Pamela Vona, Antonella Santostefano, Samantha Ciaravino, Elizabeth Miller, and Bradley Stein. 2016. Smartphone applications for mental health. *Cyberpsychol Behav Soc Netw* 19, 7 (July 2016), 465–470. <https://doi.org/10.1089/cyber.2015.0619>
- [54] Amy Leigh Rathbone and Julie Prescott. 2017. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. *J Med Internet Res* 19, 8 (August 2017), e295. <https://doi.org/10.2196/jmir.7740>
- [55] Stefan Rennick-Egglestone, Sarah Knowles, Gill Toms, Penny Bee, Karina Lovell, and Peter Bower. 2016. Health technologies 'in the wild': experiences of engagement with computerised CBT. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2124–2135. <https://doi.org/10.1145/2858036.2858128>
- [56] Hee Rin Lee, Selma Šabanović, Wan-Ling Chang, Shinichi Nagata, Jennifer Piatt, Casey Bennett, and David Hakken. 2017. Steps toward participatory design of social robots: mutual learning with older adults with depression. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. Association for Computing Machinery, New York, NY, USA, 244–253. <https://doi.org/10.1145/2909824.3020237>
- [57] Wawan Rismawan, Carla Raymondalexas Marchira, and Ibrahim Rahmat. 2021. Usability, acceptability, and adherence rates of mobile application interventions for prevention or treatment of depression: a systematic review. *Journal of Psychosocial Nursing and Mental Health Services* 59, 2 (May 2021). <https://doi.org/10.3928/02793695-20201104-02>
- [58] Darius Rohani, Andrea Quemada Lopategui, Nanna Tuxen, Maria Faurholt-Jepsen, Lars Kessing, and Jakob Bardram. 2020. MUBS: a personalized recommender system for behavioral activation in mental health. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376879>
- [59] Jonathan Rottenberg and James Gross. 2007. Emotion and emotion regulation: a map for psychotherapy researchers. *Clinical Psychology: Science and Practice* 14, 4 (December 2007), 323–328. <https://doi.org/10.1111/j.1468-2850.2007.00093.x>
- [60] Farig Sadeque, Dongfang Xu, and Steven Bethard. 2018. Measuring the latency of depression detection in social media. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*. Association for Computing Machinery, New York, NY, USA, 495–503. <https://doi.org/10.1145/3159652.3159725>
- [61] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas. 2019. HCI and Affective Health: taking stock of a decade of studies and charting future research directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Paper 245, 1–17. <https://doi.org/10.1145/3290605.3300475>
- [62] Corina Sas, Kobi Hartley, and Muhammad Umair. 2020. ManneqKit Cards: a kinesthetic empathic design tool communicating depression experiences. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 1479–1493. <https://doi.org/10.1145/3357236.3395556>
- [63] Nelson Shen, Michael-Jane Levitan, Andrew Johnson, Jacqueline Lorene Bender, Michelle Hamilton-Page, Alejandro Jadad, and David Wiljer. 2015. Finding a depression app: a review and content analysis of the depression app marketplace. *JMIR Mhealth Uhealth* 3, 1 (January 2015), e16. <https://doi.org/10.2196/mhealth.3713>
- [64] Katarzyna Stawarz, Chris Preist, Deborah Tallon, Laura Thomas, Katrina Turner, Nicola Wiles, David Kessler, Roz Shafran, and David Coyle. 2020. Integrating the digital and the traditional to deliver therapy for depression: lessons from a pragmatic study. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376510>
- [65] Katarzyna Stawarz, Chris Priest, Debbie Tallon, Nicola Wiles, and David Coyle. 2018. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. *J Med Internet Res* 20, 6 (June 2018), e10120. <https://doi.org/10.2196/10120>
- [66] Torbjorn Tannsjo. 2013. *Understanding Ethics: An Introduction to Moral Theory*. Edinburgh: Edinburgh University Press.
- [67] John Torous, Gerhard Andersson, Andrew Bertagnoli, Helen Christensen, Pim Cuijpers, Joseph Firth, Adam Haim, Honor Hsin, Chris Hollis, Shōn Lewis, David Mohr, Abhishek Pratap, Spencer Roux, Joel Sherrill, Patricia Arean. 2019. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry* 18 (January 2019), 97–98. <https://doi.org/10.1002/wps.20592>
- [68] John Torous, Michael Levin, David Ahern, and Megan Oser. 2017. Cognitive behavioral mobile applications: clinical studies, marketplace overview, and research agenda. *Cognitive and Behavioral Practice* 24, 2 (May 2017), 215–225. <https://doi.org/10.1016/j.cbpra.2016.05.007>
- [69] John Torous and Laura Weiss Roberts. 2017. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. *JAMA Psychiatry* 74, 5 (May 2017), 437–438. <https://doi.org/10.1001/jamapsychiatry.2017.0262>

- [70] Bhavika Vajawat, Prateek Varshney, and Debanjan Banerjee. 2021. Digital gaming interventions in psychiatry: evidence, applications, and challenges. *Psychiatry Research* 295 (January 2021), 113585. <https://doi.org/10.1016/j.psychres.2020.113585>
- [71] Jeroen van den Hoven. 2013. Value sensitive design and responsible innovation. In Richard Owen, John Bessant, and Maggy Heintz (Eds) *Responsible Innovation: Managing the responsible emergence of science and innovation in society*. <https://doi.org/10.1002/9781118551424.ch4>
- [72] Nikhita Vedula and Srinivasan Parthasarathy. 2017. Emotional and linguistic cues of depression from social media. In *Proceedings of the 2017 International Conference on Digital Health (DH '17)*. Association for Computing Machinery, New York, NY, USA, 127–136. <https://doi.org/10.1145/3079452.3079465>
- [73] Yameng Wang, Xiaotong Ren, Xiaoqian Liu, and Tingshao Zhu. 2021. Examining the correlation between depression and social behavior on smartphones through usage metadata: empirical study. *JMIR Mhealth Uhealth* 9, 1 (January 2021), e19046. <https://doi.org/10.2196/19046>
- [74] Daniel Wessel, Helke Kohlbrandt, and Tilo Mentler. 2019. Human-centered development of an activity diary app for people with depression. In *Proceedings of Mensch und Computer 2019 (MuC'19)*. Association for Computing Machinery, New York, NY, USA, 427–431. <https://doi.org/10.1145/3340764.3344434>
- [75] James Williamson, Elizabeth Godoy, Miriam Cha, Adrienne Schwarzentruher, Pooya Khorrani, Youngjune Gwon, Hsiang-Tsung Kung, Charlie Dagli, and Thomas Quatieri. 2016. Detecting depression using vocal, facial and semantic communication cues. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge (AVEC '16)*. Association for Computing Machinery, New York, NY, USA, 11–18. <https://doi.org/10.1145/2988257.2988263>
- [76] Hannah Wisniewski, Gang Liu, Philip Henson, Aditya Vaidyam, Narissa Karima Hajratalli, Jukka-Pekka Onnela, and John Torous. 2019. Understanding the quality, effectiveness, and attributes of top-rated smartphone health apps. *Evid Based Ment Health* 22, 1 (February 2019), 4–9. <https://doi.org/10.1136/ebmental-2018-300069>
- [77] World Health Organization. 2021. *Global strategy on digital health 2020-2025*. World Health Organization, Geneva.
- [78] Til Wykes, Jessica Lipshitz, and Stephen Schueller. 2019. Towards the design of ethical standards related to digital mental health and all its applications. *Curr Treat Options Psychiatry* 6 (July 2019), 232–242. <https://doi.org/10.1007/s40501-019-00180-0>
- [79] Til Wykes and Stephen Schueller. 2019. Why reviewing apps is not enough: Transparency for Trust (T4T) principles for responsible health app marketplaces. *J Med Internet Res* 21, 5 (May 2019), e12390. <https://doi.org/10.2196/12390>
- [80] Renwen Zhang, Jennifer Nicholas, Ashley Knapp, Andrea Graham, Elizabeth Gray, Mary Kwasny, Madhu Reddy, and David Mohr. 2019. Clinically meaningful use of mental health apps and its effects on depression: mixed methods study. *J Med Internet Res* 21, 12 (December 2019), e15644. <https://doi.org/10.2196/15644>

A APPENDICES

A.1 Purposive sample of most rated apps for depression included in analysis of user reviews

App name (Developer)	App store rating (/5)		No. of downloads (Google Play only)	No. of user ratings	
	Google Play	Apple App Store		Google Play	Apple App Store
Headspace (Headspace, Inc.)	4.6	4.8	1000000+	85394	160019
Doctor on Demand (Doctor on Demand, Inc.)	4.7	4.1	1000000+	20593	13
Simple Habit Meditation (Simple Habit, Inc.)	4.7	-	500000+	15601	-
TalkLife (TalkLife)	4.5	4.5	100000+	15498	625
Medication Reminder & Pill Tracker (MyTherapy)	4.6	-	500000+	14701	-
7 Cups (7 Cups of Tea)	4.2	4.4	500000+	14431	666
Abide: Christian Meditation & Prayers (Carpenter Code Inc.)	4.7	-	500000+	13971	-
Your.MD: Symptom Checker (Your: MD)	4.3	4.7	1000000+	12711	195
Pacifica: Stress & Anxiety (Pacifica Labs Inc.)	4.4	4.7	500000+	10778	486
Youper (Youper, Inc.)	4.8	4.9	100000+	10151	517
Relaxing Anti-Stress Sounds (Dandelion Soft)	4.0	-	1000000+	9935	-
Wysa (Touchkin)	4.5	3.6	100000+	9620	28
Stop, Breathe & Think (Stop, Breathe, Think)	4.3	-	1000000+	8889	-
Gentle Wakeup: Sleep & Alarm Clock (Dr Alexander Rieger)	4.4	-	100000+	6705	-

App name (Developer)	App store rating (/5)		No. of downloads (Google Play only)	No. of user ratings	
	Google Play	Apple App Store		Google Play	Apple App Store
Moodpath: Depression & Anxiety Test (Moodpath)	4.6	4.6	100000+	6680	2019
Secret Diary (Zheko)	3.8	-	1000000+	6601	-
Aware: Meditation & Mindfulness (zoojoo.be)	4.7	-	100000+	5157	-
SuperBetter (SuperBetter, LLC)	4.3	4.4	100000+	5036	431
Remente: Self Improvement (Remente)	4.4	4.6	100000+	4280	445
BetterHelp: Online Counselling (BetterHelp)	3.9	3.5	100000+	4082	244

A.2 Purposive sample of lowest rated apps for depression included in analysis of user reviews

App name (Developer)	App store rating (/5)		No. of downloads (Google Play only)	No. of user ratings	
	Google Play	Apple App Store		Google Play	Apple App Store
Depression Test (FXT Tech)	2.5	-	1000+	8	-
DSM-5 Differential Diagnosis (Unbound Medicine, Inc.)	2.7	-	10000+	122	-
Depression Support (MyHealth Teams)	2.9	-	1000+	23	-
MoodHacker (ORCAS)	2.9	1.0	1000+	16	1
My Possible Self (My Possible Self Ltd)	3.0	4.4	5000+	23	15
Social Force (IntelliCare) (CBITs)	3.0	-	1000+	14	-
Anxiety & Depression Symptoms (Twayesh Projects)	3.1	-	100000+	307	-
Talkspace Counselling & Therapy (Talkspace)	3.4	3.8	100000+	1639	71
WellMind (Blue Step Solutions)	3.4	3.3	10000+	82	24
IntelliCare Hub (CBITs)	3.4	Not listed	10000+	28	Not listed
Worry Knot (Intellicare) (CBITs)	3.4	Not listed	10000+	26	Not listed
UpLift for Depression (UpLift)	3.4	-	1000+	9	-
MHF (Together for Change)	3.5	-	5000+	40	-
WellTrack: Interactive Self-Help Therapy (CyberPsync)	3.5	-	5000+	31	-
Depression Test (Japps Medical)	3.5	-	100000+	1385	-

App name (Developer)	App store rating (/5)		No. of downloads (Google Play only)	No. of user ratings	
	Google Play	Apple App Store		Google Play	Apple App Store
CogniFit Brain Fitness (CogniFit Inc)	3.6	-	50000+	593	-
Depression Self-Help Guide:CBT (Xandy App Ideas)	3.6	-	5000+	8	-
Depression Test (MoodTools)	3.7	Not listed	50000+	199	Not listed
Aware (diagnose yourself) (Heretic Hammer)	3.7	-	5000+	23	-
Slumber Time (Intellicare) (CBITs)	3.7	-	5000+	14	-