Expert Review of Medical Devices

ISSN: 1743-4440 (Print) 1745-2422 (Online) Journal homepage: www.tandfonline.com/journals/ierd20

Artificially intelligent chatbots in digital mental health interventions: a review

Eliane M. Boucher, Nicole R. Harake, Haley E. Ward, Sarah Elizabeth Stoeckl, Junielly Vargas, Jared Minkel, Acacia C. Parks & Ran Zilca

To cite this article: Eliane M. Boucher, Nicole R. Harake, Haley E. Ward, Sarah Elizabeth Stoeckl, Junielly Vargas, Jared Minkel, Acacia C. Parks & Ran Zilca (2021) Artificially intelligent chatbots in digital mental health interventions: a review, Expert Review of Medical Devices, 18:sup1, 37-49, DOI: 10.1080/17434440.2021.2013200

To link to this article: https://doi.org/10.1080/17434440.2021.2013200

Published online: 31 Dec 2021.

EXPERT REVIEW OF MEDICAL DEVICES
2021, VOL. 18, NO. S1, 37-49
https://doi.org/10.1080/17434440.2021.2013200

Taylor & Francis
Taylor & Francis Group

PERSPECTIVE

OPEN ACCESS Check for updates

Artificially intelligent chatbots in digital mental health interventions: a review

Eliane M. Boucher, Nicole R. Harake, Haley E. Ward, Sarah Elizabeth Stoeckl, Junielly Vargas, Jared Minkel, Acacia C. Parks and Ran Zilca

Happify Health, New York, NY, USA

ABSTRACT

Introduction: Increasing demand for mental health services and the expanding capabilities of artificial intelligence (AI) in recent years has driven the development of digital mental health interventions (DMHIs). To date, AI-based chatbots have been integrated into DMHIs to support diagnostics and screening, symptom management and behavior change, and content delivery.

Areas covered: We summarize the current landscape of DMHIs, with a focus on AI-based chatbots. Happify Health's AI chatbot, *Anna*, serves as a case study for discussion of potential challenges and how these might be addressed, and demonstrates the promise of chatbots as effective, usable, and adoptable within DMHIs. Finally, we discuss ways in which future research can advance the field, addressing topics including perceptions of AI, the impact of individual differences, and implications for privacy and ethics.

Expert opinion: Our discussion concludes with a speculative viewpoint on the future of AI in DMHIs, including the use of chatbots, the evolution of AI, dynamic mental health systems, hyper-personalization, and human-like intervention delivery.

ARTICLE HISTORY
Received 26 August 2021
Accepted 29 November 2021

KEYWORDS
Artificial intelligence (AI); chatbots; digital mental health; digital mental health interventions (dmhis)

1. Introduction

Approximately one in five adults in the United States (U.S.) struggle with mental illness [1–3]; however, many of these individuals are not receiving treatment. For example, in one nationally representative sample, only 41.1% of U.S. adults with a diagnosis of anxiety, mood, impulse control, and/or substance disorders received treatment in the previous 12 months [4]. Worldwide, it is estimated that 70% of people with mental illness receive no formal treatment [5]. While some individuals do not seek treatment because of low perceived need [6] or attitudinal barriers like perceived stigma [7], those who desire treatment may not be able to receive it quickly due to a shortage of mental health professionals, particularly in rural and low income areas [8]. Indeed, primary care physicians report that obtaining outpatient mental health services for patients is more difficult than other common referrals [9].

Issues with the accessibility of mental health care garnered widespread attention during the COVID-19 pandemic, when access to care became more difficult. In addition to typical barriers to treatment, restrictions and lockdowns enacted to mitigate the spread of COVID-19 caused widespread disruptions to face-to-face care. In a survey conducted by the World Health Organization in Summer 2020, 60% of the 130 reporting countries indicated disruptions to mental health services, and disruptions to services for older adults and for youth were closer to 70% [10]. In particular, travel restrictions and the closure of community-based mental health services near people's homes led to even greater difficulties accessing care for those in lower-income countries [10].

Accessibility issues have been compounded by the fact that there has been greater demand for mental health services during the pandemic. In the U.S., more than one-third of the population experienced depression or anxiety symptoms, almost a three-fold increase compared to 2019 [11]. . Mental and physical health issues, such as difficulty with sleeping and/or eating, the reemergence of trauma symptoms, substance abuse, and worsening of chronic conditions, have also increased during the pandemic [12,13]. Vulnerable and marginalized populations have experienced the greatest increase in mental health concerns, perhaps due to the larger unmet need for care among these groups [14,15]

Altogether, this has created a perfect storm where an already overwhelmed system has fewer resources and greater demand, causing an increase in unmet need for mental health services worldwide. Not surprisingly, research suggests the unmet need for psychotherapy and counseling, in addition to the disruption of traditional services, has increased during the COVID-19 pandemic [16]. More specifically, over a quarter of individuals with elevated levels of depression and anxiety, and 12.8% of the U.S. population, reported an unmet need for mental health counseling or therapy [15].

2. Digital mental health interventions

As access to smartphones and the Internet continues to increase [17], digital solutions to mental health care are an important avenue to addressing accessibility issues with traditional in-person care [18]. Digital mental health interventions (DMHIs) can often be disseminated to large populations

CONTACT Eliane M. Boucher eliane@happify.com Happify Health, New York, NY, USA

Article highlights

- Digital mental health interventions (DMHIs) are an important avenue to address the treatment gap for mental health.
- Artificial intelligence (AI) offers a scalable means of providing support and improving engagement within DMHIs.
- Although AI chatbots have been designed for several mental health disorders and can serve various functions in DMHIs, adoption of chatbots in DMHIs remains limited.
- Research suggests users perceive chatbots favorably, chatbots can help improve engagement, and may help improve mental health outcomes.
- More rigorous research is needed to test the effects of chatbots within DMHIs, including understanding general perceptions of chatbots, the impact of individual and contextual factors, and whether chatbots can improve mental health outcomes beyond non-AI driven DMHIs.
- Given the variability in chatbots and their applications within DMHIs, greater differentiation of types of chatbots, and the corresponding functions they serve, is needed.

[19,20] and underserved areas [21]. In fact, in response to concerns about the impact of the pandemic on mental health and accessibility concerns for mental health services, the U.S. Food and Drug Administration (FDA) released new temporary guidance in April 2020 allowing for the distribution of digital therapeutics for some psychiatric conditions, including major depressive disorder and generalized anxiety disorder, without traditional clearance [22,23].

Although DMHIs have been on the market for over a decade [24], they have become more common in recent years. It is estimated that there are over 10,000 mental health applications currently available for download [25], though as few as 2% of these are supported by empirical evidence [26]. However, the subset of applications tested empirically has been shown to be effective. Meta-analyses of DMHIs show small to medium effects on depression and anxiety symptoms [27–30], with larger effects when programs include more engagement features [31], though other meta-analyses found significant pooled effects for depression but not for other psychological outcomes, including anxiety [32]. Similarly, a review of DMHIs among college students found that most programs (81%) were at least partially effective in improving depressive symptoms, anxiety, psychological distress, or well-being [33]. In addition, DMHIs also may be effective preventative tools (see Ebert et al., 2017, for a discussion) [21].

DMHIs are effective when used as recommended; however, these interventions have been criticized for low usage rates and high dropout [34–37]. One potential reason for the low engagement in many DMHIs may be the lack of guidance and feedback [38–40]. Indeed, people report that the lack of contact with a therapist is a disadvantage to DMHIs [41], and people using purely self-guided programs report more difficulties with using the program or remembering to engage with it [40].

Consequently, some interventions include guidance or support, including support from a therapist or from peers, or more administrative forms of support (i.e. non-clinical support, primarily logistical) offered by people other than clinicians (e.g. nurses, research coordinators, laypeople) [42]. Research suggests this is an effective mechanism for boosting engagement. Meta-analyses have shown that dropout rates are highest for unsupported interventions and lowest for therapist-supported interventions [42]. Similarly, adding therapist support via telephone coaching [43] or non-clinical support via peer-to-peer feedback [44] increases adherence and engagement. DMHIs with support from a therapist also tend to show greater improvement in mental health outcomes compared to unsupported interventions [42,45,46], although other studies show no differences in outcomes despite greater engagement [43,44].

Incorporating support from a therapist may help address the problem with engagement in DMHIs, but doing so may also reduce some perceived benefits of these interventions. For instance, people completing a DMHI for depression incorporating synchronous support by way of telephone coaching reported that scheduling these phone calls was difficult due to time constraints [43]. Asynchronous support may reduce issues with scheduling, but research suggests few patients involved in a DMHI with an option to send messages to a therapist and receive asynchronous feedback took advantage of this feature; most patients sent fewer than two messages over the course of 14 weeks [47]. Furthermore, incorporating support from a live person within DMHIs diminishes the scalability of those interventions, particularly when support is offered by a therapist, given the shortage of mental health professionals.

3. Overview of artificial intelligence in digital interventions

The digital mental health care industry has recently begun to incorporate AI into existing platforms and create AI-guided products [48]. This is being done in many ways, such as health communication, virtual reality, symptom and biomarker monitoring, mental health triage, digital phenotyping to predict outcomes, and personalization of content [49–51].

Another popular utilization of AI in DMHIs is artificially intelligent *chatbots*, also referred to as *conversational agents* or *relational agents*; these are computer programs integrated into DMHIs that are able to hold a conversation with the human user [52]. The first credited chatbot, ELIZA, was developed by Joseph Weizenbaum in 1966 [53]. ELIZA was programmed to respond based on a Rogerian psychotherapeutic approach, searching user input for keywords and then applying a rule based on those keywords to provide a response. Since ELIZA, interest in chatbots has increased considerably, particularly after 2016 [54] and within DMHIs. One review found that 39% of health chatbots focused on mental health issues [55], and another review reported that 41 mental health chatbots were developed in 2019 alone [56]. Currently, most mental health chatbots have been designed for depression or anxiety [57], though chatbots have also been developed to support patients with autism [58], suicide risk [59], substance abuse [60], post-traumatic stress disorder, stress [61], dementia [62], and acrophobia [63], and to enhance positive psychological constructs, such as psychological well-being [61], self-compassion [64], mindfulness, and quality of life [65]. Chatbots also have been developed for various groups including children, adolescents, adults, elders [66], and specific clinical populations [67]. Given this tremendous growth, we focus the remainder of our review on the potential for AI-based chatbots within DHMI rather than other applications of AI.

4. Functions of chatbots in DMHIs

Chatbots can be integrated into different platforms, including mobile applications, websites, SMS texting, smart technologies, and virtual reality. Chatbots also vary in their complexity of interaction; they can rely on systems ranging from straightforward rule-based models, like ELIZA, to more advanced AI models using natural language processing (NLP) and machine learning [48]. Typically, after analyzing user dialogue content, chatbots respond through text-based or voice-enabled conversations [68]. User input is primarily written, via open text or multiple-choice options, while output generated by the chatbot can be written, spoken, or visual [69].

Chatbots have been incorporated into DMHIss to perform various functions, ranging from assistance, screening, psychoeducation, therapeutic intervention, monitoring behavior changes, and relapse prevention [70]. We briefly review some of the most common functions: diagnosis, content delivery, and symptom management.

4.1. Diagnosis

One way chatbots can help reduce the burden on health care professionals is to diagnose and triage people with mental health concerns, which may help to prioritize in-person services for those who need it most. Chatbots have been used as a diagnostic or screening tool for dementia, substance abuse, stress, depression and suicide, anxiety disorders, and PTSD. In this context, people interact with the chatbot as they would with a human being. Through a series of queries, the chatbot identifies the user's symptoms, predicts the disease, and recommends treatment or provides information about the diagnosis to the patient [62,71–79]. This application of chatbots is controversial; in one survey of mental health professionals, 51% felt using chatbots for diagnostic purposes was problematic [80]. However, using AI for diagnostic purposes can help identify people who are at risk, allowing for earlier intervention and helping to reduce the likelihood of future problems [81].

Research on the effectiveness of chatbots in diagnostic functions remains limited, but preliminary evidence on one diagnostic chatbot, *Ada*, suggests moderate agreement between the chatbot's diagnosis and conditions depicted in a vignette. Importantly, however, agreement was higher when psychotherapists entered symptoms into the app based on the vignette, but rather low when entered by laypersons [82]. Even in cases where chatbots do not perform diagnoses themselves, they may help improve engagement with mental health assessments, increasing the likelihood of identifying people in need of care [83]. In addition, other types of AI may effectively predict the onset of certain mental health conditions by utilizing behavioral and self-report data. For example, machine learning algorithms can predict psychosis onset with 79% accuracy, and Attention-Deficit Hyperactivity Disorder and Autism Spectrum Disorder with 96% accuracy [84].

4.2. Content delivery

The most common application of chatbots within DMHIs is to deliver content. While chatbots cannot simulate traditional psychotherapy, they may be able to administer psychotherapeutic interventions that do not involve a high degree of therapeutic competence [85]. For example, some chatbots using NLP can simulate a therapeutic conversational style that implements and teaches users about various therapeutic techniques (see Fitzpatrick et al., 2017, for a discussion) [51]. Chatbots employing principles of cognitive behavioral therapy (CBT) are the most common and well-studied; one meta-analysis found that 10 out of 17 chatbots primarily used CBT [86]. One such chatbot, *Woebot* [87], delivers CBT to users via instant messaging employing NLP and mimicking human clinicians and social discourse [85].

However, other chatbots utilize a variety of therapeutic approaches, such as acceptance and commitment therapy and mindfulness [87,88]. *Wysa*, a chatbot described as 'emotionally intelligent,' draws on a variety of therapeutic approaches including CBT, dialectical behavior therapy, and motivational interviewing [89,90]. Some chatbots have been developed for more specific applications, like *Vivibot*, which helps young people learn positive psychology skills after cancer treatment to support anxiety reduction [67]; whereas others, like *MYLO*, employ general self-help strategies when users are in distress and work towards suicide prevention [91] (these are less common as they involve greater risk [52]).

4.3. Management and screening of symptoms

Chatbots may also help monitor a patient's progress or track symptoms and behaviors (e.g. physical activity, hours of sleep, time spent on social media) [92]. Chatbots are currently being used as personal health assistants to promote well-being and mental health check-ins throughout and after completing an intervention [93]. In this capacity, chatbots can help users facilitate the transfer of therapeutic content into their daily lives, assess progress, and provide personalized support by delivering additional mental health resources [52]. AI can also help improve the personalization of care by facilitating more efficient storage and processing of user information [84], which can subsequently allow users to better understand when symptoms flare up or decline. In turn, this can help improve self-management of symptoms and risk of relapse [52], particularly among people without access to a mental health professional [94–96]. This type of management and screening can also be used after traditional in-person interventions or in outpatient settings. In this case, chatbots could help maintain benefits from treatment by reminding clients of skills and practices (e.g. medication adherence, check-ups, exercise, etc.) [93,97].

5. The evidence for chatbots in DHMI

Although chatbots are becoming increasingly popular in DHMI, few chatbot-based DMHIs describe the evidence supporting their programs, and even fewer in-market chatbots have been tested in rigorous, empirical research [57]. What research exists, albeit limited, does suggest mental health chatbots can be effective, acceptable to users, and promote engagement.

5.1. Acceptability of chatbots

Installations of publicly available mental health apps with integrated chatbots are high, suggesting people are interested in these programs [57]. Users of mental health chatbots also generally report high satisfaction with chatbot interactions, positive perceptions of chatbots, prefer chatbots to information control groups, and indicate interest in using chatbots in the future [57,98]. In one study, 68% of users found *Wysa* encouraging and helpful [90]. In particular, people are more satisfied when they perceive conversations as private, reporting learning something new during the interaction when chatbot content is similar to what their therapist recommended previously and perceived to be of high quality, when there is appropriate usage of high-quality technological elements, and when the chatbot's tone or voice is consistent [57]. Moreover, compared to human beings, chatbots are perceived as less judgmental, which facilitates self-disclosure among users, and allows for more conversational flexibility [67,84,99]. In fact, some people prefer to interact with chatbots over mental health professionals [74], which may encourage people who would not normally seek therapy to receive care.

However, there is variability in how people perceive chatbots. In one study, 32% of participants reported the chatbot was unhelpful [90]. In some cases, users have even reported that their interactions with the chatbot bothered them [90], or that the chatbot was self-focused [90] or annoying [67]. Indeed, concerns regarding the accuracy, trustworthiness, and privacy of chatbots have all emerged as potential barriers to engagement and adoption [100]. Some of this variability in perceptions and satisfaction may be related to the chatbot's personality, emotional responsiveness, and empathy [101,102], which we will address in more detail later.

5.2. Effects on user engagement

Given that the lack of accountability in unsupported DMHIs is often cited as a drawback [39], one important benefit to chatbots is increased accountability, which leads to increased user engagement [103,104] and behavioral intentions to use the program [105]. For example, college students who completed web-based depression interventions with guidance from chatbots reported liking the accountability from the daily check-ins [85,98]. In another study, participants using a mobile intervention perceived the guidance and direction offered by the chatbot in the program as beneficial [61]. Thus, in addition to prompting users to engage with the intervention regularly, chatbots also help users engage with the material more deeply, which may be more important in reducing depressive symptoms than simply increasing the number of logins or activities completed [35].

However, while researchers have argued that chatbots help to increase user engagement and reduce dropout rates within DMHIs, relatively few studies have directly explored the impact of chatbots on attrition or user engagement. One study of a digital smoking cessation program found that the addition of a chatbot improved user engagement by 107% compared to their traditional program [104], but most other studies of chatbots – within DMHIs or otherwise – have tested the effects of a chatbot on engagement without a comparison to a non-chatbot-enabled DMHIs. And, to our knowledge, no studies have explored the long-term impact of chatbots on engagement or dropout rates.

5.3. Effects on mental health outcomes

Research on chatbots in DMHIs is in its infancy, and consequently, most studies have focused on the acceptability and usability of chatbots. Fewer studies have conducted rigorous tests of whether chatbots lead to improvements in mental health outcomes, and most research conducted in this domain consists of single-arm studies with no control groups, or studies with control groups that may not be adequate placebo or attention-controls. This is particularly important given that some researchers have argued that individual features of DMHIs, such as the availability of chatbots, may be associated with a greater digital placebo effect [106]. For example, although one meta-analysis found chatbots were effective for improving depression, distress, stress, and acrophobia, their evidence was considered weak due to the lack of studies, conflicting results across studies, and high estimated risk of bias in the included studies [107]. Thus, the evidence for chatbots in this domain should be considered preliminary.

In one study, researchers found that *Woebot* users had significantly greater improvements in depressive symptoms compared to participants in the control group, who read an e-book with psychoeducational content [85]. Other research has shown improvements in psychological well-being and perceived stress after two weeks, relative to a waitlist control [61]. Preliminary results also suggest significant improvements in substance use among *Woebot* users, though this study did not include a control group [60].

In regards to anxiety, results are more conflicting [107]. One study of first-year college students engaging with a healthy coping intervention delivered by a chatbot named *Atena* found significant improvements in anxiety, but only among participants with extreme anxiety scores at baseline [108]. This study also lacked a control group, and so these improvements may be a result of regression to the mean rather than engaging with *Atena*. Another study found that college students who engaged with their chatbot *Tess* reported significant improvements in anxiety symptoms, but this improvement was not significantly greater than that reported by participants in the control group, who read a psychoeducational book on depression [109].

Improvements in mental health outcomes may also depend on engagement. For instance, one study showed that users who had high engagement levels had significantly greater improvement in depressive symptoms relative to those with lower levels of engagement [90]. Research on the chatbot designed for young people treated for cancer, *Vivibot*, also found trends suggesting that greater usage led to greater improvements in anxiety [67]. Thus, much like non-AI-supported DMHIs, the effectiveness of mental health chatbots depends on the user's level of engagement, but when used as recommended, they appear to be effective.

5.4. Weaknesses of AI chatbots

Not to be mistaken for a utopian solution to existing problems with DMHIs, research points to several weaknesses in mental health chatbots. For instance, given the complexity of human language, chatbots are vulnerable to misinterpreting meaning in user responses [110]. Chatbots are not yet proficient in interpreting ellipses, metaphors, colloquialisms, and hyperbole, and these challenges will be even greater when chatbots are programmed in various languages [111]. Across several studies, a common complaint from users is that interactions with chatbots become repetitive [61,85,98,102], which makes the chatbot feel less human-like [61] and reduces users' motivation to continue the program [110]. Users also complain about misunderstandings with chatbots [61,85,90], particularly with long or complex messages that may not be understood by chatbots, leading to irrelevant or inappropriate responses which subsequently undermine the therapeutic alliance [100]. Thus, while chatbots can be beneficial, there are several ways in which they can deter users from engaging with the intervention. The risks associated with misunderstandings are even greater when chatbots or other types of AI are being used for diagnostic purposes.

6. A case study: happify health's AI chatbot Anna™

In 2019, Happify Health developed a chatbot to integrate into its commercial digital mental health platform. *Anna* is an AI-based chatbot that models the role of a therapist and delivers some Happify activities. Whereas some other DMHIs chatbots exist as a standalone app (i.e. the chatbot is the intervention), Anna was designed to complement other features of the Happify program. Specifically, Happify was developed to deliver gamified versions of evidence-based activities drawn from various therapeutic approaches, including CBT, mindfulness-based stress reduction, and positive psychology. These activities are organized into tracks designed to help users focus on a particular area of concern, like reducing stress (see Carpenter et al., 2016, for a discussion) [112].

Anna is incorporated into these tracks. Upon selecting an *Anna*-enabled track, *Anna* begins by greeting users with an introduction and explanation of the chatbot's role within the track (see Figure 1). In addition, *Anna* may ask users specific questions to gather information that can be used to better

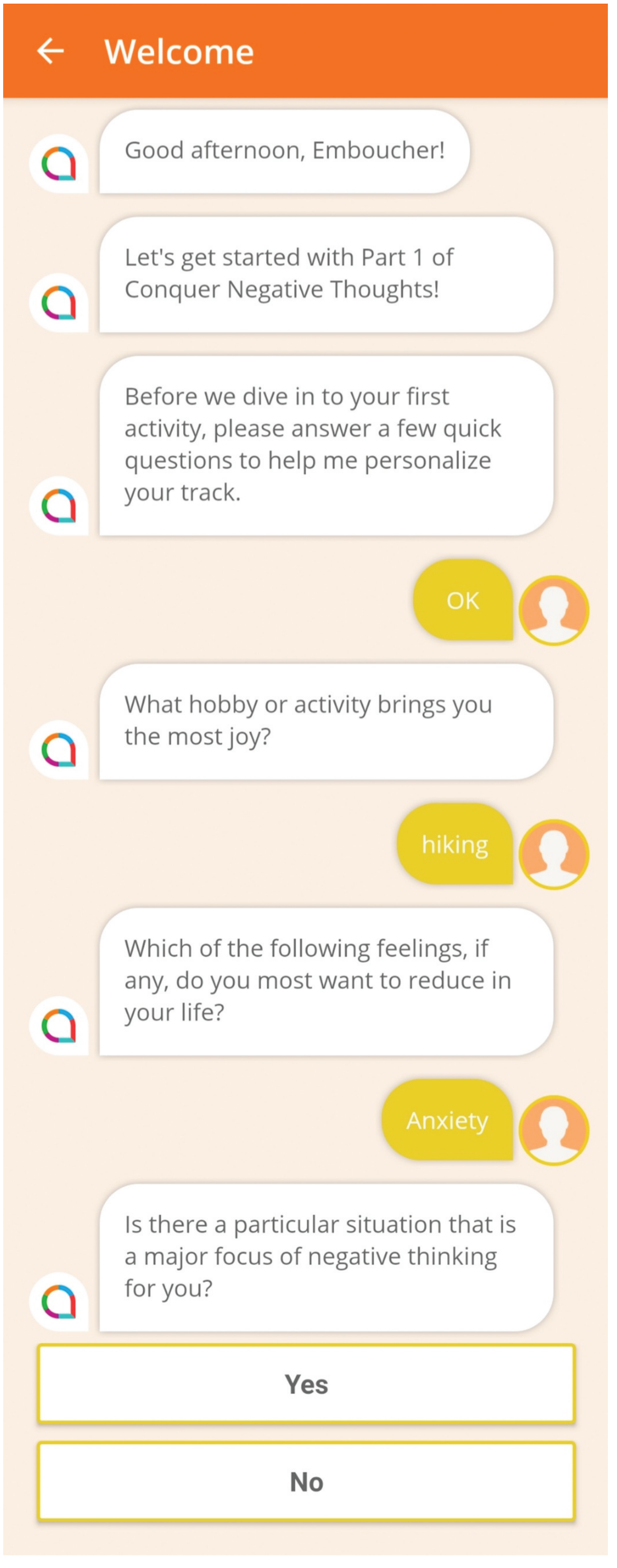

Figure 1. Screenshot of happify health's AI-based chatbot Anna's introductory dialog.

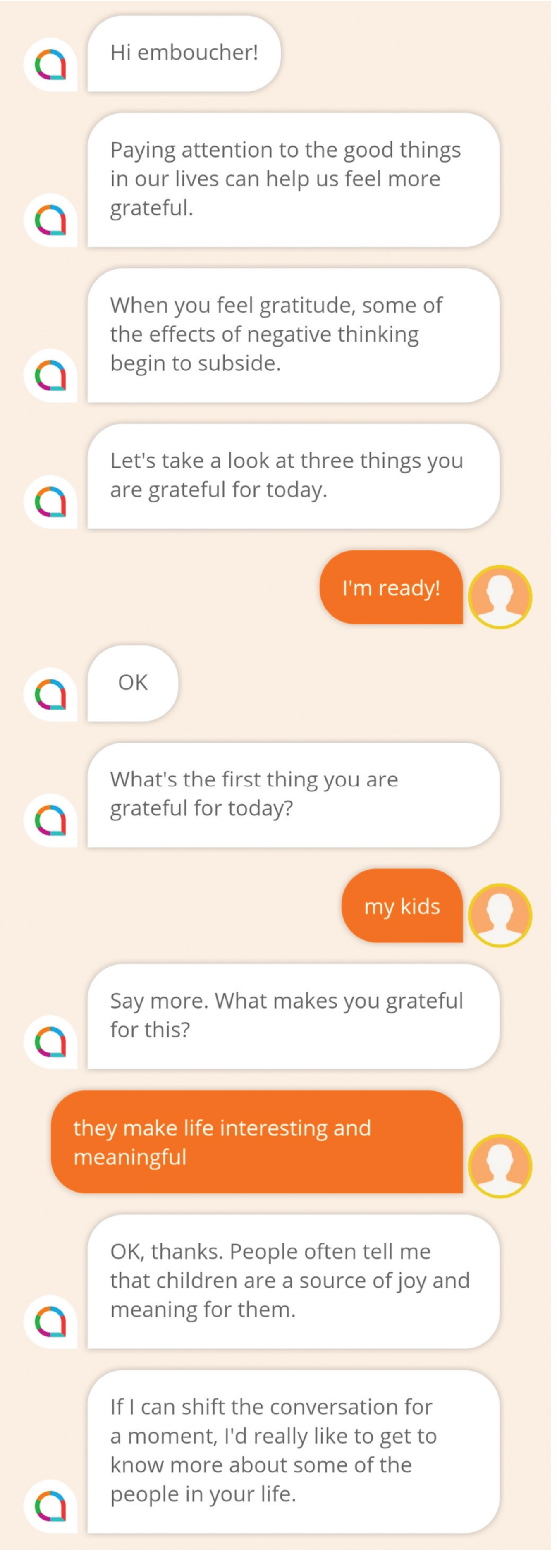

Figure 2. Screenshot of happify activity delivered by Anna.

personalize the track. *Anna* also delivers some activities within these tracks, deepening engagement with those activities and increasing the likelihood that users complete activities as intended (see Figure 2). Thus, *Anna* helps to maximize the benefits associated with that activity. To do this, *Anna* tracks criteria essential to optimal outcomes for each activity, reviews responses based on these criteria, and then prompts users to provide any expected information that was lacking from their initial response. For example, in a gratitude activity, *Anna* will coach users to express both positive emotion and meaning if these are not expressed in their initial response. While *Anna* coaches users to engage more efficiently with the platform, *Anna* also listens to the user; consequently, the course of the conversation may be steered by users and by the chatbot, addressing potential concerns that conversations with chatbots are not interactive enough [98].

Dialog content was generated by a team of clinicians and writers to ensure Anna's interaction with users effectively models therapists. Addressing previous concerns with preset responses and repetitiveness in chatbots within DMHIs [100], *Anna* employs a mix of instruction and feedback, and includes both open-ended and multiple-choice follow-up questions, using NLP models to interpret free responses. Anna's NLP models are trained by Happify Health data scientists, and labeled by a group of clinicians to ensure the interpretation of each dialog turn within the conversation is conducted from a clinical perspective.

Given that previous research revealed users view chatbots as less human-like [61], social dialogue features like empathy and meta-relational communication, which improve the working alliance, were integrated into *Anna*, following recommended best practices for AI in mental health care [102,113]. For instance, to help build rapport, *Anna* refers to users by name in conversations, and communicates curiosity about users by prompting them for background information and relationships with others when such information is mentioned in the normal course of conversation. Anna further communicates interest and understanding by using clarifying examples and content-mirroring responses to follow up user input in conversations. At the beginning of each activity, Anna also initiates a greeting dialog, which may include surface-level information (e.g. noting how often the user logs on) or reference more personal information acquired in past conversations.

The information gathered during these conversations is stored confidentially along with general user data (e.g. in-app assessment scores, activity across the platform) to form a model of the individual user. *Anna* can then reference this user model to create a more personalized experience, both relationally (e.g. celebrating an achievement) and by tailoring activities to users (e.g. recommending users perform a favorite activity with a specific person in their life, rather than suggesting they engage in a generic activity with someone).

6.1. Preliminary evidence for Anna

Although research on *Anna* is in the preliminary stages, a pilot test of how users perceived *Anna* following an interaction showed that 89.6% of the 203 surveyed users rated *Anna* as helpful. When these users were asked to rate *Anna* on a series of attributes, 74.9%, 73.3%, and 76.8% of users selected 'agree' or 'strongly agree' on the statements '*Anna listens to me*,' '*Anna is curious about me*,' and '*Anna gives me insights that I can use*,' respectively. Thus, consistent with other chatbot research, users report positive perceptions of *Anna*.

In another pilot study we conducted to explore how the addition of *Anna* within Happify activities would influence participants' engagement with those activities, we found that participants who received versions of Happify activities led by *Anna* provided more elaborate responses than those who received the regular Happify activities. We compared participants' text responses within activities over the course of four weeks based on whether they were exposed to regular Happify activities (n = 252) or those enhanced by the addition of *Anna* (n = 237). To do so, we gathered and cleaned all text data (i.e. converted contractions to separate words, removed punctuation, changed uppercase letters to lowercase, and transformed future and past tense words to present tense), and then compared overall word and character counts in participants' written responses across conditions. Participants in the Happify+*Anna* condition used significantly more words (M = 27.25, SD = 25.63) and more characters (M = 156.29, SD = 151.23) per entry overall compared to participants in the regular Happify condition (M = 14.93, SD = 14.18, and M = 83.93, SD = 82.99, respectively; ps < .001).

While writing more overall does not necessarily suggest participants were engaging more deeply with the activities, we also ran text analysis to identify what topics participants wrote most about, collapsing across the two conditions. To reduce overlap in topics, we identified the four most common themes (note, however, that there is still overlap in words across themes). The first theme included *positive relational* words like love, sure, and like. The second theme included *achievement-related* words like work, think, and feel. The third theme included *mindfulness-related* words like know, savor, and walk. And the fourth theme included *cognition-related* words like think, feel, and negative. We then compared the frequency of words from each theme in participants' responses across conditions. We found no significant difference in frequency of positive relational words, χ^2 = 0.06, p = .950. However, participants in the Happify condition used significantly more achievement-related words than those in the Happify+*Anna* condition, χ^2 = 237.18, p < .0001, but significantly fewer mindfulness-related, $\chi 2$ = 13.76, p < .001, and cognition-related, $\chi 2$ = 133.32, p < .0001, words. Thus, participants with the AI-enhanced activities were not only more elaborative, but also used more words directly relevant to the tasks Happify asks them to do, such as mindfulness and cognition. While promising, additional research considering the context of the response, rather than measuring the frequency of words, is necessary to better understand how users' responses differ when guided by a chatbot.

7. Open questions & important considerations

Despite the rapid increase of AI in DMHIs in recent years, chatbot technology remains relatively novel and experimental [52], and research on mental health chatbots is limited. Consequently, there are several important questions in regards to the use of AI-based chatbots in DMHIs, particularly as they expand to perform more functions.

7.1. Perceptions of AI in mental health care

Although research suggests that users rate chatbots favorably [98], few studies have examined the perceptions of chatbots in the general population. Consequently, it is unclear whether the integration of chatbots may encourage or inhibit uptake of DMHIs. Examining beliefs about chatbots among people who are not already using DMHIs would provide critical insight into the barriers to engaging with chatbots in DMHIs, stigmas associated with the use of AI for mental health care, and perceptions of security and privacy. Similarly, few studies have explored perceptions of mental health chatbots among mental health professionals. While some research suggests that healthcare professionals generally view chatbots favorably [80,114], other research found that 48.7% of surveyed psychiatrists from 22 countries reported that AI would only have a minimal impact on their work in the future [115]. Furthermore, personal experience with chatbots among these professionals is low and predicts whether they recommend these programs to their patients [80]. Thus, acceptance from healthcare practitioners is an important step to greater uptake of chatbot-driven DMHIs. Beyond assessing perceptions of AI, developers of mental health chatbots need to consider this task a collaboration between themselves, mental health professionals, and users/patients in order to develop chatbots that meet patients' needs, goals and lifestyles, ensure trust in AI, and improve mental health outcomes. Consequently, the development of chatbots should be viewed as an iterative process involving regular feedback from practitioners and users alike.

7.2. Empathy in chatbots

One criticism of incorporating chatbots into mental health care is that, by definition, they cannot feel empathy as human beings do [116–118]. Indeed, while psychiatrists appear to believe that AI may be capable of performing some psychiatric tasks (e.g. documentation), a global survey of psychiatrists found that 83% of respondents felt AI would never be able to match a psychiatrist in terms of empathetic care [115].

However, the extent to which chatbots are actually empathetic may be less relevant, and what likely matters most is whether patients *perceive* chatbots to be less empathetic than humans or clinicians [119]. Qualitative research suggests users perceive chatbots, like *Woebot*, as communicating empathy and emotional support [120], and other research has shown that interacting with a chatbot following ostracism can buffer against the negative effects of social exclusion [120]. Research also supports that chatbots designed to display empathic reactions are rated more positively (i.e. more enjoyable, understanding, sociable, trustworthy, intelligent) than one that is not programmed to respond empathetically [120]. Thus, users may perceive chatbots as empathetic, though the extent to which perceptions of empathy differ between chatbots and people or clinicians remains unclear. Further, although preliminary research suggests users can form a therapeutic relationship with chatbots [121], literature on how users perceive social attributes, such as personality and credibility, of chatbots is lacking. In addition to understanding general perceptions of chatbots, it is crucial to understand the impressions users form of mental health chatbots during interactions and the extent to which these compare to impressions of human practitioners.

7.3. Individual differences

The use of AI has great potential to expand and improve DMHIs [114,122]. However, individual differences, such as demographic characteristics or medical history (clinical vs. non-clinical populations), may impact how people respond to chatbots, yet this area of research remains underexplored. For example, research suggests that young people are hesitant to access mental health care due to perceived stigma [123]. Other research indicates adolescents prefer online conversations to in-person interactions for managing difficult conversations [124]. Conceivably, then, younger people may respond differently to the use of chatbots in mental health care. There is also a dearth of information regarding how race and ethnicity [7], gender, SES, and other factors may play a role in the responses to and outcomes associated with AI-supported DMHIs.

In addition to demographic variables, individual differences may also influence how people respond to chatbots in DMHIs. For example, in the pilot study we described earlier exploring the effects of integrating the chatbot *Anna* to Happify activities, we found that although the proportion of participants who dropped out of the study did not differ based on whether they had access to the chatbot or not (p = .943), the effect of mental health self-efficacy on dropout did differ based on condition. Participants who dropped out in the regular Happify condition had significantly lower baseline levels of mental health self-efficacy than those who completed the four-week study (p < .001), whereas mental health self-efficacy did not predict dropout in the Happify+*Anna* condition (p = .192). This finding might suggest that, in the absence of any support, people who are less confident in their ability to manage their mental health are at a higher risk of terminating the intervention prematurely. However, chatbots, like *Anna*, may provide sufficient support to mitigate these effects, allowing people with lower levels of mental health self-efficacy to feel more confident in their ability to manage their mental health through the intervention. This effect is preliminary and should be interpreted with caution, but is notable given that, compared to people with greater self-efficacy, these individuals may benefit more from DMHIs in terms of depression, anxiety, and overall distress [125]. In other words, the addition of an AI coach may allow people who are in greater need of an intervention to follow through with that intervention. This highlights the importance of exploring the role of other individual differences to better understand when, and for whom, chatbots are most beneficial. Without consideration of demographic and individual differences, the use of chatbots in DMHIs remains limited and introduces concerns regarding risks, safety, and effectiveness [126,127].

7.4. Ethical considerations

Considering the importance of trust for therapeutic relationships [128], garnering users' trust is an important consideration in DMHIs [100,129]. Maintaining trust requires a discussion of ethics, privacy, confidentiality, and safety [130]. The highly personal and sensitive nature of mental health information highlights the need to ensure that sensitive patient data is adequately secured and protected. Further, AI in DMHIs introduces additional risk for potential harms, including racial prejudice due to the potential for algorithmic bias [131], crisis response limitations [132], and safety concerns [132]. In some cases, for example, DMHIs have lacked evidence to support their claims regarding improvements in mental health [132], provided inaccurate health education [134], failed to spot signs of sexual abuse [133, 134], and encouraged unsafe behavior [135,136]. Therefore, challenges with respect to regulating AI in DMHIs for safety and effectiveness are a particular concern for the relatively novel field of digital mental health [137].

Given the dynamic and iterative nature of AI-supported DMHIs, standards for evaluation are especially challenging [138]. The existing literature on mental health chatbots includes many calls for considerations of AI-specific ethics [129], but definitions of ethical guidelines have yet to be established and widely adopted. Furthermore, research suggests that ethical and regulatory concerns associated with AI are not often considered by psychiatrists when evaluating the role of AI in mental health [136]. As the FDA and American Psychiatric Association [139,140] issue guidance on these critical topics, there exists an opportunity for those applying AI in mental health care to incorporate ethical, safety, and efficacy standards and values into DMHIs.

8. Conclusion

The availability of effective AI-supported interventions is an important avenue to reduce the longstanding burden on practitioners and improve the increasing shortage of mental health professionals [8]. Although preliminary research suggests chatbots are perceived favorably and may help to improve engagement and mental health outcomes, more rigorous tests of chatbots within DMHIs are needed. In particular, more research on how chatbots may help to improve mental health outcomes compared to other digital interventions without chatbots is an important next step, as is considering how individual and contextual factors might influence the impact of mental health chatbots.

9. Expert opinion: AI as the future of digital mental health

To be truly scalable, while optimizing engagement and adherence, digital health solutions must embrace AI, including chatbots, as a means of offering support. Indeed, in one survey of mental health professionals, the majority of respondents indicated healthcare chatbots would play a more significant role than healthcare providers in the future [80]. While skeptics remain, the integration of AI in DMHIs inevitably expands the possibilities of what can be done within digital health. Recent advances in AI technology, paired with the evolution of digital healthcare systems, are making it possible to take digital healthcare to the next level by making it dynamic, transparent, hyper-personalized, and human-like. This will result in levels of engagement and efficacy that are not yet possible today.

As technology improves and people become more digitally connected, the streams of data available for AI algorithms to analyze are no longer limited to isolated tests and care sessions. Instead, they feed on ongoing streams of data from multiple sources, particularly as the use of wearables and sensors expands. This data can be integrated in real-time to determine a patient's current state, detect deviations from usual patterns, and make probabilistic predictions of an individual's future health state, thereby enabling digital mental health systems to respond to events as they take place (or even before they take place).

Such integration of various data sources will permit a level of personalization and responsiveness that is impossible from human beings alone. Unlike human practitioners, AI algorithms can quickly analyze large amounts of user data to understand and respond to users. For instance, chatbots can predict personality based on a user's language and subsequently adapt their own personality to match the user's. Much like mimicry in face-to-face interactions, matching chatbot personalities to user personalities leads to more productive interactions [141].

AI algorithms also have access to vast amounts of normative data that can be used to train accurate models that infer a patient's current state, and these models can be revised as new data becomes available. This could result in fine-tuned, responsive models of care that adapt quickly based on where the patient is on their care journey, what is happening in their lives, and even how their day is going.

Conceivably, while chatbots may never be truly empathetic, access to this normative data along with recent advances in generative deep learning will allow AI-based chatbots to interface with patients in a way that communicates empathy. As we learn from existing chatbots and user feedback, chatbots will become more effective at delivering interactions that are conversational and natural. More importantly, they will become more sensitive to patient states and able to express the appropriate emotions and actions during these interactions. Chatbots will exhibit empathy, curiosity, understanding, and a sense of working collaboratively with patients rather than being one-sided and authoritative, resulting in deeper intrinsic patient motivation to engage and adhere. As such, true therapeutic alliances between users and chatbots will be possible.

To reach these goals, we need to learn more about the existing strengths and limitations of AI-based chatbots in DMHIs. That requires more widespread adoption of chatbots within DHMI. However, numerous barriers exist to widespread adoption. One important barrier is that there are many misconceptions about the intended role of chatbots in DMHIs. Although critics argue chatbots cannot replace human interaction/therapy [142], most creators of chatbots in DMHIss never intended for their chatbots to replace human therapists. Rather, chatbots were primarily developed to help increase engagement, support the prevention of mental disorders by delivering more engaging and adaptive interventions [30,143], and to re-engage users they identify at risk of dropout. Research also suggests that people may self-disclose more to a mental health practitioner when it is facilitated via self-disclosing chatbots [144]. Thus, chatbots can be effective adjunctive options and may help make in-person therapy more effective. In other words, in its best form, AI algorithms would not replace human intelligence, but augment it (hence why some use the term 'augmented intelligence' over 'artificial intelligence'). Future research exploring the extent to which a combination of AI chatbots, DMHIs, and clinician support compared to DMHIs or clinician support alone will be an important step to understanding how AI chatbots can augment both DMHIs as well as traditional clinician-led therapy.

An important step to reducing these misconceptions is to focus more on working with and educating mental health professionals about AI in mental health settings. While practitioners generally agree that mental health chatbots are beneficial and important, they also express trepidation about using AI in domains like diagnostics, counseling, and delivering CBT [80,114]. Personal experience with mental health chatbots also tends to be quite low, although practitioners with personal experience are more likely to recommend chatbot-driven DMHIs to their patients. Consequently, ensuring the digital therapeutic industry is working hand-in-hand with healthcare professionals as we expand the use of AI in digital interventions, while also encouraging healthcare professionals to engage with these interventions themselves, will be crucial to securing support from the broader healthcare community.

Part of what contributes to the misconceptions is the low levels of adoption of chatbots in large-scale DMHIs. As few as 24% of the top-funded DMHIs companies include a conversational agent [70], so few people (or practitioners) have interacted with chatbots in this particular domain. This is compounded by the fact that research on mental health chatbots is limited. Most studies conducted thus far do not include a control group or compare the effects of interacting with a chatbot to a very different approach, like reading an e-book with psychoeducational content. Few, if any, studies have used an adequate sham condition as the control group and, as a result, the extent to which chatbots enhance the effects of DMHIs is unclear. This data will be important to clarify what AI chatbots can and cannot do in this context, and reduce misconceptions about AI. There is a pressing need for more rigorous research on AI in digital therapeutics, and more transparency and communication among AI developers in this domain. Ultimately, the extent to which the digital therapeutics industry engages in more widespread adoption of AI within their interventions, that they engage in a patient-centered approach to developing and referring chatbots, and that they commit to conducting research using best practices will determine the future of AI in DMHIs.

Funding

This paper was not funded, and the pilot study referred to herein was sponsored by Happify Health.

Declaration of interest

All authors of this manuscript are current employees of Happify Health. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

ORCID

Eliane M. Boucher http://orcid.org/0000-0002-1384-7177
Nicole R. Harake http://orcid.org/0000-0003-3750-8185
Haley E. Ward http://orcid.org/0000-0001-5423-8928
Sarah Elizabeth Stoeckl http://orcid.org/0000-0003-2719-2957
Junielly Vargas http://orcid.org/0000-0002-1806-7428
Jared Minkel http://orcid.org/0000-0001-7979-3098
Acacia C. Parks http://orcid.org/0000-0001-6643-0116
Ran Zilca http://orcid.org/0000-0001-8521-5438

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Merikangas KR, He JP, Burstein M, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry. 2010;49(10):980–989. DOI:10.1016/j.jaac.2010.05.017.
2. Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: results from the 2019 national survey on drug use and health. Rockville, MD: SAMHSA. 2020.
3. NIMH-mental illness [Internet]. Bethesda MD; NIMH. [cited 2021 Aug 26]. Available from: https://www.nimh.nih.gov/health/statistics/mental-illness#part_2555
4. Wang PS, Lane M, Olfson M, et al. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62(6):629–640. DOI:10.1001/archpsyc.62.6.629.
5. Wang PS, Aguilar-Gaxiola S, Alonso J, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet. 2007;370(9590):841–850. DOI:10.1016/S0140-6736(07)61414-7.
6. Mojtabai R, Olfson M, Sampson NA, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. Psychol Med. 2011;41(8):1751–1761. DOI:10.1017/S0033291710002291.
7. Conner KO, Copeland VC, Grote NK, et al. Mental health treatment seeking among older adults with depression: the impact of stigma and race. Am J Geriatr Psychiatry. 2010;18(6):531–543. DOI:10.1097/JGP.0b013e3181cc0366.
8. Thomas KC, Ellis AR, Konrad TR, et al. County-level estimates of mental health professional shortage in the United States. Psychiatr Serv. 2009;60(10):1323–1328. DOI:10.1176/ps.2009.60.10.1323.
9. Cunningham PJ. Beyond parity: primary care physicians' perspectives on access to mental health care. Health Aff (Millwood). 2009;28(3):w490–w501.
10. Mental Health and Substance Use. The impact of COVID-19 on mental, neurological and substance use services. World Health Organization. 2020.
11. CDC. Mental health household pulse survey COVID-19. CDC. 2021.
12. Hamel L, Kearney A, Kirzinger A, et al. . KFF. 2020. https://www.kff.org/coronavirus-covid-19/report/kff-health-tracking-poll-july-2020/
13. Czeisler MÉ, Lane RI, Petrosky E, et al. Mental health, substance use, and suicidal ideation during the covid-19 pandemic — United States. MMWR Morb Mortal Wkly Rep. 2020 June 24–30;69(32):1049–1057. DOI:10.15585/mmwr.mm6932a1.
14. Ettman CK, Abdalla SM, Cohen GH, et al. Prevalence of depression symptoms in us adults before and during the covid-19 pandemic. JAMA Network Open. 2020;3(9):e2019686. DOI:10.1001/jamanetworkopen.2020.19686.
15. Nagata JM, Ganson KT, Bonin SL, et al. Prevalence and sociodemographic correlates of unmet need for mental health counseling among adults during the covid-19 pandemic [published online ahead of print, 2021 jun 30]. Psychiatr Serv. 2021;appips202100111. DOI:10.1176/appi.ps.202100111.
16. Kaiser family foundation [Internet]. Unmet Need for Counseling or Therapy Among Adults Reporting Symptoms of Anxiety and/or Depressive Disorder During the COVID-19 Pandemic. State Health Facts. 2021. [cited 6 August 2021]. Available from: https://www.kff.org/other/state-indicator/unmet-need-for-counseling-or-therapy-among-adults-reporting-symptoms-of-anxiety-and-or-depressive-disorder-during-the-covid-19-pandemic/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D
17. Perrin A. Mobile technology and home broadband 2021 [Internet]. Research Topics: [Internet & Technology]. Washington D.C.: Pew Research Center; 2021 [2021 Aug 26]. Available from: https://www.pewresearch.org/internet/2021/06/03/mobile-technology-and-home-broadband-2021/
18. Anthes E. Mental health: there's an app for that. Nature. 2016;532(7597):20–23.
19. Renton T, Tang H, Ennis N, et al. Web-based intervention programs for depression: a scoping review and evaluation. J Med Internet Res. 2014;16(9):e209. DOI:10.2196/jmir.3147.
20. Ebert DD, Cuijpers P, Muñoz RF, et al. Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research. Front Psychiatry. 2017;8:116.
21. Schueller SM, Hunter JF, Figueroa C, et al. Use of digital mental health for marginalized and underserved populations. Curr Treat Options Psych. 2019;6(3):243–255. DOI:10.1007/s40501-019-00181-z
   - **Review of feasibility and acceptability of digital mental health interventions for underserved populations.**
22. Kadakia K, Patel B, Shah A. Advancing digital health: FDA innovation during COVID-19. NPJ Digit Med. 2020;3(1):161. DOI:10.1038/s41746-020-00371-7.
23. Enforcement policy for digital health devices for treating psychiatric disorders during the coronavirus disease 2019 (covid-19) public health emergency guidance for industry and food and drug administration staff. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health (CDRH), Office of Product Evaluation and Quality (OPEQ). Rockville, MD: FDA; 2020.
24. Parks AC, Della Porta MD, Pierce RS, et al. Pursuing happiness in everyday life: the characteristics and behaviors of online happiness seekers. Emotion. 2012;12(6):1222–1234. DOI:10.1037/a0028587.
25. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. JAMA Psychiatry. 2017;74(5):437–438.
26. Lau N, O'Daffer A, Colt S, et al. Android and iPhone mobile apps for psychosocial wellness and stress management: systematic search

in app stores and literature review. JMIR Mhealth Uhealth. 2020;8 (5):e17798. DOI:10.2196/17798.

27. Firth J, Torous J, Nicholas J, et al. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord. 2017;218:15–22.
  - **This meta-analysis suggests that psychological interventions delivered using smartphone devices can reduce anxiety.**
28. Firth J, Torous J, Nicholas J, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry. 2017;16(3):287–298. DOI:10.1002/wps.20472.
  - **This meta-analysis suggests that psychological interventions delivered via smartphone have small to medium effects on depression.**
29. Linardon J, Cuijpers P, Carlbring P, et al. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry. 2019;18(3):325–336. DOI:10.1002/wps.20673.
30. Sander L, Rausch L, Baumeister H. Effectiveness of internet-based interventions for the prevention of mental disorders: a systematic review and meta-analysis [published correction appears in jmir ment health. 2016;3(3):e41]. JMIR Ment Health. 2016;3(3):e38.
31. Wu A, Scult MA, Barnes ED, et al. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. NPJ Digit Med. 2021;4(1):20. DOI:10.1038/s41746-021-00386-8
32. Weisel KK, Fuhrmann LM, Berking M, et al. Standalone smartphone apps for mental health—a systematic review and meta-analysis. NPJ Digit Med. 2019;2(1):118. DOI:10.1038/s41746-019-0188-8.
33. Lattie EG, Adkins EC, Winquist N, et al. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res. 2019;21(7):e12869. DOI:10.2196/12869.
34. Rudd BN, Beidas RS. Digital mental health: the answer to the global mental health crisis?. JMIR Ment Health. 2020;7(6):e18472.
35. Donkin L, Christensen H, Naismith SL, et al. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res. 2011;13(3):e52. DOI:10.2196/jmir.1772.
36. Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med. 2008;31(2):169–177.
37. Eysenbach G. The law of attrition. J Med Internet Res. 2005;7(1):e11.
  •• **Cornerstone article discussing the characteristics and methodological challenges in evaluating eHealth applications, urging the need for a science of attrition for measuring digital health interventions.**
38. Andersson G. Using the Internet to provide cognitive behaviour therapy. Behav Res Ther. 2009;47(3):175–180.
39. Gerhards SA, Abma TA, Arntz A, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. J Affect Disord. 2011;129(1–3):117–125. DOI:10.1016/j.jad.2010.09.012.
40. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res. 2021;23(3):e24387. DOI:10.2196/24387.
41. Ferwerda M, van Beugen S, van Burik A, et al. What patients think about E-health: patients' perspective on internet-based cognitive behavioral treatment for patients with rheumatoid arthritis and psoriasis. Clin Rheumatol. 2013;32(6):869–873. DOI:10.1007/s10067-013-2175-9.
42. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev. 2012;32(4):329–342.
43. Mohr DC, Duffecy J, Ho J, et al. A randomized controlled trial evaluating a manualized telecoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One. 2013;8(8):e70086. DOI:10.1371/journal.pone.0070086.
44. Morris RR, Schueller SM, Picard RW. Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. J Med Internet Res. 2015;17(3):e72.
45. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther. 2009;38(4):196–205.
46. Hoermann S, McCabe KL, Milne DN, et al. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res. 2017;19(8):e267.
47. Wallin E, Norlund F, Olsson EMG, et al. Treatment activity, user satisfaction, and experienced usability of internet-based cognitive behavioral therapy for adults with depression and anxiety after a myocardial infarction: mixed-methods study. J Med Internet Res. 2018;20(3):e87. DOI:10.2196/jmir.9690.
48. D'Alfonso S. AI in mental health. Curr Opin Psychol. 2020;36:112–117.
49. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology. 2016;41(7):1691–1696.
50. Abashev A, Grigoryev R, Grigorian K, et al. Programming tools for messenger-based chatbot system organization: implication for outpatient and translational medicines. BioNanoScience. 2017;7 (2):403–407. DOI:10.1007/s12668-016-0376-9.
51. Baker A, Perov Y, Middleton K, et al. A comparison of artificial intelligence and human doctors for the purpose of triage and diagnosis. Front Artif Intell. 2020;3:543405.
52. Bendig E, Erb B, Schulze-Thuesing L, et al. The next generation: chatbots in clinical psychology and psychotherapy to foster mental health – a scoping review. Verhaltenstherapie. 2019;1–13. DOI:10.1159/000501812.
  - **This review offers a thorough discussion of the opportunities, limits, risks, and challenges of chatbots in clinical psychology.**
53. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. Commun ACM. 1966;9(1):36–45.
  - **Cornerstone paper written by the developer of the first chatbot, ELIZA, describing the first natural language processing program.**
54. Almalki M, Azeez F. Health chatbots for fighting covid-19: a scoping review. Acta Inform Med. 2020;28(4):241–247.
55. Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. J Med Internet Res. 2020;22(10):e20346. DOI:10.2196/20346.
56. Abd-Alrazaq AA, Bewick BM, Farragher T, et al. Factors that affect the use of electronic personal health records among patients: a systematic review. Int J Med Inform. 2019;126:164–175.
57. Ahmed A, Ali N, Aziz S, et al. A review of mobile chatbot apps for anxiety and depression and their self-care features. Comput Meth Prog Bio. 2021;1:1–8.
  - **This review reports the lack of evidence and need for further study of chatbot effectiveness and safety.**
58. Sana M, Muhammad HJ, Tayyaba A. Aquabot: a diagnostic chatbot for achluophobia and autism. IJACSA. 2017;8:9.
59. Wibhowo C, Sanjaya R. Virtual assistant to suicide prevention in individuals with borderline personality disorder. Kuching Malaysia: ICCOINS; 2021 July 13 15. p. 234–237.
60. Prochaska JJ, Vogel EA, Chieng A, et al. A therapeutic relational agent for reducing problematic substance use (woebot): development and usability study. J Med Internet Res. 2021;23(3):e24850. DOI:10.2196/24850.
61. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Interv. 2017;10:39–46.
62. Ujiro T, Tanaka H, Adachi H, et al. Detection of dementia from responses to atypical questions asked by embodied conversational agents. Proc Interspeech. 2018;2018:1691–1695.
63. Kidder GW, Montgomery CW. Oxygenation of frog gastric mucosa in vitro. Am J Physiol Cell Physiol. 1975;229(6):1510–1513.
64. Lee M, Ackermans S, As NV, et al. Caring for Vincent: a chatbot for self-compassion; Proceedings of the 2019 CHI Conference on

Human Factors in Computing Systems; 2019 May 4-9; Scotland, UK: Glasgow.
65. Valtolina S, Charlie: HL. A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness. In: CHItaly 2021: 14th Biannual Conference of the Italian Sigchi Chapter. New York (NY): Association for Computing Machinery; 2021. p. 1–5.
66. Wargnier P, Benveniste S, Jouvelot P, et al. Usability assessment of interaction management support in LOUISE, an ECA-based user interface for elders with cognitive impairment. Technol Disabil. 2018;30(30):105–126. DOI:10.3233/TAD-180189.
67. Greer S, Ramo D, Chang YJ, et al. Use of the chatbot "vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. JMIR Mhealth Uhealth. 2019;7(10):e15018. DOI:10.2196/15018.
68. Ivanovic M, Semnic M. The role of agent technologies in personalized medicine. Proceedings of the 2018 5th International Conference on Systems and Informatics (ICSAI); Nanjing, China; 2018. p. 299–304.
69. Shum HY, He XD, Li D. From Eliza to XiaoIce: challenges and opportunities with social chatbots. Frontiers Inf Technol Electronic Eng. 2018;19(1):10–26.10.1631/FITEE.1700826.
70. Alattas A, Teepe G, Leidenberger K, et al. To what scale are conversational agents used by top-funded companies offering digital mental health services for depression?. Proceedings of the 14th International Joint Conference on BIOSTEC. 2021; 2021 Feb 5; HEALTHINF p. 801–808.
71. Mirheidari B. Detecting early signs of dementia in conversation [dissertation]. Sheffield England: University of Sheffield; 2018.
72. Auriacombe M, Moriceau S, Serre F, et al. Development and validation of a virtual agent to screen tobacco and alcohol use disorders. Drug Alcohol Depend. 2018;193:1–6.
73. Huang JQ, Li XY, Cheng TS , et al. Teenchat: achatterbot system for sensing and releasing adolescents' stress. International Conference on Health Information Science. Cham: Springer; 2015. 9085. p. 133–145.
74. Bickmore TW, Mitchell SE, Jack BW, et al. Response to a relational agent by hospital patients with depressive symptoms. Interact Comput. 2010;22(4):289–298. DOI:10.1016/j.intcom.2009.12.001.
75. Breso A, Martinez-Miranda J, Botella C, et al. Usability and acceptability assessment of an empathic virtual agent to prevent major depression. Expert Syst: Int J of Knowl Eng and Neural Networks. 2016;33(4):297–312. DOI:10.1111/exsy.12151.
76. DeVault D, Artstein R, Benn GT, et al. SimSensei kiosk: a virtual human interviewer for healthcare decision support. Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (AAMAS '14); Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems; 2014. p. 1061–1068.
77. Lucas GM, Rizzo AJ, Gratch S, et al. Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers, front. Robot. AI. 2017. p. 4.
78. Wortwein T, Scherer S. What really matters-an information gain analysis of questions and reactions in automated PTSD screenings. in Proceedings of the 2017 Seventh International Conference on ACII; USA; San Antonio, TX; 2017. p. 15–20.
79. Divya S, Indumathi V, Ishwarya S, et al. A self-diagnosis medical chatbot using artificial intelligence. WDJWD. 2018;3(1):1–7.
80. Sweeney C, Potts C, Ennis E, et al. Can chatbots help support a person's mental health?: perceptions and views from mental healthcare professionals and experts. ACM HEALTH. 2021;2(3): 1–16
81. Balaskas A, Schueller SM, Cox AL, et al. Ecological momentary interventions for mental health: a scoping review. PLoS One. 2021;16(3):e0248152.
82. Jungmann SM, Klan T, Kuhn S, et al. Accuracy of a chatbot (ada) in the diagnosis of mental disorders: comparative case study with lay and expert users. JMIR Form Res. 2019;3(4):e13863.
83. Hungerbuehler I, Daley K, Cavanagh K, et al. Chatbot-based assessment of employees' mental health: design process and pilot implementation. JMIR Form Res. 2021;5(4):e21678.
84. Lovejoy CA, Buch V, Maruthappu M. Technology and mental health: the role of artificial intelligence. Eur Psychiatry. 2019;55:1–3.
85. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. JMIR Ment Health. 2017;4(2):e19.
86. Abd-Alrazaq AA, Alajlani M, Alalwan AA, et al. An overview of the features of chatbots in mental health: a scoping review. Int J Med Inform. 2019;132:103978.
 - **This review of 41 unique chatbots provides a comprehensive overview of the characteristics of chatbots in mental health research, including: purpose, platform, response generation, dialogue initiative, input and output modalities, embodiment, and targeted disorder.**
87. Woebot [Internet]. Woebot Health. 2021 [2021 Aug 26]. Available from: https://woebothealth.com/
88. Montenegro JZ, Da Costa CA, Da Rosa Righi R. Survey of conversational agents in health. Expert Syst Appl. 2019;129:56–67.
89. Wysa [Internet]. Wysa; 2021 [2021 Aug 26]. Available from: https://www.wysa.io/
90. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth. 2018;6(11):e12106.
91. Gaffney H, Mansell W, Edwards R, et al. Manage your life online (MYLO): a pilot trial of a conversational computer-based intervention for problem solving in a student sample. Behav Cogn Psychother. 2014;42(6):731–746.
92. Abashev A, Grigoryev R, Grigorian K, et al. Programming tools for messenger-based chatbot system organization: implication for outpatient and translational medicines. BioNanoScience. 2017;7(2):403–407.
93. Cr D, Kj M, Rg C. A process evaluation examining the performance, adherence, and acceptability of a physical activity and diet artificial intelligence virtual health assistant. Int J Environ Res Public Health. 2020;17(23):9137.
94. Burton C, Szentagotai Tatar A, McKinstry B, et al. Pilot randomised controlled trial of help4mood, an embodied virtual agent-based system to support treatment of depression. J Telemed Telecare. 2016;22(6):348–355. DOI:10.1177/1357633X15609793.
95. Cameron G, Cameron D, Megaw RG, et al. McTear, assessing the usability of a chatbot for mental health care. Proceedings of the 5th International Conference on Internet Science, INSCI 2018; 2018 Oct 24-26; St. Petersburg: Russian Federation; St. Petersburg, Russia. INSCI Workshops; 2018. p 121–132.
96. Ghosh S, Bhatia S, Bhatia A. Quro: facilitating user symptom check using a personalised chatbot-oriented dialogue system. Stud Health Technol Inform. 2018;252:51–56.
97. Feijt MA, de Kort YA, Bongers IM, et al. Perceived drivers and barriers to the adoption of emental health by psychologists: the construction of the levels of adoption of emental health model. J Med Internet Res. 2018;20(4):e153.
98. Fulmer R, Joerin A, Gentile B, et al. Using psychological artificial intelligence (tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health. 2018;5(4):e64.
99. Radziwill NM, Benton MC. Evaluating quality of chatbots and intelligent conversational agents. arXiv. 2017 [2021 Aug 26]:21.
100. Kretzschmar K, Tyroll H, Pavarini G, et al. NeurOx young people's advisory group. 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. 2019;11:1178222619829083.
101. Yu Z, Papangelis A, Rudnicky A. TickTock: anon-goal-oriented multimodal dialog system with engagement awareness. AAAI Spring symposium series. Pittsburg: PA; 2015.
102. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns. 2005;59(1):21–30.
103. Bickmore TW, Pfeifer LM, Byron D, et al. Usability of conversational agents by patients with inadequate health literacy: evidence from

two clinical trials. J Health Comm. 2010;15(Suppl 2):197–210. DOI:10.1080/10810730.2010.499991.

104. Perski O, Crane D, Beard E, et al. Does the addition of a supportive chatbot promote user engagement with a smoking cessation app? an experimental study [published correction appears in digit health. 2020;6:2055207620930958]. Digit Health. 2019;30 (5):2055207619880676.
105. Kamita T, Ito T, Matsumoto A, et al. A chatbot system for mental healthcare based on SAT counseling method. Mob Inf Syst. 2019;9517321:1-9517321: 11.
106. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry. 2016;3(2):100–102.
107. Abd-Alrazaq AA, Rababeh A, Alajlani M, et al. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res. 2020;22(7):e16021.
108. Gabrielli S, Rizzi S, Bassi G, et al. Engagement and effectiveness of a healthy-coping intervention via chatbot for university students during the covid-19 pandemic: mixed methods proof-of-concept study. JMIR Mhealth Uhealth. 2021;9(5):e27965. DOI:10.2196/27965.
109. Klos MC, Escoredo M, Joerin A, et al. Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial. JMIR Form Res. 2021;5(8):e20678.
110. Adamopoulou E, Moussiades L. Chatbots: history, technology, and applications. ML with Appl. 2020; 2:100006. DOI:10.1016/j.mlwa.2020.100006.
•• **Comprehensive overview of the history of chatbots, including technical approaches, algorithms, and techniques.**
111. Neidlein A, Wiesenbach P, Markert K An analysis of language models for metaphor recognition. Proceedings of the 28th International Conference on Computational Linguistics 2020; 2020 December 8-13; Barcelona, Spain; p. 3722–3736.
112. Carpenter J, Crutchley P, Rd Z, et al. Seeing the “Big” picture: big data methods for exploring relationships between usage, language, and outcome in internet intervention data [published correction appears. J Med Internet Res. 2017Dec19; 19(12):e347.J Med Internet Res. 2016;18(8):e241. 10.2196/jmir.8099.
113. Cameron G, Cameron D, Megaw G, et al. Best practices for designing chatbots in mental healthcare–A case study on iHelpr. Proceedings of the 32nd International BCS HCI Conference 2018; 2018 July 4-6; Belfast, UK; p. 1–5.
114. Palanica A, Flaschner P, Thommandram A, et al. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. J Med Internet Res. 2019;21(4):e12887.
115. Doraiswamy PM, Blease C, Bodner K. Artificial intelligence and the future of psychiatry: insights from a global physician survey. Artif Intell Med. 2020;102:101753.
116. Montemayor C, Halpern J, Fairweather A. In principle obstacles for empathic AI: why we can't replace human empathy in healthcare. AI Soc. 2021; [2021 Aug 26]; 1–7. 10.1007/s00146-021-01230-z.
117. Fadhil A, Schiavo G. Designing for health chatbots. ArXiv. 2019; 2021 Aug 26; abs/1902.09022.
118. Lisetti C, Amini R, Yasavur U, et al. You change! an empathic virtual agent delivers behavior change health interventions. ACM Trans Inf Syst. 2013;4(19):1–28.10.1145/2544103.
119. Howick J, Morley J, Floridi L. An empathy imitation game: empathy turing test for care- and chat-bots. Minds Mach. 2021;31(3):457–461.
120. de Gennaro M, Krumhuber EG, Lucas G. Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. Front Psychol. 2020;10:3061.
121. Darcy A, Daniels J, Salinger D, et al. Evidence of human-level bonds established with a digital conversational agent: cross-sectional, retrospective observational study. JMIR Form Res. 2021;5(5):e27868.
122. Andersson G. Internet-delivered psychological treatments. Annu Rev Clin Psychol. 2016;12(1):157–179.
123. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry. 2010;10(1):113.
124. Dubow EF, Lovko KR Jr, Kausch DF. Demographic differences in adolescents' health concerns and perceptions of helping agents. J Clin Child Psychol. 1990;19(1):44–54.10.1207/s15374424jccp1901_6.
125. Clarke J, Proudfoot J, Birch MR, et al. Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: secondary analysis of a randomised controlled trial. BMC Psychiatry. 2014;14 (1):272. DOI:10.1186/s12888-014-0272-1.
126. Daniel F, Cappiello C, Benatallah B. Bots acting like humans: understanding and preventing harm. IEEE Internet Comput. 2019;23 (2):40–49.
127. De Filippis ML, Federici S, Mele ML, et al. Preliminary results of a systematic review: quality assessment of conversational agents (chatbots) for people with disabilities or special needs. In: International Conference on Computers Helping People with Special Needs. (p. 250–257). Cham: Springer). 2020.
128. Birkhäuer J, Gaab J, Kossowsky J, et al. Trust in the health care professional and health outcome: a meta-analysis. PLoS One. 2017;12(2):e0170988. DOI:10.1371/journal.pone.0170988.
129. Martinez-Martin N. Trusting the bot: addressing the ethical challenges of consumer digital mental health therapy. In: Bárd I, and Hildt E, editors. Developments in neuroethics and bioethics. Vol. 3. United Kingdom: Academic Press; 2020. p. 63–91.
•• **This article provides an overview of the ethical considerations and challenges in consumer digital mental health.**
130. Asan O. Contextual differences in the dynamic measurement of trust in web sites across domains. IIJCSE. 2012;5(2):91–110.
131. Towards trustable machine learning. Nat Biomed Eng. 2018 Oct;2 (10):709–710. DOI:10.1038/s41551-018-0315-x.
132. Larsen ME, Nicholas J, Christensen H, et al. Assessment of smartphone tools for suicide prevention. PLoS One. 2016;11(4): e0152285.
133. Akbar S, Coiera E, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. J Am Med Inform Assoc. 2020;27(2):330–340.
134. White G. Child advice chatbots fail to spot sexual abuse. BBC News [Internet]. 2018 Dec 11; [2021 Aug 26]. Technology. Available from: https://www.bbc.com/news/technology-46507900
135. Nicholas J, Larsen ME, Proudfoot J, et al. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res. 2015;17(8):e198.
136. Blease C, Locher C, Leon-Carlyle M, et al. Artificial intelligence and the future of psychiatry: qualitative findings from a global physician survey. Digit Health. 2020;6:205520762096835.
137. Torous J, Andersson G, Bertagnoli A, et al. Towards a consensus around standards for smartphone apps and digital mental health. World J Psychiatry. 2019;18(1):97–98. DOI:10.1002/wps.20592.
138. Price II, Nicholson W. Regulating black-box medicine. Mich L Rev. 2017;116(3):421.
139. Harvey HB, Gowda V. How the FDA regulates AI. Acad Radiol. 2020;27(1):58–61.
140. Gerke S, Babic B, Evgeniou T, et al. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. NPJ Digit Med. 2020;3(1):53. DOI:10.1038/s41746-020-0262-2.
141. Shumanov M, Johnson L. Making conversations with chatbots more personalized. Comput Hum Behav. 2021;117:106627. DOI:10.1016/j.chb.2020.106627.
142. Sachan D. Self-help robots drive blues away. Lancet Psychiatry. 2018;5(7):547.
143. Christensen H, Batterham P, Mackinnon A, et al. Prevention of generalized anxiety disorder using a web intervention, iChill: randomized controlled trial. J Med Internet Res. 2014;16(9):e199. DOI:10.2196/jmir.3507.
144. Lee YC, Yamashita N, Huang Y Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. Proceedings of the ACM on HCI; 2020. p. 1–27. https://dl.acm.org/doi/abs/10.1145/3392836