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

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 effects, 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.

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 [73]. Among these common mental health problems, depression, defined by persistent negative mood and diminished positive affect [95], 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) [5]. 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 [124]. 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 [13, 52, 117].

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, health care systems, and society [89, 123]. 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 [53, 59, 68, 97]. These concerns are particularly relevant for the subfield of mobile mental health, specifically publicly available 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 [46, 59, 60].

While work in this space has started to emerge focusing on evaluating the content [19, 20, 46, 69, 88], functions [58, 86, 104], and quality [59, 110, 123, 125] 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' perceptions and 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 aspects 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 mental health and 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 also 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 [23, 44] and social media behaviors [7, 36] to detecting or predicting depressive symptomatology from social media data [64, 96, 119].

Efforts to develop more objective methods of detecting depression include multimodal systems [17, 37] integrating for instance audio with facial or body data [122]. HCI work on the detection of depression also includes the use of commonly available technologies such as smartphone sensors [10, 74, 120] 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 [56], leveraging memory technologies [87], game interventions [34], virtual reality [12], social robots [92], or chatbots [43]. Most commonly, technological interventions for depression are delivered online or via apps [94, 121, 128]. 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 health care systems [16] or in blended care [103].

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 [97]. 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 [51, 81, 100].

HCI work in depression has also sought to improve digital intervention by exploring factors impacting engagement and adoption [35, 66] and improved clinical outcomes [27]. Factors impacting use and effectiveness of mental health technologies and self-management include (but are not limited to) appropriate client support [27, 91], managing expectations [49], provision of social support [61], and designing for flexibility in use [35] 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 [9, 38] and systematic reviews [89, 111] 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 [46, 69, 58, 59], screening and assessment [46, 58, 86, 102], symptom management [46, 88, 102], interactive interventions [58], and tracking of moods, thoughts, or behaviors [38, 69, 88]. Publicly available app interventions for depression may be based on existing evidence-based treatments such as cognitive behavioral therapy (CBT) [38, 111], and behavioral activation and dialectical behavior therapy [111]. However, few of these apps demonstrated high fidelity to the adopted treatment approach [20, 104], 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 [29-31], hypnosis, and sound or music therapy [20].

Publicly available app interventions for depression also appear to adopt innovative uses of traditional treatment approaches and strategies, with frequent use of integrative, multi-theoretical, or transdiagnostic approaches [20]. These innovations reflect the eclectic delivery of clinical interventions in the wild [63] but unfortunately often lack research evidence supporting apps' specific design and use in treating depression [20]. 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 [14, 93]. 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 [28]. There is therefore need for more evidence on the use of apps for depression in the wild to better understand user experiences and implications for design to improve the quality and acceptance of these apps.

2.3 Perspectives in user reviews of mental health apps

App store user reviews provide concise expressions of users' perspectives and experiences of the app reviewed. In the case of apps for depression, user reviews can provide unique insights into app use, benefits, and challenges. Previous studies on user reviews in mobile mental health have analyzed reviews of general mental health apps [2, 3, 84], mood tracking apps [25], chatbot app for social support [105], CBT apps [108, 109], apps for bipolar disorder [77], and CBT apps for depression [104]. To the best of our knowledge, only one study has focused specifically on apps for depression [104]. These studies mainly aimed to explore user expectations [2, 77, 84, 104], needs [2, 77, 84], and experiences [2, 25, 77, 84, 104]. User review studies also explored specific app features such as social support [105], persuasive design elements [109], or usability [3]. A few of these studies analyzed how apps are used [77] and their main features related to adoption [108]. User review studies varied in the number of apps included (ranging from one [105] to 106 [2, 3]) and the number of reviews analyzed (ranging from 1,000 [25] to 88,125 [84], with a median of 1,287 [104]).

Users' reasons for using apps varied across the range of app types reviewed and included use for tracking mood and monitoring symptoms [25, 77], self-reflection and insight [25, 77], managing mental health conditions [25], supporting in-person treatment [25, 77, 104], and as a replacement to real-world care or support [104]. User reviews described both positive and negative user experiences, but were predominantly positive [2, 84, 104, 108]. Previous findings indicated that the most significant facilitator of positive user experiences was usability [2, 3, 77, 84, 108, 109], particularly app stability and ease of use. Other features associated with positive user experiences included monitoring and tracking [77, 84, 104, 108, 109], meditation and relaxation [104, 108], in-app communities [2, 77, 108, 109] or support [84, 105], goal setting [108], reminders or notifications [84, 108, 109], information resources [2, 108], personalization and customization [2, 104], and persuasive design features [109].

For CBT apps for depression, users appreciated both CBT and non-CBT features, such as tracking and meditation, but some users voiced the importance of evidence-based content [104]. Across all apps, the scientific or evidence-base was infrequently discussed in user reviews [2, 77, 84, 104, 109]. Discussion of specific features of intervention design was most detailed in studies that focused on a specific type of app [25, 77, 104]. For example, Nicholas et al. [77] highlighted specific intervention needs of people with bipolar disorder, such as the need for additional moods or symptoms to be tracked in the app. Similarly, Stawarz et al. [104] captured several aspects of CBT apps for depression that were important to users, including the ability to share data with their therapist and the importance of apps to focus also on positive aspects instead of merely on negative thoughts or feelings (i.e., lack of positivity).

The main negative user experiences of mental health apps were often related to poor usability [2, 3, 77, 84, 108], with users reporting technical issues and bugs [2, 3, 84, 108, 109]; interface design issues such as poor layout, readability, or navigation [3, 84, 108]; limited guidance on use [2, 3, 77]; data loss [2]; and negative effects on device battery or memory usage [2, 3, 77, 84]. Other factors which negatively impacted user experience included concerns with app content, such as poor quality or missing content [2, 77, 84], issues with privacy and security [2, 77, 84, 104, 108], poor customer service [2, 84, 108, 109], and issues with costs or billing [2, 77, 84]. Concerns with app privacy and security focused on issues with data storage and security [77, 84, 104], app permissions [2, 77],

passwords [77, 104], privacy policies [2, 84], and data sharing with third parties [77, 84]. Some apps offered better privacy in paid compared to free versions [77], which many users believed to be unethical as they considered privacy to be a right that should not have to be bought.

Only two user review studies of mental health apps explicitly mentioned ethics [77, 84]. Discussion of ethics was limited and focused mostly on privacy [77, 84], and excessive in-app advertisements [84]. This is surprising considering the importance of ethics in mental health care to ensure the safe and effective delivery of interventions. Moreover, it is surprising that much of the findings from user review studies of mental health apps focused on the usability and functioning of apps, with less discussion of elements of mental health care such as intervention design, clinical benefits, and adverse effects. Although there has been greater discussion of the ethics of mobile mental health in the literature, to the best of our knowledge, there are no studies that have used user reviews to discuss how these ethical issues may be experienced by users of publicly available mental health apps in the wild. More specifically, there is a gap in user review studies of apps for depression outside of CBT, and the ethical experiences that may be unique to users of these apps. Given the high prevalence and burden of disease associated with depression [73], we believe there is a need for more studies to inform the ethical design of apps for depression in order to increase people's access to safe and effective app interventions.

2.4 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 [6, 22]. Related to this, a rather distinct body of work has focused on extensive discussion of the ethics [50, 53, 86] and evaluation [4] of mobile mental health including apps for depression. Ethical issues commonly discussed include concerns regarding privacy and data security [39, 45, 50, 53, 57, 68, 79, 97, 110, 113, 114, 125-127], particularly related to how users' data is protected, how matters related to privacy and data security are communicated to users, and the use of users' data without their informed consent. This is related to ethical issues in areas of risks and safety [50, 53, 57, 68, 71, 85, 97, 110, 112, 114, 125, 126], and benefits and evidence [45, 50, 53, 57, 68, 70, 71, 85, 97, 110, 112, 114, 125-127]. The lack of evidence on the use, effectiveness, and adverse effects of most publicly available apps for depression [19, 20, 104] increases the risks to users of these technologies. Moreover, it demonstrates a lack of transparency in the intervention being delivered and its effects [19, 20, 57, 125-127], which may negatively impact users' trust [19, 20, 57, 113, 125, 126] and their ability to give truly informed consent [19, 50, 68, 79, 97, 113, 114, 125, 126]. Yet, these issues were largely absent in the reviews of user reviews of apps for depression or more general mental health apps. Other ethical issues emerging from the literature but not evident in user review studies included the importance of user involvement in the ethical development of mental health technologies [110, 125-127], respect for human rights and diversity [97, 127], and challenges with standards and regulation [53, 68, 113, 125].

Few researchers have framed their discussions of mobile mental health within existing ethical frameworks. Jones and Moffitt [50] and Karcher and Presser [53] referenced the professional ethical principles of the American Psychological Association [6] to provide guidance for app development and the use of mobile health in clinical practice, respectively. These codes reflect principlism or principle-based ethics [90], which is an approach to applied ethics involving the use of moral principles to analyze concrete cases and issues. Although evident in professional ethical codes [6, 22], this approach is best known by the principles of biomedical ethics [15], which are widely used to discuss ethical issues in clinical medicine. There are four principles of biomedical ethics: (1) respect

for autonomy, i.e., respect for the decision-making capacity of autonomous people, (2) beneficence, i.e., providing benefits and balancing risks, (3) nonmaleficence, i.e., avoiding harm, and (4) justice, i.e., fairness in distribution of benefits and risks for all people. These principles are used to assist with ethical decision-making and to resolve ethical conflicts.

Sanches et al. [97] used bioethics [15] as a lens to present their review of the ethics of HCI research on affective health, with a focus on how the principles were reflected in HCI research and the design of affective health technologies. Their findings showed that most papers on affective health in HCI research did not discuss any ethical concerns. Of those that did, most discussed matters related to autonomy, specifically respect for the self-determination of people with affective disorders, and their data privacy. Nonmaleficence was also evident and related to the involvement of people with affective disorders in research, diagnostic claim, and providing feedback on negative affective states. With respect to the latter, the authors found possible risks associated with monitoring negative experiences as is done in some apps for mood tracking. Fewer papers reviewed discussed benefits or justice. Bowie-DaBreo et al. [19, 20] also adapted ethical frameworks in their review of ethical issues within app store descriptions of apps for depression. Their findings highlighted the relevance of principles of beneficence, nonmaleficence, responsibility, integrity, autonomy, and justice [19]. In addition to previously discussed findings related to the lack of evidence of claimed benefits, apps for depression were also found to be poorly aligned with existing clinical guidelines [20]. This was associated with nonmaleficence, with concerns related to the limited guidance and disclaimers in app store descriptions to guide safe and effective selection and use of apps for depression [19]. Issues of responsibility and integrity included inadequate involvement of multisector expertise in app development, and limited reporting of apps' regulatory status, sources of funding, and business models. These challenges had potential impact on users' autonomy in making treatment decisions and fair access to care. Based on these findings, Bowie-DaBreo et al. [19] advocated for the application of these ideals using a responsible innovation approach [83], which encourages a process of anticipation, reflection, inclusive deliberation, and responsiveness in the design and development of new technologies [118].

These reviews show the utility of principlism in framing ethical evaluations of apps for depression. 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) [107]. Yet, although 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 [47]. This is relevant for HCI work in depression, as researchers and designers can often feel limited when discussions of ethics and associated guidelines arise. The present study therefore aims to shed light on the users' voice and experiences of ethical issues pertaining to apps for depression, and to use this perspective to shape actionable guidance for the design and development of ethical apps for depression. We approached ethics in the broadest sense, as relating to individual and social good and universal standards of right and wrong [107]. 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. [102] and Stawarz et al. [104]. 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). Of these apps, 89% (316/353) explicitly mentioned use for depression in their app store descriptions. The remaining were marketed for mood management (4.8%, 17/353), mental health problems implicitly including depression (2.5%, 9/353), low mood/mood disorder (2.0%, 7/353), distress (0.8%, 3/353), and negative thoughts (0.3%, 1/353).

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 a wide breadth of user reviews and ethical experiences, we sampled 40 apps for depression that were rated by many users but varied in their users' ratings. Users' ratings are scores (1 to 5) given to an app by users in the app stores, with 1 being the lowest rating. We chose not to sample apps based on high user ratings as positive user ratings may not accurately reflect user experiences [112]. We ranked all 353 apps according first to the numbers of users rating them, and second according to users' ratings. Firstly, we identified the 20 most rated apps for depression, or those with the highest number of user ratings across the app stores. We considered that apps with the most user ratings would likely also have high users' ratings, and thus may not fully reflect the range of ethical issues. To address this, we also identified the 20 lowest rated apps for depression which were downloaded by at least 1,000 users. Apps were not sampled based on treatment approach as we wished to capture ethical experiences across the range of interventions marketed for depression. By sampling 40 apps, we hoped to capture different user perspectives and ethical experiences across the many apps for depression in the app marketplaces. We felt a smaller sample of apps limited by treatment approach (e.g., CBT) would not truly reflect the ethical issues in this direct-to -consumer space. 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. For apps with fewer reviews, all reviews were sampled. User reviews were excluded from selection if they lacked

content, for example consisted of only 'emoji' symbols, or did not discuss user experiences, for example stated only what the app did ("This is a diary"). We did not specifically search for ethical-related posts because we consider ethics to be intrinsic in mental health care. Therefore, we expected users' reviews of their experiences of using apps for depression to include some discussion of individual or social good, harms, or other ethical issues. 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 [21]. After sampling, 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. 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. Coding, thematic development, and mapping were done 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. We considered it important to include quotes from the user reviews to accurately capture the user's voice. But we have kept quotes as short as possible and extracted only the parts that are relevant to the theme being discussed. Although user reviews are in the public domain, we have taken care to not include any information that could be used to identify the user or locate their review online. For this reason, we do not name the apps being reviewed in the quotes provided. This research received institutional ethics approval.
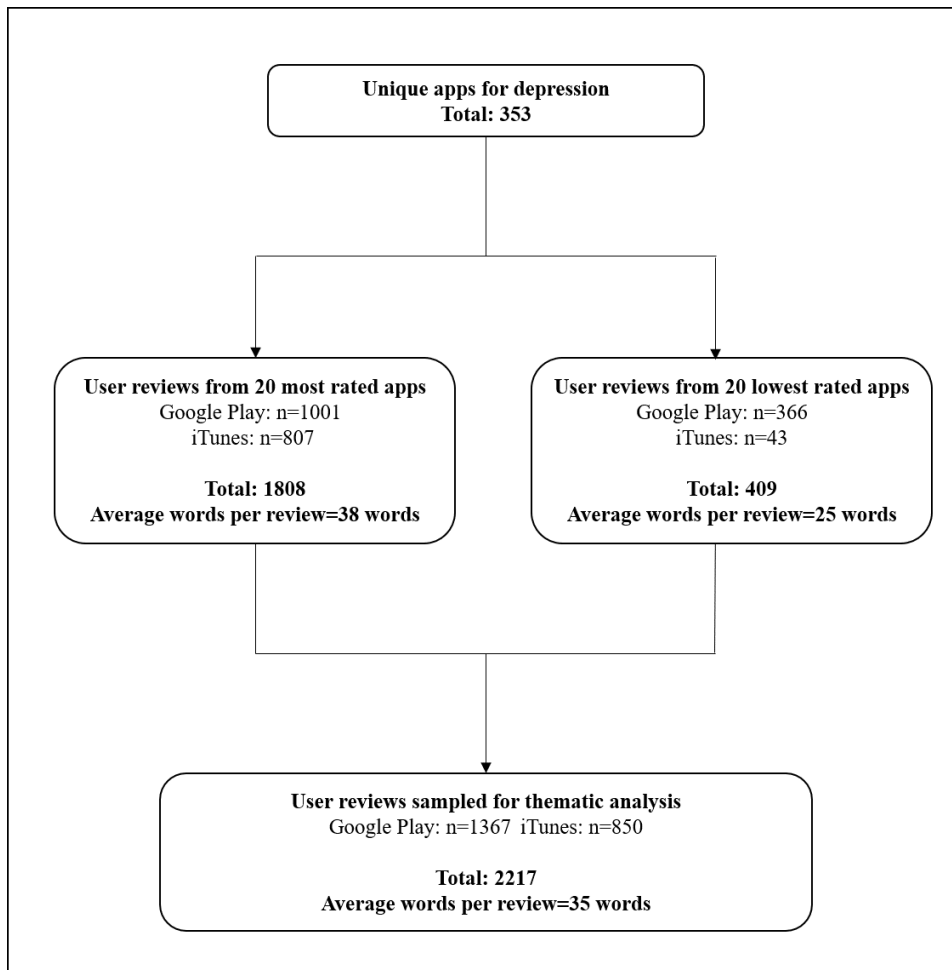
Figure 1: Sampling flowchart for user reviews

4 FINDINGS

4.1 Summary of sample of apps for depression

The sampled apps included a range of interventions representative of the different types of apps for depression available in the app marketplace. Based on their app store descriptions, approximately a third of the apps (35%, 14/40) provided a multi-theoretical intervention incorporating more than one treatment approach. The most common interventions were assessment (25%, 10/40), CBT (20%, 8/40), mindfulness or meditation (20%, 8/40), psychosocial interventions (20%, 8/40), and psychoeducation (18%, 7/40) (Appendix A.3).

The 20 most rated apps for depression ranged in number of ratings from 4,082 to 85,394 (n=20; median 10,043) in Google Play and 13 to 160,019 (n=12; median 465.5) in Apple App Store. The mean users' rating for these apps in Google Play was 4.4 (n=20; SD 0.3; range 3.8-4.8), with apps rated similarly in Apple App Store (n=12; mean 4.4,

SD 0.4; range 3.5-4.9). The most common intervention in this group of apps was mindfulness or meditation (40%, 8/20), followed by CBT (20%, 4/20) and psychosocial approaches (20%, 4/20) (Appendix A.3).

For the lowest rated apps, the number of ratings ranged from 8 to 1,639 (n=20; median 27) in Google Play and 1 to 71 (n=4; median 19.5) in Apple App Store. The mean users' rating was 3.3 (n=20; SD 0.4; range 2.5-3.7) in Google Play and 3.1 (n=4; SD 1.5; range 1.0-3.8) in Apple App Store. The most common interventions in the lowest rated apps were assessment (40%, 8/20), followed by psychoeducation (25%, 5/20), and CBT (20%, 4/20) and psychosocial approaches (20%, 4/20).

4.2 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. This is not surprising given that there were more user reviews written for the most rated apps compared with the lowest rated apps, with the former also having higher users' ratings. In our sample, the apps with higher users' ratings also tended to have more positive reviews (Appendix A.3). Interestingly however, all apps bar one had negative reviews with 85% of apps (34/40) having a range of positive, negative, and ambivalent reviews. 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).

4.3 Benefits and harms of apps for depression

4.3.1 Benefits of apps for depression

Benefits of apps for depression describe the positive effects of apps on users' mental health, wellbeing, and development. Almost half of all reviews (42%, 921/2217) across 36 of the 40 sampled apps described at least one benefit of apps for depression. In particular, apps helped users to manage their mental health problems (including depression, anxiety, stress, and bipolar disorder) at different stages of their disorders, from prevention through to treatment, recovery, and relapse prevention. Apps also helped with managing difficult situations, providing support or connections to services during low moments. Users often expressed feeling that they had no one else to turn to and sought comfort in the app to help them cope with challenges or moments of distress:

> "[It] 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."

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

> "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!"

Emotion regulation was also supported through encouragement for simple activities and small steps to feeling better, i.e., behavioral activation. This included goal setting and the creation of new habits and routines:

"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."

More generally, apps helped users with their overall wellbeing, including personal development. Users described apps as having a positive impact on their balance and focus, gratitude, motivation, engagement, and openness. For some, this 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."

4.3.2 Risks, adverse effects, and safety

An important finding is the reporting of risks, adverse effects and safety concerns of apps for depression (10%, 220/2217). We define risks as anything which could potentially harm users, while adverse effects are negative outcomes experienced by users from using the app. Risks associated with apps for depression included the provision of harmful advice and the potential for misdiagnosis, failure or errors in delivering important information, or the risk of users' overreliance on the app. With respect to misdiagnosis, reviews pointed out the mismatch between the tracked data or app's prediction of their experience and their actual experience:

"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 *[which is issued by the app for the user to share with their healthcare professional]* says I haven't been having problems with sleep when I certainly do."

App errors posed additional risks in terms of delivering 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 lead users to become over reliant on them:

"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.

For some, using apps led to adverse effects such as feeling worse after use, harassment, or mistreatment. Some users reported discomfort with completing aspects of the intervention, stress because of difficulties accessing the intervention or poor app or intervention design, and poor quality of support. For example, negative mood tracking can become problematic as it can increase user awareness of it and feelings of self-deprecation:

"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."

Other negative outcomes of using apps for depression were experiences of bullying, harassment, or mistreatment by in-app or peer 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?"

Alleged harassment or mistreatment of users included unwanted sexual advances, rude or judgemental comments, and trolling behaviour from in-app or peer support:

"I needed help and got nothing but creepy guys who pressured me for snap chat adds and kept asking for virtual hugs...Avoid this especially if like me you've lost faith in people..."

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.

Risks and adverse effects 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. A key finding is the limited number of reviews that described apps for depression as providing a safe space or being designed with safety in mind. These concerns relate mostly to proper monitoring and moderation of peer support to prevent misuse by users, insufficient communication of apps' limitations, and insufficient support for safety and safeguarding.

Regarding misuse, several reviews expressed concerns that apps for depression may fail to provide a safe place for all their 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."

As this quote illustrates, misuse can be arguably linked to users' limited psychoeducation, and more importantly to apps' limited moderation and regulation.

Reviewers also highlighted the importance of being aware of app limitations in providing adequate support or access to therapist whenever needed. This was demonstrated in user reviews where reviewers advised others to seek in-person care or to contact emergency services when in crisis. In one instance, this was reiterated in a developer's reply to a review:

"[This app] is NOT a replacement for a real therapist, if professional help is required. [It] is NOT a crisis tool. Do seek professional help and create an emergency plan if you ever have such thoughts again."

This example demonstrates some attempt by developers to safeguard users, though more direct referral to appropriate mental health or crisis services is warranted for better and more responsible safeguarding. It also highlights the need for monitoring and safeguarding of app store user reviews, with safety extending beyond apps and their content. As our sample of user reviews demonstrates, users post very personal and at times sensitive and triggering content in app store reviews:

> "For the legit kids who aren't sad over one thing. If you're like me and you never feel hunger or happy just sad and you want to cut yourself or harm yourself everyday and sometimes are so tempted you just sleep then yes you have depression"

Yet only 63% (25/40) of the sampled apps had at least one developer reply to user reviews in the app stores. It is unclear whether users who posted reviews containing safety concerns were contacted privately by developers or app stores. There was also a call from a small number of users for app stores to take more responsibility to regulate or ban apps that were deemed to be exploitative or unsafe.

Users also demanded greater support for safety and safeguarding, in particular protection 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 to real-world treatment when needed and explanations for why some users were refused care. The latter may reflect in-app attempts to safeguard users who were not suited for 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."

4.4 Facilitators of benefits and harms

4.4.1 Usability

Benefits, avoidance of harms and risks, and safety are the foundation of ethical apps for depression. These areas are also indirectly affected by apps' usability, design, and support. Usability was the most common theme emerging from user reviews (43%, 940/2217), with users reporting both positive and negative experiences. We consider usability to be how well an app provided its intervention to ensure users were able to use it effectively and safely. In this context, poor usability which affected access to support or accurate data collection and reporting has 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. Usability concerns include those of app stability, ease of use, device functionality, and app customization.

Most reviews on app stability highlighted issues with technical difficulties, app failures, and poor responsiveness affecting the use of the app. Issues with app stability occurred at early stages of signing up, logging in, as well as crashing throughout use. For some apps, multiple user reviews reported the same technical difficulties without reply from the developer. Other issues included lag in the running of the app, frequent crashing or freezing, or challenges with specific elements of apps not functioning as expected, such as sounds not working in meditation recordings, notifications not appearing, and in-app communications failing to be sent or received.

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."

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 and hence less easy to use and with limited user support to navigate such complexity.

Discussions on usability also pertained to device functionality. This includes how the app functioned on specific mobile devices, such as smartphones versus tablets, as well as how apps affected a device's functioning, with some apps interrupting phone calls, sounds and volume settings, or battery life and memory. 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 tracked data or in-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, for example being unable to change previously selected intervention goals. 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.

4.4.2 Design of intervention and design of app interface

Like usability, design (30%, 657/2217) had indirect ethical implications related to how the design of the therapeutic intervention delivered through the app and the design of the app interface affected intervention validity, app safety, or accuracy of data.

Positive reviews on intervention design related to the intervention approach and the content provided. But negative experiences stemmed from users' disagreement with the intervention content, including beliefs that the intervention was too generic and lacked depth:

> "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."

Some users believed in the validity of the apps, but others expressed concerns with the theoretical orientation or evidence base: "As they say it's not professional...I don't think questions/results are based on some concrete research." Other reviews reflected users' critique of the assessment tools aimed to predict their level of depression:

"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."

Related to this, there were also concerns with using apps for self-diagnosis, with a few users highlighting the potential for invalid results due to false reporting by users: "You will get whatever answer your looking for. Not an accurate method of diagnosis".

Some user reviews highlighted areas of app design and particularly its interface. These were mostly positive, reflecting enhanced user experiences and treatment delivery. These consist of several persuasive design features [80], such as notifications and reminders, tailored interventions, tunnelling using stepped tasks, rewards and gamification, or self-monitoring.

Some users also expressed a desire for more persuasive app design features. Others wished for more meaningful data, ranging from data on app usage, before-after data to monitor the effects of intervention activities on their emotional states, more open-ended data entry for better self-expression and accuracy, and improved data storage and long-term data collection:

"One thing that maybe could be improved is extending the mood average period on the screen where it shows average mood over 7, 14, and 21 days to maybe show over a longer period. I have noticed that I have several really low days followed by 2 moderate days and it would be really good to see if this trend has persisted over a longer time frame (i.e. monthly, 3 monthly and 6 monthly)."

4.4.3 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 also 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:

"I love the idea of this game! Installed it on my phone a couple of years ago, but it kept crashing. Trying it again. Hope it works better this time. Update: Ran into my first problem. None of the links in the app work. Disappointing. Update: Tried to email the developers. Message wouldn't go through, so I suspect they may have abandoned the app. Super disappointing."

Users also equally discussed the importance of therapeutic support. Therapeutic support in apps for depression consisted of three types: (1) online therapy with qualified counsellors, (2) peer support, and (3) in-app support provided by chatbots. Several users were satisfied with the therapeutic support received but others noted

shortcomings and additional support needs, including the need for human support, more frequent and consistent delivery of support, and provision of support at specific points in the intervention such as after self-assessment or when dealing with adversity:

> "I dont recommend. People download this app with the hopes that someone (a live person) will chat with them and help them deal with their problems, but literally no one is ever available and if anything, it kind of made me feel even worse. It felt like they were too busy to talk and I found no one fit enough to help me cope with my depression. The bot was honestly more helpful than the people. I suggest finding a hotline."

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 also concerns with chatbot communications. The most notable issue was the inappropriateness of chatbot responses which users thought showed a lack of understanding and insufficient technological capacity to provide adequate support:

> "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:

> "I had spent many years suffering with some sporadic bouts of depression…I had some trepidation, but my counsellor felt like a perfect fit straight away and I have now been speaking to her on average once a week for the last year. Knowing I have her support, it has helped guide me through the times where I used to feel real darkness and emptiness. When I don't have a session scheduled, I can still message her through the message facility and she has been an absolute professional throughout."

While social support could be also problematic, as discussed in Section 4.3.2, 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. Users described receiving help from 'like-minded' persons and how this helped them to feel less alone. Users also benefitted from providing support to their peers, describing feelings of empowerment, empathy, and improved social relations. 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 health care provider. Users described previously having difficulty getting their primary care provider to understand their challenges, and felt that the app made this easier by legitimizing their mental health concerns or providing a summary report of their data which they could share with their doctor:

"The referral letter *[provided by the app]* 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."

4.5 Justice and rights

4.5.1 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 user's voice.

With respect to app choice, the depression app marketplace allows potential users to take an active role in
researching and selecting apps and their 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. But less transparent
costing affects this:

"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 provide users with greater choice regarding the type of interventions available and facilitate their
engagement in treatment planning and decisions, 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."

In addition to treatment choices, apps also offered users in-app options for customization (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 apps 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 by providing explicit

recommendations for use, or recommendations for alternative treatments. In this manner, user reviews functioned as a community of peer referrals and support.

4.5.2 Access

Apps for depression provided many users with increased access to care with almost 30% of reviews (645/2217) discussing some aspect of this theme in their review. Three main concepts were discussed, namely the barriers to real-world care, preference of apps over real-world care, and barriers to apps for depression.

User reviews showed that apps for depression have the potential 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 the difficulties of physically accessing in–person care, often due to work or childcare demands, or treatment costs. 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: "If you're really in need and looking for a chat with a therapist, then be prepared that one might not always be available. It is a shame."

Access to care was also negatively impacted by disruptions to the continuity of care. Some users reported unexpected disruptions to app interventions due to technical difficulties, app updates, unnotified termination of in–app support, and increased app costs. While some users found apps to be more affordable than other treatment options, many reviews highlighted app costs or paywalls as barriers to care:

> "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."

4.5.3 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 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 health care 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 health care 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, but now, not so much."

Apps business models and their transparent 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. However, a small number of users called on the app stores to take more responsibility to regulate or ban apps that are perceived as exploitative and deceitful.

> "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."

4.5.4 Privacy

A significant finding is that less than 5% of user reviews (4.8%, 107/2217) mentioned privacy. In the context of apps for depression, privacy pertains to respect and protection of users' information, including personal details, identifiable 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 also appreciated anonymity which they believed helped them be more open in expressing themselves and seeking help, while also making them feel safe:

> "How I describe this app is like a Facebook for people that need help. When you go to post there is also a way you can post anonymously so you feel safer."

However, almost 80% of user reviews discussing privacy highlighted concerns ranging from a need for greater secrecy or anonymity, concerns with the collection of sensitive user data before confirming access to the intervention, and concerns with data security and the sharing of user data with third parties. While some users found app privacy policies to be accessible and easy to understand, they did not always agree with the practices outlined in relation to their use of personal data:

> "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 users' desire 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."

4.5.5 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 apps' 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."

4.6 Virtue of apps for depression

4.6.1 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 under insufficient key information and reduced trust in apps or developers. 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."

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, perceived intentions, and ethical and legal compliance. Fidelity is related to truthfulness, i.e., the app does as it says it would. This involved 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."

Less surprising, 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."

4.6.2 Social impact

A minor theme in user reviews was the wider social impact or social good 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 health care:

"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 perspectives and ethical experiences, and key elements of ethically designed apps for depression. This study provides a novel contribution to the literature on the ethics of apps for depression in particular, and mobile mental health more broadly. 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 design implications are explored in greater detail in the following section.

The sample of user reviews of apps for depression was largely positive, with less than half of all reviews categorized as negative or ambivalent. This may reflect our sampling method of selecting reviews from the most rated apps, followed by the lowest rated apps. Despite this, even highly rated apps had some negative reviews and discussion of ethical issues. We were therefore successful in sampling a diverse range of user reviews to convey the breadth of experiences and potential ethical issues. 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 echoing some of previous findings [2, 84]. Ease of use and good product usability were key to positive user experiences, with apps for depression being 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, an outcome also confirming previous ones [84].

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 risks and adverse effects. User perspectives and ethical 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 on user values in apps for depression and broader mobile mental health, and how these impact user expectations, use, and adoption.

Our study extended the findings of the one previous user review study on apps for depression [104]. Stawarz et al. [104] however, limited their review to user experiences of CBT apps for depression. But 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. This is important as although CBT is considered the gold-standard in psychotherapy for depression, it is not the only digital intervention accessed and used by people with depression [20]. Moreover, although Stawarz et al. [104] mention some ethical concepts, such as privacy and trust, they do not discuss their ethical implications. Our study not only explicitly focuses on the ethical experiences and issues of apps for depression, but shows how traditionally non-ethical concepts, such as usability and design, also have great impact on benefits, harms, and safety. Apps for depression are by nature apps designed and developed to be used by people with mental health difficulties who may be more vulnerable and have greater needs for guidance and support. We therefore consider our study's focus on the ethics of these apps to be of great importance and to fill a glaring gap in the existing evidence.

5.2 Implications for the design of ethical apps for depression

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 [47]. 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. In this section, we reflect on our key findings and how they can inform the design of apps for depression. We contrast our implications for design of such apps with design recommendations

or guidelines previously suggested and articulate how ours extend them in new ways. We organize these design implications under the four principles of biomedical ethics, starting with the most challenging one: nonmaleficence, followed by beneficence, justice, and respect for autonomy. We have chosen to frame this discussion using principlism [15] given its relevance to mental health care and mobile mental health. However, we use the user reviews to make these concepts more accessible and to present practical considerations for the design of ethical apps for depression. In addition, also discuss the ethical concept of virtue [107] which resonated in user reviews but is not strongly highlighted in traditional discussions on biomedical ethics [15]. Within each principle, the most novel findings and their implications are unpacked first.

5.2.1 Nonmaleficence: problematic peer support, in-app diagnosis, tracking negative moods, and problematic usability

Nonmaleficence is the ethical concept of avoiding harm [15]. A key outcome concerning this principle is that our findings provide a more cautious view on the benefit of social support compared to previous studies on users reviews of mental health apps in general [108, 109], and bipolar disorder apps [77] or chatbots [105], particularly with regard to peer support in users' online communities [77]. While some apps for depression leveraging such communities were perceived as beneficial for strengthening users' social connections and sense of community, for other apps, the peer support was perceived as not only limited but particularly harmful leading to users feeling ignored when not supported promptly, feeling worse when listening to other's mental health challenges, feeling judged when sharing their own experiences, or even bullied or targeted through inappropriate sexual comments. These are particularly important findings, especially since previous research has provided limited evidence of the harm done by mental health apps [71]. They therefore require strong design interventions. Here we suggest the value of apps for depression which involve peer support to provide ground rules for sharing experiences, and psychoeducation on how to adequately respond to other users' negative shared experiences, together with the option of training users to facilitate emotional co-regulation among peers. We extend previous suggestions for trained moderators to ensure confidentiality of online mental health interventions [62], with design implications for depression apps to moderate peer support through both trained (or training) facilitators, and machine learning and natural language processing algorithms which have been already explored for the detection of online antisocial behavior [18, 78].

Also related to adverse effects is our finding that some apps for depression led to automatic misdiagnosis based on user generated data. This is a novel outcome for apps for depression, although previous findings have highlighted the ethical concerns of how, and to whom diagnostic claims are communicated more broadly through affective technologies [97], as they can perpetuate stigma, discrimination, and worsening of mental health difficulties. Hence design suggestions to address these concerns could include those previously suggested ones for affective health technologies such avoidance of diagnostic labels, or explicit advice within the app for users to seek support and interpretation of the tracked data with a mental health professional [32, 33, 103]. But this is not sufficient, as seen in our finding that app generated reports intended to be shared with a health care professional contained errors in the data recorded and the summary provided. Apps with user generated data should be designed with multiple opportunities for data entry as well as the checking and correcting of errors in data. In-app error reporting should be clearly visible to allow users to easily flag errors in their data to developers who should provide a quick response to address these concerns. Our findings also indicate additional users' concerns regarding the validity of the diagnosis tools, or the data used to inform such diagnoses. This indicates the value of psychoeducation with respect

to the validity of any such tools, their scientific underpinning, and the population for use. Any outcomes need to be sensitively communicated as highlighted above.

Another important outcome related to the principle of nonmaleficence concerns the tracking of negative emotional content. Previous user reviews studies suggested tracking moods and thoughts, and activities for mood regulation as key features for mental health apps [108, 109], which our findings extend to user reviews of apps for depression. One distinction in our findings is the challenge experienced by some users with tracking and visualization of negative feelings and thoughts which could negatively impact or worsen users' experience, and ultimately depression symptoms. These outcomes extend to commercial apps for depression the previous findings from academic research on affective health technologies [97] and the suggestion for sensitive design of such feedback to support reflection or users' actionability [101] and agency through DIY approaches for better understanding of their data [115]. To address this concern, another design implication for apps for depression could be to provide the option of more ambiguous visualization [98] of the moods and thoughts being tracked, with emphasis on supporting self-compassion, mindfulness, and acceptance rather than the precise and detailed capture of negative patterns. This design implication also leverages findings on the value of compassion theory for depression [55] allowing users to benefit from self-compassion while engaging in a softer form of monitoring which we call *compassionate self-monitoring*. Moreover, the design of interventions for apps for depression should consider the aim of the intervention (e.g., the purpose of the user tracking their moods), the design needed to facilitate this, and the evidence-base behind this design. Further to this, apps should be researched to see if they work as intended, how they compare to standard care, and if there are any risks or adverse effects associated with their use.

We also support previous suggestions regarding the application of ethics models, frameworks for mental health apps, or reputable websites [26, 71, 79, 85] such as American Psychiatric Association (APA) App Evaluation Model [4] or ORCHA [82]. These can be followed not only by therapists recommending apps to their patients or users to inform their apps' selection, but more importantly we suggest that they should be also followed by developers to inform their design of apps for depression and by app stores to check apps' suitability before being uploaded. With respect to the latter, in the light of our findings on some of the apps' harmful impacts, it is paramount to ensuring that only apps which are safe to be used by people with depression or other mental health conditions should be uploaded to app stores, as safety is arguably even more important that apps' effectiveness [4, 71].

5.2.2 Beneficence: users' limited awareness of app business models and science- and evidence-basis, and problematic usability

Beneficence relates to both benefits and minimization of risk [15]. It is therefore closely linked to nonmaleficence. Our findings showed that beneficence is challenged by four key issues, which prevent users from reaping the full benefits of apps for depression. There is also the risk that if not addressed, the most serious of these issues can become harmful.

The first issue is users' limited awareness of the business models underpinning apps for depression and how, as they stand, these models require the monetization of users' mental health-related data. Given the prevalence of user reviews demanding free apps and their limited privacy literacy [41, 72, 79], it is likely that users of apps for depression operate under two distinct yet conflicting mental models. On the one hand, they hold the assumption that apps for depression, like many mobile apps, are affordable tools with valuable functionalities. On the other hand, they hold the assumption that apps for depression should not share their sensitive data with third parties

and thus monetize it. This is akin to the traditional health care context, regulated by strong professional standards, legal and fiduciary protection of patients' best interest and therefore their private data [41].

Our findings support the idea of users having confounding mental models, with them expecting for instance respect of their private and sensitive data on apps for depression, despite these apps operating outside the traditional health care context [41]. In fact, commercial apps for depression are geared towards revenue and are largely unregulated. Thus, they are not obliged to follow the fiduciary protection of users' best interest [85] but rather maximize their own revenue. Therefore, the trade-off that some users make, often with limited awareness, is that the cost of free apps is their private data [26, 41, 85]. For context, in 2021 the mental health apps market reached between USD500 million [11] and USD1 billion, experiencing a massive growth at an annual rate of over 20%. It is estimated to reach over USD3 billion by 2027 [UnivDatos]. This is impressive, given that most of these apps are currently free or low cost, relying on in-app purchasing of extra features (which our user reviews show dislike for) and tailored advertisements leveraging user data [Grundy, Tang] with new forms of monetizing such data through personalized and customized features also being explored [11]. This is a notable ethical issue due to the potential exploitation of people with depression who may be vulnerable and seek care [8].

To address this challenge, we encourage supporting users' understanding of apps' business models and potential trade-off of apps' cost and their monetization of user data. For this, we can think of novel designs for materializing apps' business models, to ensure transparency of what data is being collected and for what purposes. One can also think of providing in-app support for educational content on business models, or generic training provided in app stores. In addition, we also suggest that apps for depression are provided not only as "free" albeit in exchange for users' data, but also in paid versions to ensure users' choice over the ownership, security, and privacy of their data. Indeed, our findings and user reviews of apps for bipolar condition show that users are willing to pay for good apps [77]. With better understanding of the commercial side of apps for depression, users would be better equipped to benefit from their use.

We also suggest the value of better managing users' expectations with upfront and complete information on app costs. Related to this is the two main sources of app costs: the cost of the app itself and the cost of therapists' support, which appear entangled, shaping users' expectations of the apps. Our findings indicate that users' mental modes of free app use appear to extend towards also accessing free therapeutic support. Users' interest in circumventing the cost of such support is reflected in their wishes for free access to apps for depression and the therapeutic support provided. This leads to an interesting tension between developers' commodification of commercial mental health apps and users' expectation of their right to free mental health care [1, 48]. This perceived right to care is probably greater in countries with national health care systems, such as the NHS in the UK, which does not charge patients for mental health treatment. However, the growing demand for professional therapeutic support is met with a significant shortage of mental health practitioners worldwide [24]. In this context, mental health apps in general, and those for depression in particular, appear to offer their users the opportunity to fill this gap by allowing access to therapists' support. It is a model more in line with private mental health care. Professional therapists' support is a costly resource, irrespectively of being delivered face to face or through mental health apps. The challenge of apps for depression is managing users' expectations of these products and their place in the health care system, as well as the cost of the app and the cost of therapeutic support provided in-app. One way to address this challenge is the clear decoupling of these costs by providing transparent information and user education to help them more realistically manage expectations that quality private therapeutic support is not a free, although the use of the app may be.

In relation to beneficence principle, our findings also confirm previous ones on the limited scientific underpinning and evidence-based effectiveness of mental health apps [77, 108, 109] showing also that only a few users of apps for depression share such concerns. While many user reviews relayed experiencing benefits of apps for depression, the lack of evidence into these benefits makes it difficult for users to easily find the most suitable and effective apps to meet their needs. This was seen in our finding that many users shopped around to find the best apps for depression. To address this, we also argue that more work is needed from developers to address the continued shortfall in evidence-based apps for depression [104] which in turn can increase users' perception of apps' credibility and trustworthiness [104, 109]. We also suggest the value of educating users about the importance of apps' scientific underpinning and evidence-based effectiveness, by designing new tools which apply existing ethical models and frameworks to empower users make more informed choices of their apps for depression.

While previous findings also indicated usability concerns of mental health apps such as bugs [3] or technical issues [77, 108, 109], our outcomes show how these issues negatively impact benefits of apps for depression and risk potential harm. Poor usability can hinder user engagement and long-term adoption of mental health apps in general [75, 76] and those for depression in particular, and thus limit apps' potential benefits. This outcome is surprising given the ubiquity of mobile technologies and the focus on user experience in both industry and academia which assumes that usability issues have been adequately ironed out. Given however the massive growth of mental health apps including those for depression, one possible explanation is that developers may prematurely upload or update their apps on the marketplace without being fully tested and stable. To address this issue, both app developers and app stores should be responsible to ensure that apps are stable and updated regularly with stable new versions. Moreover, Torous and colleagues [112] pointed out that apps should "stand the test of time" providing continuing support. Based on our findings, we add to this the need for developers to be responsible for providing exit strategy for discontinuing apps which people with mental health conditions have used, as this may lead to loss of data and harmful consequences due to feelings of abandonment. Here we build on the previous suggested option of extending use of research prototypes deployed in the wild [99] until alternative solutions are provided, for instance of easy transfer to another available app with similar functionalities. We also encourage designers and developers to engage in responsible innovation [83, 118] to anticipate potential usability and design risks throughout the app's lifecycle and to engage in participatory design to better understand how usability and design can negatively impact users with depression. Developers should also iteratively assess risks and safety of their app and promptly correct errors and user reported concerns.

5.2.3 Justice: refusal of care due to screening and affordable apps for depression

Justice concerns fairness in how benefits and risks are distributed for all people [15]. Under this principle are ethical issues related to access, equality, and rights. An important outcome is that while most reviews appreciated therapeutic support, others expressed concerns regarding refusal of care. Refusal of care is a problematic barrier to access even if it is being informed by the need to safeguard users. This outcome extends previous ones on affective health technologies where researchers employed screening to exclude people whose mental health may be negatively impacted by taking part in research [97]. In the context of apps for depression, it would be useful that not only the apps' functionalities are described in app store listings, but also their limitations and inappropriateness for people with severe mental health conditions who should be directed towards real-world mental health services. We argue for the importance of transparent information for setting right expectations of what apps for depression do, and more importantly what they cannot do in terms of treatment details, expected outcomes, and support. User

reviews suggested that users were more understanding of limits to care when clearly explained, but they expressed strong disapproval and feelings of rejection when this was unclear. In line with traditional mental health care standards [6, 22], it is crucial to actively involve users in safeguarding decisions and practices. This is key for supporting prospective users to make more informed choices [49] and to easily access the right level and type of intervention for depression to meet their individual needs.

The issue of accessing professional therapists' support through apps for depression and the cost of such support raises interesting ethical challenges under the principle of justice. Sadly, covering the cost to access such therapeutic support may be particularly problematic for economically disadvantaged users, who arguably are at higher risk of mental health conditions such as depression. Addressing this issue may require policy makers' efforts towards covering the cost of mental health apps by health insurance providers, or collaborations with health care providers towards supporting subsidized subscriptions and affordable payment plans for recommended apps for depression. In the UK, apps for depression will need robust evidence of clinical and cost effectiveness compared with standard care to be considered and recommended for use in the NHS [20]. This raises an ethical conundrum as developers may not invest in evidence-based design and research of apps for depression, resulting in few apps being available through the NHS. This further proliferates the unregulated direct-to-consumer marketplace, where some users may struggle with selecting and paying for quality apps for depression. We therefore encourage developers of apps for depression to explore how they can make their apps more accessible to users in their respective health care systems, and to invest in evidence-based design and research to gain entry to these markets.

Another avenue to address this tension is exploring new ways in which users' data can provide revenue to the users themselves, which can then be used to purchase therapeutic support. This is grounded in work of sociology of health scholars [67] who have highlighted the unethical monetization of data of users' experience of illness whose gathering is not always automatic but requires users' labor. As prosumers (users of apps and creators of data), users contribute significantly to growing the new data economy of digital patient experience with only the indirect benefit of their apps being improved. Users receive no financial compensation for their crucial contribution which is commodified by the app developers. This opens up an interesting design space where HCI researchers can work closely with developers, users, and business studies researchers to co-design new and fairer business models towards more ethical distribution of profit generated from user data.

5.2.4 Respect for autonomy: supporting users' privacy literacy, choices, and flexible use of apps for depression

The principle of respect for autonomy acknowledges people's right to make their own decisions or to act on their own freewill without pressure or coercion [15]. Our findings confirm previous ones from user reviews of mental health apps [62, 77, 104, 109] on the value of supporting various user choices, including that of specific treatment options or customization of apps' features. Our findings indicate that balance is needed to provide sufficient options: not too few which users may feel lack flexibility, and not too many which may lead to users feeling overwhelmed. This echoes findings of a systematic review on computerized CBT indicating that high level of autonomy can be both empowering and demanding [54].

Our findings showed that increased user choice and autonomy also meant increased user responsibility. 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 [28], little is known of the potential impact of this responsibility on treatment outcomes, potential risks and harms, and the type of support needed to ensure safe and effective autonomous use of apps for depression. More evidence is also needed on how much user responsibility is appropriate in apps for depression, and where in the intervention this may be best applied. People with depression may need more support at the start of the intervention when motivation, moods, and behavioral activation may be lowest. Too much user responsibility at this stage of treatment may therefore become a barrier to care but could be used to propel treatment goals and progress if increased appropriately as the intervention progresses. Through participatory design with people with lived experience of depression, apps for depression can be created to balance self-determination with support.

This is closely related to issues of users' privacy and informed decisions around sharing their personal health data. Although the literature tells of well acknowledged privacy violations of mental health apps [79], our findings surprisingly indicate that less than 5% of reviews mentioned privacy concerns. This contrasts with previous outcomes from a smaller study of user reviews of CBT apps for depression [104], indicating a concerning lack of awareness of users of apps for depression regarding privacy implication. Besides previously suggested recommendations for clear and transparent communication of privacy policies [25, 26, 71, 79, 109], in the light of our findings, we also suggest supporting users' privacy literacy through, for instance in-app training or app store descriptions. Following previous recommendations that mental health apps should implement security and privacy mechanisms [79], we also encourage app stores to request that developers explicitly state in their app store descriptions if and what security extensions [65] are in place. We also encourage explicit informed consent practices for apps for depression as has been previously suggested for online mental health interventions [62]. Inspired from security research [65] and ethical practices in traditional mental health care [6], apps for depression should present users with a brief comprehensible privacy summary and terms of use to be reviewed and agreed upon before the app's first use. Only then can a user be said to be truly informed and autonomous in their decision-making.

5.2.5 Virtue: foundations of transparency for trust, and wider social good

Our findings showed how user experiences were impacted by more than just the app itself, with users also commenting on apps' purpose or developers' intentions. 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 is akin to the ethical concepts of virtue, which considers how a person's character or values ought to be [107].

Within this theme, we found issues related to transparency and trust. The concepts of transparency and trust were more nuanced in the user reviews than in the literature, which is focused largely on issues of privacy, security, benefits, and safety [113, 126]. 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, our findings show a disparity between users' concerns in this area and the focus on privacy in the literature [45, 68]. 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 apps for depression.

When designing apps for depression, developers must consider their target audience and the importance of truthfulness, fidelity, and trust. It is important that apps for depression do as they promise, and that all elements of apps are clearly and openly communicated to users. Transparent app design and honest communication of apps' offerings and limitations is key to fostering positive perceptions and interactions with the app and in-app support. This is akin to therapeutic alliance which is the bedrock of effective interventions for depression. All aspects of apps for depression should be freely communicated to users, from costs, billing practices, treatment approach, risks, privacy, or evidence. Apps for depression should also be designed so that not only is the app easy to use, but information about the app is easy to access and to query.

5.3 Ethical tensions in the design of apps for depression

A key finding from our analysis of user reviews was the interrelations between ethical themes, such that a shortcoming in one area often negatively impacted others. For example, a lack of app updates affected 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 also impacted their perceptions of developers' intentions and motives, leading to diminished trust in the app. 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 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. This opens up an interesting HCI research space for novel design methods grounded in ethics models that can better support designers and developers to create more ethical 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 [15, 118]. 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 [20, 86]. 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 [83] which encourages designers to use moral conflicts to inspire, rather than hinder innovation [118]. 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 [40], and be granted access accordingly. 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.

5.4 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. Our analysis included 2,217 user reviews of apps for depression. While a few user review studies have included larger samples [2, 84], our sample of user reviews was more than most studies in this area [3, 25, 77, 104, 105, 108]. As with these studies, our findings contained rich user data and a wealth of information on user perspectives and ethical experiences. We suggest additional qualitative research using focus groups or interviews to further explore ethical experiences of apps for depression and to advance our preliminary findings in this area.

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. Our sampling methods captured a wider range of negative and ambivalent reviews than reported in previous studies [104] and was successful in exploring a wide range of user perspectives and ethical experiences. However, the disparity should still be kept in mind when interpreting the valence in user reviews.

Lastly, our sample of apps for depression included apps targeting depression or depressive symptoms. We chose a broader definition of 'apps for depression' as we considered that people with depression, particularly those without a formal diagnosis, may not always search for apps using the term 'depression' but rather may seek apps to help with low mood or feelings of distress. We acknowledge that this may have resulted in apps targeting more general mental health difficulties being included. In our sample of 40 apps for depression, 38 apps explicitly mentioned use for depression. Of the remaining two apps, one targeted low mood while the other was a professional diagnostic tool for all mental health conditions including depression. We therefore believe the apps included in our analysis to reflect apps marketed for depression in app stores. Apps for depression are often transdiagnostic and may target multiple mental health conditions along with depression [14]. It is therefore possible that some user reviews for these apps may be from users who have other mental health difficulties. We consider these perspectives to be valuable in reflecting the range of experiences of people who use apps for depression, particularly considering the notable mental health comorbidities associated with depression [73]. This increases the generalizability of our findings and design recommendations to other mental health apps which may also have some of the same ethical issues. Further research is recommended to explore ethical issues specific to apps for depression.

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 apps for depression 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]   Neil Krishan Aggarwal Michael Rowe, and Michael A Sernyak. 2010. Is health care a right or a commodity? Implementing mental health reform in a recession. *Psychiatr Serv* 61, 11 (November 2010), 1144-1145. https://doi.org/10.1176/ps.2010.61.11.1144

[2]   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

[3]   Felwah Alqahtani and Rita Orji. 2019. Usability issues in mental health applications. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*. Association for Computing Machinery, New York, NY, USA, 343-348. https://doi.org/10.1145/3314183.3323676

[4]   American Psychiatric Association. 2018. *App evaluation model*. https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model. Accessed 28 November 2021

[5]   American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders (5th ed.)*. American Psychiatric Association, Washington DC.

[6]   American Psychological Association. 2017. *Ethical principles of psychologists and code of conduct*. American Psychological Association, Washington DC.

[7]   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

[8]   Gil Appel, Barak Libai, Eitan Muller, and Ron Shachar. 2020. On the monetization of mobile apps. *International Journal of Research in Marketing* 37, 1 (March 2020), 93-107. https://doi.org/10.1016/j.ijresmar.2019.07.007

[9]   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

[10]   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

[11]   Brooke Auxier, Ariane Bucaille, and Kevin Westcott. 2021*. Mental health goes mobile: the mental health app market will keep on growing.* https://www2.deloitte.com/us/en/insights/industry/technology/technology-media-and-telecom-predictions/2022/mental-health-app-market.html

[12]   Nilufar Baghaei, Lehan Stemmet, Andrej Hlasnik, Konstantin Emanov, Sylvia Hach, John Naslund, Mark Billinghurst, 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

[13]   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

[14]   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

[15]   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

[16]   Alfred Philip Benigno, Samuel Lomugdang, 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

[17]   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

[18]   Ushma Bhatt, Divya Iyyani, Keshin Jani, and Swati Mali. 2018. Troll-detection systems limitations of troll detection systems and AI/ML anti-trolling solution. In *2018 3rd International Conference for Convergence in Technology (I2CT)*, Institute of Electrical and Electronics Engineers, New Jersey, USA, 1-6. https://doi.org/10.1109/I2CT.2018.8529342

[19]   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.

[20]   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

[21]   Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2, 77-101. https://doi.org/10.1191/1478088706qp063oa

[22]   The British Psychological Society. 2018. *Codes of ethics and conduct.* The British Psychological Society, Leicester, UK

[23]   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

[24]   Tracy Butryn, Leah Bryant, Christine Marchionni, and Farhad Sholevar. 2017. The shortage of psychiatrists and other mental health providers: causes, current state, and potential solutions. *Int J Acad Med* 3, 1 (July 2017), 5-9. https://doi.org/10.4103/IJAM.IJAM_49_17

[25]   Clara Caldeira, Yu Chen, Lesley Chan, Vivian Pham, Yunan Chen, and Kai Zheng. 2018. Mobile apps for mood tracking: an analysis of features and user reviews. *AMIA Annu Symp Proc 2017* (April 2018), 495-504. Hannah Capon, Wayne Hall, Craig Fry, and Adrian Carter. 2016. Realising the technological promise of smartphones in addiction research and treatment: an ethical review. *Int J Drug Policy* 36 (October 2016), 47-57. https://doi.org/10.1016/j.drugpo.2016.05.013

[26]   Steven Chan, John Torous, Ladson Hinton, and Peter Yellowlees. 2015. Towards a framework for evaluating mobile mental health apps. *Telemedicine and e-Health* 21, 12 (December 2015), 1038-1041. https://doi.org/10.1089/tmj.2015.0002

[27]   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

[28]   Liam Crosby and Oliver Bonnington. 2020. Experiences and implications of smartphone apps for depression and anxiety. *Sociol Health Illn* 42

(March 2020), 925-942. https://doi.org/10.1111/1467-9566.13076

[29] Claudia Daudén Roquet and Corina Sas. 2021. Interoceptive interaction: an embodied metaphor inspired approach to designing for meditation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 265, 1–17. https://doi.org/10.1145/3411764.3445137

[30] Claudia Daudén Roquet and Corina Sas. 2021. A mindfulness-based brain-computer interface to augment mandala coloring for depression: protocol for a single-case experimental design. *JMIR Res Protoc* 10, 1 (January 2021), e20819. https://doi.org/10.2196/20819

[31] Claudia Daudén Roquet, Corina Sas, and Dominic Potts. 2021. Exploring Anima: a brain–computer interface for peripheral materialization of mindfulness states during mandala coloring, *Human–Computer Interaction* (November 2021), 1-41. https://doi.org/10.1080/07370024.2021.1968864   Fernando Estrada Martinez de Alva, Greg Wadley, and Reeva Lederman. 2015. It feels different from real life: users' opinions of mobile applications for mental health. In *Proceedings of the annual meeting of the Australian special interest group for computer human interaction (OzCHI '15)*. Association for Computing Machinery, New York, NY, USA, 598-602. https://doi.org/10.1145/2838739.2838806

[32] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Major life changes and behavioral markers in social media: case of childbirth. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 1431–1442. https://doi.org/10.1145/2441776.244193

[33] Munmun De Choudhury, Sanket S Sharma, Tomaz Logar, Wouter Eekhout, and René Clausen Nielsen. 2017. Gender and cross-cultural differences in social media disclosures of mental illness. 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, 353–369. https://doi.org/10.1145/2998181.299822

[34] Steven Denisevicz 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

[35] 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

[36] 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

[37] 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

[38] 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

[39] Kyriaki G Giota and George Kleftaras. 2014. Mental health apps: innovations, risks and ethical considerations. *Ehealth Telecommun Syst Netw* 3, 3 (September 2014), 19-23. https://doi.org/10.4236/etsn.2014.33003

[40] Richard Griffith. 2016. What is Gillick competence? *Hum Vaccin Immunother* 12, 1 (January 2016), 244-247. https://doi.org/10.1080/21645515.2015.1091548

[41] Marissa S Gross, Amelia Hood, and Bethany Corbin. 2021. Pay no attention to that man behind the curtain: an ethical analysis of the monetization of menstruation app data. *IJFAB: International Journal of Feminist Approaches to Bioethics* 14, 2 (April 2021), 144-156. https://doi.org/10.3138/ijfab-2021-03-22

[42] Quinn Grundy, Kellia Chiu, and Lisa Bero. 2019. Commercialization of user data by developers of medicines-related apps: a content analysis. *J Gen Intern Med* 34, 12 (December 2019), 2833-2841. https://doi.org/10.1007/s11606-019-05214-0

[43] 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

[44] 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

[45] Kit Huckvale, John Torous, and Mark E Larsen. 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

[46] 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

[47] Richard Huxtable. 2013. For and against the four principles of biomedical ethics. *Clinical Ethics* 8, 2/3 (September 2013), 39-43. https://doi.org/10.1177/1477750913486245

[48] Terry Hyland. 2016. The erosion of right livelihood: counter-educational aspects of the commodification of mindfulness practice. *Person-Centered & Experiential Psychotherapies* 15, 3 (July 2016), 177-189. https://doi.org/10.1080/14779757.2016.1179666

[49] 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

[50]   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

[51]   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

[52]   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

[53]   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

[54]   Sarah E Knowles, Gill Toms, Caroline Sanders, Penny Bee, Karina Lovell, Stefan Rennick-Egglestone, David Coyle, Catriona M Kennedy, Elizabeth Littlewood, David Kessler, Simon Gilbody, and Peter Bower. 2014. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. *PLoS ONE* 9, 1 January 2014), e84323. https://doi.org/10.1371/journal.pone.0084323

[55]   Tobias Krieger, David Altenstein, Isabelle Baettig, Nadja Doerig, and Martin Grosse Holtforth. 2013. Self-compassion in depression: associations with depressive symptoms, rumination, and avoidance in depressed outpatients. *Behav Ther* 44, 3 (September 2013), 501-513. https://doi.org/10.1016/j.beth.2013.04.004

[56]   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

[57]   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

[58]   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

[59]   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

[60]   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

[61]   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

[62]   Reeva Lederman, Simon D'Alfonso, Simon Rice, Simon Coghlan, Greg Wadley, and Mario Alvarez-Jimenez. 2020. Ethical issues in online mental health interventions. In *European Conference on Information Systems (ECIS) 2020 Research Papers*. Association for Information Systems, Atlanta, GA, USA. https://aisel.aisnet.org/ecis2020_rp/66/

[63]   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

[64]   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

[65]   Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 501–510. https://doi.org/10.1145/2370216.2370290

[66]   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

[67]   Deborah Lupton. 2014. The commodification of patient opinion: the digital patient experience economy in the age of big data. *Sociol Health Illn* 36, 6 (August 2014), 856-869. https://doi.org/10.1111/1467-9566.12109

[68]   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

[69]   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

[70]   Jaime M Marshall, Debra A Dunstan, and Warren Bartik. 2020. Clinical or gimmickal: the use and effectiveness of mobile mental health apps for treating anxiety and depression. Aust N Z J Psychiatry 54, 1 (January 2020), 20-28. https://doi.org/10.1177/0004867419876700

[71]   Jaime M Marshall, Debra A Dunstan, and Warren Bartik. 2020. Smartphone psychology: new approaches towards safe and efficacious mobile mental health apps. *Professional Psychology: Research and Practice* 51, 3 (June 2020), 214-222. https://doi.org/10.1037/pro0000278

[72] Wei Meng, Ren Ding, Simon P Chung, Steven Han, and Wenke Lee. 2016. The price of free: privacy leakage in personalized mobile in-apps ads. In *Network and Distributed System Security Symposium (NDSS'16)*. Internet Society, Reston, VA, USA. https://doi.org/10.14722/ndss.2016.23353

[73] Mental Health Foundation. 2016. *Fundamental facts about mental health 2016*. Mental Health Foundation, London.

[74] Isaac Moshe, Yannik Terhorst, Kennedy Opoku Asare, Lasse Bosse Sander, Denzil Ferreira, Harald Baumeister, David Mohr, and Laura Pulkki-Råback. 2021. Predicting symptoms of depression and anxiety using smartphone and wearable data. *Front. Psychiatry* 12 (January 2021), 625247. https://doi.org/10.3389/fpsyt.2021.625247

[75] Camille Nadal, Shane McCully, Kevin Doherty, Corina Sas, and Gavin Doherty. 2022. The TAC toolkit: supporting the design for user acceptance of health technologies from a macro-temporal perspective. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI'22)*. Association for Computing Machinery, New York, NY, USA.

[76] Camille Nadal, Corina Sas, and Gavin Doherty. 2020. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. *J Med Internet Res* 22, 7 (July 2020), e17256. https://doi.org/10.2196/17256

[77] Jennifer Nicholas, Andrea S Fogarty, Katherine Boydell, and Helen Christensen. 2017. The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. *J Med Internet Res* 19, 4 (April 2017), e105. https://doi.org/10.2196/jmir.7273

[78] Stefanos Nikiforos, Spyros Tzanavaris, and Katia-Lida Kermanidis. 2020. Virtual learning communities (VLCs) rethinking: influence on behavior modification—bullying detection through machine learning and natural language processing. *Journal of Computers in Education* 7 (May 2020), 531-551. https://doi.org/10.1007/s40692-020-00166-5

[79] Leysan Nurgalieva, David O'Callaghan, and Gavin Doherty. 2020. Security and privacy of mHealth applications: a scoping review. *IEEE Access* 8 (June 2020), 104247-104268. https://doi.org/10.1109/ACCESS.2020.2999934

[80] Harri Oinas-Kukkonen and Marja Harjumaa. 2009. Persuasive system design: key issues, process model, and system features. *Communications of the Association for Information Systems 24*, 28 (March 2009), 485–500. https://doi.org/10.17705/1CAIS.02428

[81] 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

[82] ORCHA. https://appfinder.orcha.co.uk/. Accessed 10 January 2022.

[83] 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

[84] Oladapo Oyebode, 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.

[85] Kathleen M Palmer and Vanessa Burrows. 2021. Ethical and safety concerns regarding the use of mental health–related apps in counseling: considerations for counselors. *J Technol Behav Sci* 6 (March 2021), 137-150. https://doi.org/10.1007/s41347-020-00160-9

[86] 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

[87] 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

[88] 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

[89] 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

[90] Oliver Rauprich. 2012. Principlism. In Ruth Chadwick (Ed) *Encyclopedia of Applied Ethics (Second Edition)*. https://doi.org/10.1016/B978-0-12-373932-2.00244-1

[91] 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

[92] 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

[93] 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

[94] 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

[95] 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

[96] 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

[97] 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

[98] Pedro Sanches, Kristina Höök, Corina Sas, and Anna Ståhl. 2019. Ambiguity as a resource to inform proto-practices: the case of skin conductance. *ACM Transactions on Computer-Human Interaction* 26, 4 (August 2019), article 21, 1-32. https://doi.org/10.1145/3318143

[99] Corina Sas, Nigel Davies, Sarah Clinch, Peter Shaw, Mateusz Mikusz, Madeleine Steeds, and Lukas Nohrer. 2020. Supporting stimulation needs in dementia care through wall-sized displays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376361

[100] 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

[101] Corina Sas, Steve Whittaker, Steven Dow, Jodi Forlizzi, and John Zimmerman. 2014. Generating implications for design through design research. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1971–1980. https://doi.org/10.1145/2556288.255735

[102] 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

[103] 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

[104] 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

[105] Vivian Ta, Caroline Griffith, Carolynn Boatfield, Xinyu Wang, Maria Civitello, Haley Bader, Esther DeCero, and Alexia Loggarakis. 2020. User experiences of social support from companion chatbots in everyday contexts: thematic analysis. *J Med Internet Res* 22, 3 (March 2020), e16235. https://doi.org/10.2196/16235

[106] Ailie K Y Tang. 2016. Mobile app monetization: app business models in the digital era. *IJIMT* 7, 5 (October 2016), 224-227. https://doi.org/10.18178/ijimt.2016.7.5.677

[107] Torbjorn Tannsjo. 2013. *Understanding Ethics: An Introduction to Moral Theory*. Edinburgh: Edinburgh University Press.

[108] Kong Saoane Thach. 2018. User's perception on mental health applications: a qualitative analysis of user reviews. In *2018 5th NAFOSTED Conference on Information and Computer Science (NICS)*. Institute of Electrical and Electronics Engineers, New Jersey, USA, 47-52. https://doi.org/ 10.1109/NICS.2018.8606901

[109] Kong Saoane Thach and Thi Phuong Nam Phan. 2019. Persuasive design principles in mental health apps: a qualitative analysis of user reviews. In *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*. Institute of Electrical and Electronics Engineers, New Jersey, USA, 1-6. https://doi.org/ 10.1109/RIVF.2019.8713753

[110] John Torous, Gerhard Andersson, Andrew Bertagnoli, Helen Christensen, Pim Cuijpers, Joseph Firth, Adam Haim, Honor Hsin, Chris Hollis, Shôn 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

[111] 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

[112] John Torous, John Luo, and Steven R Chan. 2018. Mental health apps: what to tell patients. *Curr Psychiatry* 17, 3 (March 2018), 21-25.

[113] 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

[114] John Torous and Laura Weiss Roberts. 2017. The ethical use of mobile health technology in clinical psychiatry. *J Nerv Ment Dis* 205, 1 (January 2017), 4-8. https://doi.org/10.1097/NMD.0000000000000596

[115] Muhammad Umair, Corina Sas, and Miquel Alfaras. 2020. ThermoPixels: toolkit for personalizing arousal-based interfaces through hybrid crafting. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1017–1032. https://doi.org/10.1145/3357236.3395512

[116] UnivDatos Market Insights. 2021. *Mental health apps market: current analysis and forecast (2021-2027)*. https://univdatos.com/report/mental-health-apps-market/

[117] 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

[118] 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

[119] 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

[120] 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

[121] 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

[122] James Williamson, Elizabeth Godoy, Miriam Cha, Adrianne Schwarzentruber, Pooya Khorrami, 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

[123] 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

[124] World Health Organization. 2021. *Global strategy on digital health 2020-2025*. World Health Organization, Geneva.

[125] 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

[126] 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

[127] Jennifer Zelmer, Krystle van Hoof, MaryAnn Notarianni, Trevor van Mierlo, Megan Schellenberg, and Cara Tannenbaum. 2018. An assessment framework for e-Mental health apps in Canada: results of a modified Delphi process. *JMIR Mhealth Uhealth* 6, 7 (July 2018), e10016. https://doi.org/10.2196/10016

[128] 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	Users' rating (/5)		No. of user ratings	
		Google Play	*Apple App Store*	*Google Play*	*Apple App Store*
Headspace	Headspace, Inc.	4.6	4.8	85394	160019
Doctor on Demand	Doctor on Demand, Inc.	4.7	4.1	20593	13
Simple Habit Meditation	Simple Habit, Inc.	4.7	-	15601	-
TalkLife	TalkLife	4.5	4.5	15498	625
Medication Reminder & Pill Tracker	MyTherapy	4.6	-	14701	-
7 Cups	7 Cups of Tea	4.2	4.4	14431	666
Abide: Christian Meditation & Prayers	Carpenter Code Inc.	4.7	-	13971	-
Your.MD: Symptom Checker	Your: MD	4.3	4.7	12711	195
Pacifica: Stress & Anxiety	Pacifica Labs Inc.	4.4	4.7	10778	486
Youper	Youper, Inc.	4.8	4.9	10151	517
Relaxing Anti-Stress Sounds	Dandelion Soft	4.0	-	9935	-
Wysa	Touchkin	4.5	3.6	9620	28
Stop, Breathe & Think	Stop, Breathe, Think	4.3	-	8889	-
Gentle Wakeup: Sleep & Alarm Clock	Dr Alexander Rieger	4.4	-	6705	-
Moodpath: Depression & Anxiety Test	Moodpath	4.6	4.6	6680	2019
Secret Diary	Zheko	3.8	-	6601	-
Aware: Meditation & Mindfulness	zoojoo.be	4.7	-	5157	-
SuperBetter	SuperBetter, LLC	4.3	4.4	5036	431
Remente: Self Improvement	Remente	4.4	4.6	4280	445
BetterHelp: Online Counselling	BetterHelp	3.9	3.5	4082	244

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

App name	Developer	Users' rating (/5)		No. of user ratings	
		Google Play	*Apple App Store*	*Google Play*	*Apple App Store*
Depression Test	FXT Tech	2.5	-	8	-
DSM-5 Differential Diagnosis	Unbound Medicine, Inc.	2.7	-	122	-
Depression Support	MyHealth Teams	2.9	-	23	-
MoodHacker	ORCAS	2.9	1.0	16	1
My Possible Self	My Possible Self Ltd	3.0	4.4	23	15
Social Force (IntelliCare)	CBITs	3.0	-	14	-
Anxiety & Depression Symptoms	Twayesh Projects	3.1	-	307	-
Talkspace Counselling & Therapy	Talkspace	3.4	3.8	1639	71
WellMind	Blue Step Solutions	3.4	3.3	82	24
IntelliCare Hub	CBITs	3.4	Not listed	28	Not listed
Worry Knot (Intellicare)	CBITs	3.4	Not listed	26	Not listed
UpLift for Depression	UpLift	3.4	-	9	-
MHF	Together for Change	3.5	-	40	-
WellTrack: Interactive Self-Help Therapy	CyberPsyc	3.5	-	31	-
Depression Test	Japps Medical	3.5	-	1385	-
CogniFit Brain Fitness	CogniFit Inc	3.6	-	593	-
Depression Self-Help Guide:CBT	Xandy App Ideas	3.6	-	8	-
Depression Test	MoodTools	3.7	Not listed	199	Not listed
Aware (diagnose yourself)	Heretic Hammer	3.7	-	23	-
Slumber Time (Intellicare)	CBITs	3.7	-	14	-

A.3 Summary of app interventions and reviews included in analysis (sorted from high to low ratings)

App name	Intervention (from app store description)	Type of review [n (%)]		
		+ve	Ambivalent	-ve
Youper	ACT | CBT	112 (79)	22 (15)	7 (5)
Abide: Christian Meditation & Prayers	Mindfulness or meditation	30 (61)	15 (31)	4 (8)
Aware: Meditation & Mindfulness	Mindfulness or meditation	36 (72)	12 (24)	2 (4)
Simple Habit Meditation	Mindfulness or meditation	40 (80)	8 (16)	2 (4)
Headspace	Mindfulness or meditation	62 (38)	39 (24)	62 (38)
Medication Reminder & Pill Tracker	Medication management	30 (60)	10 (20)	10 (20)
Moodpath: Depression & Anxiety Test	Assessment | CBT | Mindfulness or meditation | Psychoeducation	63 (56)	37 (33)	13 (11)
Pacifica: Stress & Anxiety	CBT | Mindfulness or meditation | Psychosocial | Self-help	96 (69)	27 (19)	16 (12)
TalkLife	Psychosocial	76 (47)	36 (22)	49 (31)
Remente: Self Improvement	Psychoeducation | Self-help	60 (71)	17 (20)	8 (9)
Your.MD: Symptom Checker	Assessment | Psychoeducation	79 (75)	5 (5)	21 (20)
Doctor on Demand	Online therapy	46 (84)	4 (7)	5 (9)
Gentle Wakeup: Sleep & Alarm Clock	Mindfulness or meditation	38 (76)	7 (14)	5 (10)
SuperBetter	BT | Positive psychology | Psychosocial	46 (39)	30 (25)	43 (36)
Stop, Breathe & Think	Mindfulness or meditation	36 (74)	5 (10)	8 (16)
7 Cups	Online therapy | Psychosocial	42 (29)	46 (32)	56 (39)
Wysa	CBT | DBT | Exercise	36 (60)	16 (27)	8 (13)
Relaxing Anti-Stress Sounds	Sound therapy	29 (58)	12 (24)	9 (18)
Secret Diary	Emotional awareness | Monitoring and tracking	32 (65)	16 (33)	1 (2)
BetterHelp: Online Counselling	Online therapy	57 (43)	14 (10)	63 (47)
Aware (diagnose yourself)	Assessment	2 (25)	0	6 (75)
Depression Test (MoodTools)	Assessment	14 (70)	3 (15)	3 (15)
Slumber Time (IntelliCare)	Sound therapy	1 (20)	2 (40)	2 (40)
My Possible Self	CBT | IPT | Positive psychology | PST | Self-help	3 (17)	3 (17)	12 (66)
CogniFit Brain Fitness	Assessment | Cognitive training	7 (19)	6 (16)	24 (65)
Depression Self-Help Guide:CBT	Psychoeducation | Self-help	1 (33)	0	2 (67)
Talkspace Counselling & Therapy	Online therapy	23 (31)	10 (14)	41 (55)

App name	Intervention (from app store description)	Type of review [n (%)]		
		+ve	*Ambivalent*	*-ve*
Depression Test (Japps Medical)	Assessment	27 (54)	19 (38)	4 (8)
MHF	Psychoeducation | Psychosocial	5 (25)	4 (20)	11 (55)
WellTrack: Interactive Self-Help Therapy	ACT | Assessment | CBT | Self-help	12 (52)	3 (13)	8 (35)
IntelliCare Hub	Assessment | Psychoeducation	1 (8)	2 (15)	10 (77)
UpLift for Depression	CBT	2 (11)	4 (22)	12 (67)
Worry Knot (Intellicare)	Cognitive reappraisal | Distraction or grounding | Monitoring and tracking | Problem solving | Skills building	3 (50)	0	3 (50)
WellMind	Psychoeducation | Self-help	13 (35)	7 (19)	17 (46)
Anxiety & Depression Symptoms	Psychoeducation	9 (56)	1 (6)	6 (38)
Social Force (IntelliCare)	Psychosocial	0	0	2 (100)
Depression Support	Psychosocial	2 (12)	2 (12)	13 (76)
DSM-5 Differential Diagnosis	Assessment	6 (20)	1 (3)	23 (77)
Depression Test (FXT Tech)	Assessment	1 (100)	0	0
MoodHacker	CBT | Positive psychology | Psychosocial | Self-help	0	2 (67)	1 (33)